INTERNATIONAL JOURNAL OF PSYCHOLOGY, 1998, *33* (2), 107–122

Influence of Global Structure on Musical Target Detection and Recognition

Barbara Tillmann and Emmanuel Bigand

Laboratoire d'Etudes des Apprentissages et du Développement, Université de Bourgogne, France

The present study adapted a paradigm used in visual perception by Biederman, Glass, and Stacy (1973) and analyzed the influence of a coherent global context on the detection and recognition of musical target excerpts. Global coherence was modified by segmenting minuets into chunks of four, two, or one bar. These chunks were either reordered (Experiments 1, 3, 4, 5) or transposed to different keys (Experiment 2). The results indicate that target detection is influenced only by a reorganization on a very local level (i.e. chunks of one bar). Context incoherence did not influence the recognition of the real targets, but rendered the rejection of wrong target excerpts (foils) more difficult. The present findings revealed only a weak effect of global context on target identification and only for extremely modified structures.

Cette étude reprend un pardigme utilisé dans la perception visuelle par Biederman, Glass, et Stacy (1973) pour analyser l'influence du contexte global sur la détection et la reconnaissance de cibles musicales. Pour modifier la cohérence globale du contexte, des menuets sont segmentés en unités de quatre, deux ou une mésure. Ces groupes sont enchaînés dans des ordres differents (Expériences 1, 3, 4 et 5) ou transposés dans différentes tonalités (Expérience 2). Seule une structure globale fortement détruite rend la détection des cibles plus difficile. L'incohérence du contexte global n'influence pas la reconnaissance de vraies cibles, mais gêne le rejet des fausses cibles mélodiques (leurres). Les résultats révèlent que la cohérence globale du contexte a un faible effet sur l'identification des cibles et seulement pour des structures extrêmement modifiées.

INTRODUCTION

In the real world, events are generally integrated in a coherent structural organization; meaningful auditory verbal information comes to us in the form of sentences, meaningful visual information in the form of scenes. Perception and comprehension require not only the identification of single events, but also the specification of the relations and structures among these entities. Knowledge about context and its structures permits us to develop expectancies about what is to come and to identify patterns. It is more likely that we will recognize an event in an appropriate, structured context. Despite the presence of noise interference, a word presented in the context of a well-formed sentence is more intelligible than the same word presented in a random string of words (Miller, 1962). The influence of context on pattern recognition has also been demonstrated with figures (Weisstein & Harris, 1974), letters in words (Reicher, 1969; Warren, 1970), or objects in real-world scenes (Biederman et al., 1973).

In the Biederman et al. study (1973), coherency of context was manipulated through a scene-jumbling procedure. Pictures of scenes were presented in a normal, coherent version or in a jumbled version. In the jumbled version, the scene was cut into sixths and was rearranged. The jumbling destroyed the natural spatial relations of scene components. It was assumed that this modification manipulated the meaningfulness of the object's setting independently of the complexity of the scene. The effect of jumbling was studied with a speeded search task. A target object was first presented, followed by the original or the

Requests for reprints should be addressed to Barbara Tillmann, LEAD–CNRS ESA 5022, 6, Boulevard Gabriel, F–21000 Dijon, France (E-mail: Barbara.Tillmann@u-bourgogne.fr).

jumbled scene. Participants had to judge if the presented object appeared in the scene or not. Three kinds of objects were used: (1) the object was, in fact, in the scene (referred to below as "yes responses"), (2) it was not in the scene, but could have been in it ("possible-no responses"), or (3) the object was highly unlikely to have appeared in the scene ("impossible-no responses"). The results revealed longer response times for "possible-no responses" than for "yes-responses", and the shortest response times for "impossible-no responses". Jumbling increased response times to the greatest extent for objects that were not in the scene but that had a high probability of occurring in that kind of scene ("possible-no responses"). Objects that had really appeared in the scene were detected nearly as quickly in the jumbled as in the coherent version. The results were discussed in terms of schema activation. A scene schema is taken to be an overall internal representation integrating the scene's entities and their relations and allowing access to semantic information. Achieving a schema would be insufficient for "possible-no" objects. In order to determine if the object was present in the scene, participants would have to perform detailed feature processing and object identification. Biederman et al.'s findings suggest that this process was facilitated by a coherent, structured context corresponding to the schema, but that it was made more difficult by a jumbled, incoherent context.

Biederman et al. (1973) worked on structures and relations between objects in the perception of real-world scenes. Structures linking events exist not only in visual stimuli but also in sound environments and in music. The principal goal of the present study was to investigate the effect of global musical structures on target detection and recognition by using Biederman et al.'s jumbling procedure.

In Western tonal music, the global form of musical pieces relies on both motivic and harmonic structures (D'Indy, 1987; Rosen, 1972). Motivic relations are defined by a thematic development. Musical themes are exposed, varied, developed and, finally, exposed again at the end of the piece (Francès, 1958). Harmonic progressions also follow a formal organization and are constrained by a number of rules. Tonal unity represents one of the formal determinants. Musical pieces start in a main key, move through other keys (relatively near to or far from the starting key) and then finally return to the main key. Temporary modulation should create, in the listener, the expectancy of a return to the main key. Harmonic structures confer a strong unity on tonal pieces because they establish hierarchical relationships between all the musical events of the piece (Lerdahl & Jackendoff, 1983; Meyer, 1973; Schenker, 1935). The combination of motivic and harmonic structures ideally enables listeners to integrate local information into more global patterns (Dowling & Harwood, 1986; Francès, 1958). Studies using short and simple musical sequences have revealed the psychological reality of hierarchical structure in memorization tasks (Bigand, 1990, Exp. 2; Deutsch & Feroe, 1981), in phrase completion and tension judgements (Bigand, 1997; Boltz, 1989; Palmer & Krumhansl, 1987a, 1987b), in similarity judgements about musical sequences (Dibben, 1994; Serafine, Glassman, & Overbeeke, 1989), and in the development of expectations about following events (Boltz, 1993; Schmuckler, 1989, 1990; Schmuckler & Boltz, 1994).

However, the importance of global structure for the listener has been challenged by provocative results obtained with longer and more complex musical sequences. The influence of large-scale tonal closure on listeners' feeling of coherence and completion has been shown to be weak and restricted to fairly short time spans (Cook, 1987). Karno and Konecni (1992) reported that permuting the sections within a symphonic movement did not alter ratings of pleasure, interest, or the desire to own a copy. Segmenting piano pieces into short chunks and linking them in a backward order represents a drastic modification to the global structure. However, listeners' responses concerning expressiveness and coherence were not influenced by these inverted versions (Tillmann & Bigand, 1996). In short minuets, listeners understood authentic cadences in the dominant key locally (i.e. as a definitive ending) rather than globally (i.e. as a temporary ending) (Tillmann, Bigand, & Madurell, in press). These results raise the question of the extent to which global structures predominate over local structures.

In the present study, the global and local processing of musical structures were further investigated with short minuets. The main purpose was to analyse the influence of a coherent global context on the detection and recognition of musical target excerpts. In minuets, a short excerpt of circa 2.5 seconds was selected as the target. For the manipulation of context coherency, the Biederman et al. (1973) jumbling procedure was adapted for use with musical stimuli. The minuets were presented in their original version and in three modified ones. In the modified versions, the minuets were cut into chunks of four bars, two bars, or one bar. For each version, the order of the chunks was rearranged (Experiment 1, 3, 4, 5) or the chunks were transposed to different keys (Experiment 2). Biederman et al.'s visual speed search task was adapted with a target detection and a target recognition task. In the target detection task, the target was presented first, followed by the presentation of the whole minuet (Experiments 1, 2, 3). In the recognition task, the minuet was presented first, followed by a target excerpt: a real target or a foil (Experiment 4, 5). Biederman et al.'s results suggest that a coherent, structured context should facilitate target detection and recognition whereas a disorganized context should not. The present study went one step further than Biederman et al. (1973) by using three different versions with different levels of coherence (see Appendix 1). These variations also help us to analyze the level at which a more global organization influences target detection or recognition.

EXPERIMENT 1

Method

Participants

Thirty-four students from the University of Dijon participated in this experiment: 17 musically untrained students (referred to below as nonmusicians) and 17 students with an average of 10.7 years of instrumental instruction and formal musical training (referred to below as musicians).

Material

Eight minuets of a length of 16 bars were chosen (see Appendix 2). They were played with a sampled piano sound produced by a Yamaha EMT10 Sound Expander; the sampler was controlled through a MIDI interface and by a Macintosh computer running Performer software. In order to make the pieces more expressive, the dynamics and velocity of several tones were modified, but no rubato was performed. The tempo of the minuets was adjusted to produce a standard length of 20 seconds. In every minuet a target excerpt of a length of two bars, i.e. 2.5 seconds, was chosen. The target was situated in the second half of the minuet. It was not repeated inside the minuet and did not contain the same starting notes as the other bars. In four of the minuets, the target was situated in the 9th and 10th bars and in the other four minuets in the 13th and 14th bars (see Appendix 2). Three modified versions were constructed for each minuet. The minuets were cut into chunks of a length of four bars for one modified version (referred to below as "four-bar version"), into chunks of two bars for the second modified version ("two-bar version"), and into chunks of one bar for the third modified version ("one-bar version"). The four-bar versions were composed of 4 chunks, the two-bar versions of 8 chunks, and the one-bar versions of 15 chunks (including the target). As far as possible, the order of the chunks was rearranged in order to prevent adjacent chunks from appearing side-by-side in the new version. The target maintained its original serial position; this condition was set up in accordance with Biederman et al.'s method, in which the visual target object maintained its original spatial position. The sound stimuli were captured by Sound Edit Pro software. The experiment was run by Psyscope software (Cohen, MacWhinney, Flatt, & Provost, 1993). The software recorded participants' responses and reaction times. The reaction times were measured by Button Box's clock with an accuracy of 1msec.

Procedure

The target was presented twice. The participants started the presentation of the whole minuet by pressing a button. When listening to the whole minuet, participants were required to press a button as soon as they detected the target. They were encouraged to respond as quickly as possible during the target itself and not to wait for it to end. Correct responses (hits) and errors (false alarms) were indicated by two different sound signals. False alarms signified answers given outside the duration of the target excerpt. Modified versions of two musical examples were used for training and to demonstrate the material and task. All the participants listened to the original versions and the three modified versions of the eight minuets. The order of presentation of the minuets and versions was randomized for each participant.

Results

Overall, the percentage of hits was high in the musicians' responses (88.6%) and 94% of these hits were not preceded by false alarms. Hits were

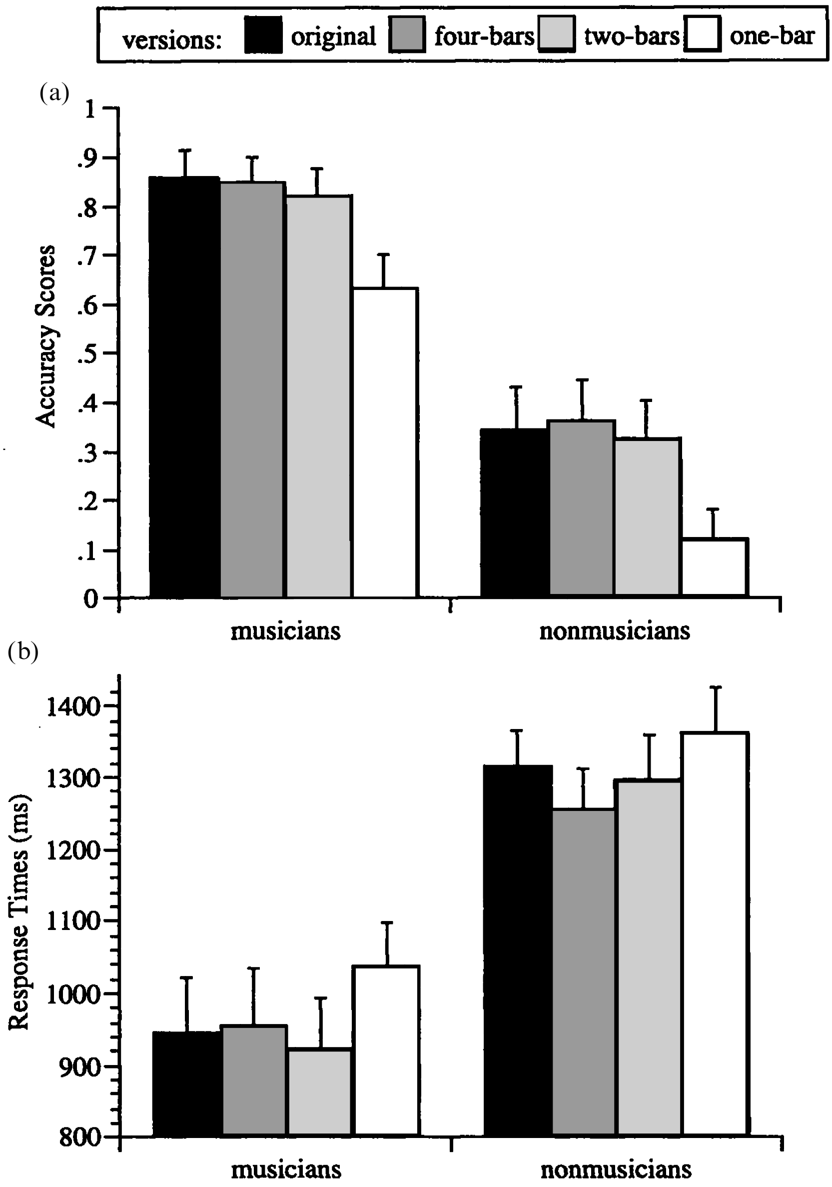

FIG. 1. Mean accuracy scores (a) and mean response times (b) for the four versions of the minuets and the two levels of musical expertise.

less numerous with nonmusicians (61.6%) and 71% of these hits were not preceded by false alarms. Response accuracy scores for the relationship between hits, false alarms, and the number of given responses were calculated for each participant and for each of the four versions: (hits − false alarms) / number of responses. The response times for all the correct responses were analyzed. Correct responses were defined as hits that were preceded by no more than three false alarms. Figure 1 depicts the mean accuracy scores and the mean response times for each version and the two levels of musical expertise. Musicians had a better response pattern than nonmusicians, but for all participants response accuracy scores decreased and response times increased for the one-bar versions only.

Accuracy scores and response times were both analyzed using a musical expertise (2) × versions (4) analysis of variance (ANOVA), with musical expertise as the between-subject factor and versions as the within-subject factor. The ANOVAs were performed by subjects (F_1) and by items (F_2). Since subject analyses were mirrored in the item analyses, only the subject analyses will be reported here.

The main effect of version was significant for accuracy scores, $F(3, 96) = 9.2$, $P < .0001$, and for response times, $F(3, 96) = 3.01$, $P < .05$. For the one-bar versions, accuracy scores decreased, $F(1, 32) = 25.94$, $P < .0001$, and response times increased, $F(1, 32) = 6.88$, $P < .05$, when compared with the original versions, and the four-bar and two-bar versions. This comparison explained 98.3% of the experimental variance for accuracy scores and 93.1% for response times. None of the contrasts between the original, the four-bar, and the two-bar versions were significant.

In addition, there was a significant effect of musical expertise for accuracy scores and for response times. Musicians' scores were higher than nonmusicians'; they gave more correct responses than nonmusicians and made relatively few false alarms, $F(1, 32) = 39.2$, $P < .0001$. Musicians responded on average 340msec faster than nonmusicians, $F(1, 32) = 17.7$, $P < .001$. The interaction between musical expertise and versions was never significant.

Discussion

In this study, Biederman et al.'s jumbling procedure was adapted for use with musical pieces. Together, the accuracy scores and response times revealed the influence of jumbling in the most incoherent versions. When the structural organization of the piece was broken on a very local level, i.e. a length of one bar, target detection was more difficult: accuracy scores decreased and response times increased. A jumbling in groups of four or two bars did not influence response patterns when compared with the original pieces. The effect of incoherent versions seems to be limited to a very local level.

The various coherent versions were constructed with the help of the jumbling procedure. Consequently, the local, superficial features surrounding the target also changed with the level of incongruence. Changes in pitch intervals or in duration values preceding the target might be responsible for modulating the emphasis placed on the target in the different contexts. Three analyses were performed to address this issue.

First, a multiple regression analysis was run in order to predict response time data with four variables: the level of context coherence, the duration value of the last soprano note before the target, the duration value of the first soprano note of the target, and the pitch interval between these two notes. A linear combination of these four variables provided a good fit for the response time data, $R^2 = .75$, $F(4, 27) = 19.6$, $P < .0001$. There was a significant contribution of the duration value of the first target note, $t = 7.99$, $P < .0001$, and of the size of the pitch interval, $t = 2.4$, $P < .05$, but not of the note preceding the target. This indicates that the longer the first note of the target and the greater the pitch interval, the longer the response times. The contribution of the level of context coherence was only marginally significant, $t = 1.9$, $P = .07$, with a more incoherent context tending to result in longer response times. However, local features, especially the value of the first note of the target, have a greater effect on response times than does the global coherence of the context.

In two of the eight minuets, the target starts with a long note, which could of course result in prolonged detection times. A second multiple regression analysis was conducted without these two minuets. As previously, a linear combination of the four variables provided a good fit for the data, $R^2 = .51$, $F(4, 19) = 4.9$, $P < .01$. The level of context coherence did not contribute significantly to the detection data. The contribution of the pitch interval size was significant, $t = 3.3$, $P < .01$, and the duration of the last note before the target was marginally significant, $t = 2.1$, $P = .052$. Local features continue to be a more important factor for response time data than the global coherence of the different versions. This suggests that the salience of the local boundaries preceding the target is more important for target detection than global context coherence: the greater the pitch interval and the longer the note preceding the target, the easier the detection of the target.

Given the importance of the boundary preceding the target, it was necessary to test in a third analysis that the salience of this boundary did not covary with the change in global coherence. The most problematic confound for the present findings would have taken the form of an increase in the salience of the boundary with the level of structural incoherence. The former factor would facilitate target detection, whereas the latter would work in the opposite way. As a consequence, the sum of these factors could lead to a very small effect of global structure. For the eight minuets, the size of the pitch interval preceding the target (in the soprano voice) and the duration of the tone preceding the target were computed for each experimental condition. Two ANOVAs were performed with the minuets as the random factors and the pitch interval and the duration as the dependent variables. Only the duration of the last note before the target varied significantly with the level of structural incoherence, $F(3, 21) = 6.97$, $P < .001$), and this effect is mainly explained by a quadratic tendency, $F(1, 7) = 13.75$, $P < .01$): the duration of the last note was longer in the four-bar version minuets (mean = 3 beats) than in the original version (mean = 2 beats) or the two-bar version (mean = 1.91 beats). They were shortest in the one-bar version (mean = 1.45 beats). Such a quadratic trend should have strengthened the effect of the global structure. Previous analysis revealed that this was not the case, since this effect remained weak. Therefore, the weak effect of the global structure reported in Experiment 1 cannot be explained in terms of a confound with local incoherence factors.

Despite the fact that local transitions cannot be the cause of the findings reported here, Experiment 2 was designed to manipulate global coherence by keeping the superficial features surrounding the target constant. The tonal unity of the minuets was manipulated by transposing the chunks to different keys. In the same way as for the jumbling in Experiment 1, transpositions were performed 4 times in the four-bar versions, 8 times in the two-bar versions, and 15 times in the one-bar versions. These modifications should strongly affect the global coherence of the context and should influence the detection of the target.

EXPERIMENT 2

Method

Participants

Thirty-two students from the University of Dijon participated in this experiment: 16 had never received formal musical training, nor learned a musical instrument (referred to below as nonmusicians), and the other 16 had an average of 6.7 years of formal musical training and 9.4 years of instrumental instruction (referred to below as musicians). None had participated in Experiment 1.

Material and Procedure

The eight minuets and targets of Experiment 1 were used and played as explained earlier. In the modified versions, the minuets were cut into chunks of four bars, two bars, or one bar (“four-bar version”, “two-bar version” and “one-bar version”). For each version, the different chunks were played in one of four tonalities: the original one, one semitone above, one semitone below, and one whole tone below the original. This type of transposition minimized changes in pitch interval and avoided changes of duration values between the different versions. Each of the four tonalities was used in the same frequency. The chunks that contained the target were played in the original tonality. The construction of the stimuli and the procedure were the same as in Experiment 1.

Results

Overall, the percentage of hits in the musicians’ responses was high (93.1%) and 96.6% of these hits were not preceded by false alarms. Hits were less numerous with nonmusicians (73.3%) and 84.9% of these hits were not preceded by false alarms. The accuracy scores were calculated and the response times for all the correct responses were analyzed. As Table 1 shows, accuracy scores and response times were not altered by the structural modifications. Two musical expertise (2) × versions (4) analyses of variance (ANOVAs), with musical expertise as the between-subject factor and versions as the within-subject factor, were conducted on response times and accuracy scores by subjects (F_1) and by items (F_2). Since subject analyses were mirrored in the item analyses, only the subject analyses will be reported here. For accuracy scores and response times, there was no significant effect of version. Analyses of accuracy scores revealed a significant effect of musical expertise, $F(1, 30) = 18.75$, $P < .001$. Nonmusicians made more false alarms and gave fewer correct responses than musicians. The response time analyses revealed no significant effect of musical expertise. The interaction between the factors of version and musical expertise was never significant.

Discussion

In this experiment, global coherence was manipulated by keeping the salience of the boundaries preceding the target constant. An incoherent tonal context did not slow down response times and neither did it decrease accuracy scores. The destruction of tonal unity, even at a very local level, did not increase the difficulty of the target detection task. Experiment 2 confirms the very

TABLE 1
Mean Accuracy Scores and Mean Response Times (in msec) for the Four Versions and the Two Levels of Musical Expertise in Experiment 2 (Standard Errors in Brackets)

	Versions			
	Original	*4-bar*	*2-bar*	*1-bar*
Accuracy Scores				
Musicians	.91	.86	.90	.81
	(.04)	(.04)	(.03)	(.05)
Nonmusicians	.49	.56	.43	.57
	(.09)	(.06)	(.11)	(.07)
Response Times				
Musicians	1103.1	1053.53	1105.78	1040.57
	(44.9)	(97.3)	(64.2)	(60.3)
Nonmusicians	997.52	1056.36	1065.6	1089.57
	(68.2)	(72.5)	(79.3)	(68.7)

weak influence of global context on target detection already reported in Experiment 1. It further indicates that our previous findings could not have been caused by any interference influences due to local coherence factors.

It may be argued that this weak effect might also have been caused by the experimental design used in the first two experiments. The fact that the target was always present in the minuets could have encouraged the participants to attenuate the processing of the musical information. If this is the case, the absence of a "possible-no responses" condition may have counteracted the influence of global structures. To address this issue, Experiment 1 was re-run with a new condition including foils. It was assumed that manipulating the global structure would have a stronger effect on performance than in Experiment 1.

EXPERIMENT 3

Method

Participants

Thirty-two students from the University of Dijon participated in this experiment: 16 musically untrained students (referred to below as nonmusicians) and 16 students with an average of 10.7 years of instrumental instruction and formal musical training (referred to below as musicians). None had participated in Experiments 1 and 2.

Material and Procedure

The eight minuets of Experiment 1 were used again and were played as described earlier. For half of the participants, four minuets were presented with foils, and the others were presented with targets. For the other half of the participants, the first group of minuets was presented with targets, and the second with foils. A foil was created for each target: the pitches of the target were played in a backward order, and certain changes in rhythmic structure were introduced in order to render the foil musically plausible. In a pre-test, real targets and foils were judged on a subjective scale ranging from 1 (slightly surprising) to 7 (very surprising). An ANOVA with target excerpts and minuets as within-subject factors revealed no effect of target excerpts, $F(1, 9) = 1.1$. Foils (2.1) were judged to be as musically coherent as targets (2.4). Targets or foils were presented twice before the minuets. The participants were informed that the target excerpts might or might not be present in the minuet. For each participant, the order of presentation of the four versions of the eight minuets was randomized.

Results

The first analysis considered the data of the minuets presented with the real target. Overall, the percentage of hits in the musicians' responses was very high (95.78%) and 98.7% of these hits were not preceded by false alarms. Hits were less numerous with nonmusicians (65.9%) and 78.8% of these hits were not preceded by false alarms. Response accuracy scores for the relationship between hits, false alarms, and the number of given responses were calculated for each participant and for each of the four versions: (hits − false alarms)/number of responses. The response

TABLE 2
Mean Accuracy Scores and Mean Response Times (in msec) for the Four Versions and the Two Levels of Musical Expertise in Experiment 3 (Standard Errors in Brackets)

	Versions			
	Original	*4-bar*	*2-bar*	*1-bar*
Accuracy Scores				
Musicians	.82	.92	.86	.84
	(.07)	(.05)	(.06)	(.07)
Nonmusicians	.39	.51	.41	.19
	(.09)	(.08)	(.13)	(.13)
Response Times				
Musicians	1241.1	1214.5	1163.6	1332.3
	(72.8)	(77.4)	(83.6)	(91.1)
Nonmusicians	1363.7	1353.9	1409.7	1387.2
	(124.8)	(59.9)	(103.4)	(124.5)

times of all hits were analyzed. Accuracy scores and response times were both analyzed using a musical expertise (2) × versions (4) analysis of variance (ANOVA), with musical expertise as the between-subject factor and versions as the within-subject factor. The effect of version failed to reach a significant level for accuracy scores, $F(3, 90) = 2.08$, $P = .11$, and was not significant for response times, $F(3, 90) = < 1$. For both groups of participants, response times were longer than in Experiment 1 and accuracy scores were similar (Table 2). The critical finding was that the effect of version remained very weak, and was even less pronounced than in Experiment 1. The sole evidence for this effect came from the one-bar version. In this condition, accuracy scores were the lowest for the nonmusicians, $F(1, 30) = 4.4$, $P < .05$, and response times tended to be longer for the musicians, $F(1, 30) = 3.40$, $P = .07$.

The second analysis concerned the number of false alarms committed for minuets presented with foils. False alarms were 3.7 times higher for nonmusicians than for musicians $F(1, 30) = 48.09$; $P < .001$, but there was no effect of version, $F(3, 90) < 1$, nor any significant interaction.

Discussion

Contrary to what has been expected, adding a foil condition to the experimental design did not increase the effect of global coherence. Thus Experiment 3 confirmed Experiment 1: the global organization of the piece seems to have no major influence on target detection. Only when the global organization was broken on a very local level did target detection tend to be more difficult. To counter the possible argument that a detection task is not appropriate for revealing the effect of global structure, the experimental task was changed in Experiment 4. A recognition task requires that participants memorize the whole minuet and thus consider the global organization of the piece. Memory capacity can be enhanced by chunking the incoming information. Structured material permits a more economical encoding and better retrieval than does material lacking structural organization (Bower, 1970; Thorndyke, 1977). The original versions with a coherent context would assist in the memorization of the whole piece and would facilitate recognition of a target excerpt. This is not the case for the incoherent versions.

EXPERIMENT 4

Method

Participants

Thirty students participated in this experiment: 15 musically untrained students (referred to below as nonmusicians) and 15 students at Dijon music conservatory (referred to below as musicians). None had participated in Experiments 1, 2, or 3.

Material and Procedure

The original versions, the two-bar versions, and the one-bar versions, as well as the real targets and foils of Experiment 3, were used. The experiment was run by Psyscope software (Cohen et al., 1993) and reaction time was measured by a Button Box clock.

The minuet was first presented, then followed by a silence of 2 seconds, a white noise of 250msec, a silence of 750msec, and by the target excerpt (the real target or the foil)[1]. The noise mark was added in order to eliminate the possible influence of sensory memory (Kallman & Massaro, 1983). Participants were asked to judge as fast as possible if the target excerpt had occurred in the minuet by pressing a button. They could respond from the very beginning of the target excerpt until 2 seconds after it had ended. Correct responses and errors were indicated by two different sound signals. Each minuet and each version was tested with the real target and the foil. The order of presentation of the minuets and the different versions was randomized.

Results

The recognition accuracy data and the response time data were analyzed by means of separate musical expertise (2) × versions (4) × target excerpts (2) analyses of variance, with musical expertise as the between-subject factor and versions and target excerpts as the within-subject factors. Subjects and items were treated as random effects. As far as subject analysis mirrored item analysis, only the former was reported.

Figure 2 shows the means of the correct responses as a function of versions, target excerpts, and the level of musical expertise. More correct responses were observed for real targets than for foils. The different versions influenced only the rejection of the foils, but not the real targets. Analyses of recognition accuracy data revealed a significant main effect of version, $F(2, 56) = 4.4$, $P < .05$; of targets, $F(1, 28) = 27.4$, $P < .001$, and of musical expertise, $F(1, 28) = 4.5$, $P < .05$. The three-way interaction between these factors was significant, $F(2, 56) = 3.4$, $P < .05$. Planned comparisons for real targets indicated no significant differences between the three versions or the two levels of musical expertise. In the case of foils, there were significantly more correct responses for the original versions than for the one-bar versions, $F(1, 28) = 5.98$, $P < .05$; independently of musical expertise. For nonmusicians, the number of correct responses decreased for two-bar versions, $F(1, 26) = 6.1$, $P < .05$, but not for one-bar versions. For musicians, the number of correct responses for original versions was the same as for two-bar versions, but it decreased significantly for one-bar versions, $F(1, 28) = 7.2$, $P < .05$.

Figure 3 depicts the mean response times for correct responses for each version, target excerpts, and the two levels of musical expertise. There was no effect of versions, although there was a significant interaction between versions and target excerpts, $F(2, 56) = 4.2$, $P < .05$. For real targets, the response times for the three versions did not differ significantly. For foils, the response times for the one-bar versions were significantly longer than for the original versions, $F(1, 28) = 6.2$, $P < .05$. Only in the case of nonmusicians was the difference in response times between the original versions and the two-bar versions marginally significant, $F(1, 28) = 3.0$, $P = .09$. Furthermore, response times for real targets were shorter than for foils, $F(1, 28) = 68.45$, $P < .0001$. Musicians responded faster than nonmusicians. However, this influence was not significant in the subject analyses, $F_1(1, 28) = 2.78$, $P = .11$, although it was so in the item analyses, $F_2(1, 12) = 8.74$, $P < .05$.

In order to test the possible role of local features surrounding the target on recognition (see Experiment 1), a further multiple regression analysis was performed on the recognition data. It revealed no influence of local features preceding the target on response time data. The salience of local boundaries seems to be less important in target recognition than in target detection.

Discussion

This result provides evidence that a global context had only a weak effect on recognition performance. No influence on real target recognition was observed, although the level of foil rejection was affected. Beyond that, only an incoherent context with chunks of one bar makes foil rejection more difficult. Independently of the level of musical expertise, the number of correct responses decreased and the response times increased for one-bar versions when compared with the original versions.

For nonmusicians, the influence of context on foil rejection is also observable for the two-bar versions, although it is less marked than for the one-bar versions. However, the observation that musicians' responses for the two-bar versions did

[1] Due to sound reverberation in this context the very end of the target and foil was 2900msec.

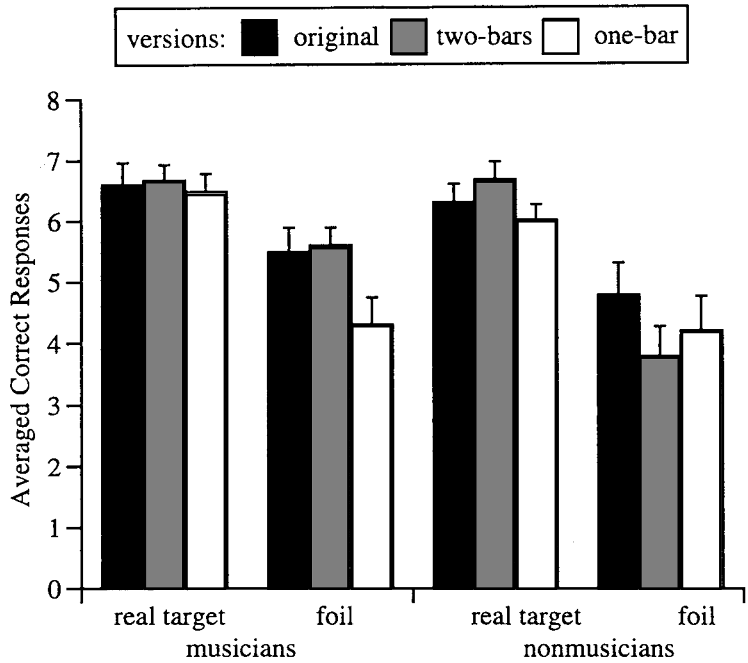

FIG. 2. Averaged correct responses for the two target excerpts, the three versions, and the two levels of musical expertise.

not differ from their original version responses was surprising. Musicians seem to resist this level of contextual incoherence. Participants listened to all three versions during the experimental session. The repeated hearing of the minuets and the target excerpts could have reduced the effect of an incoherent context on memorization, especially for musicians. To examine this point in greater detail,

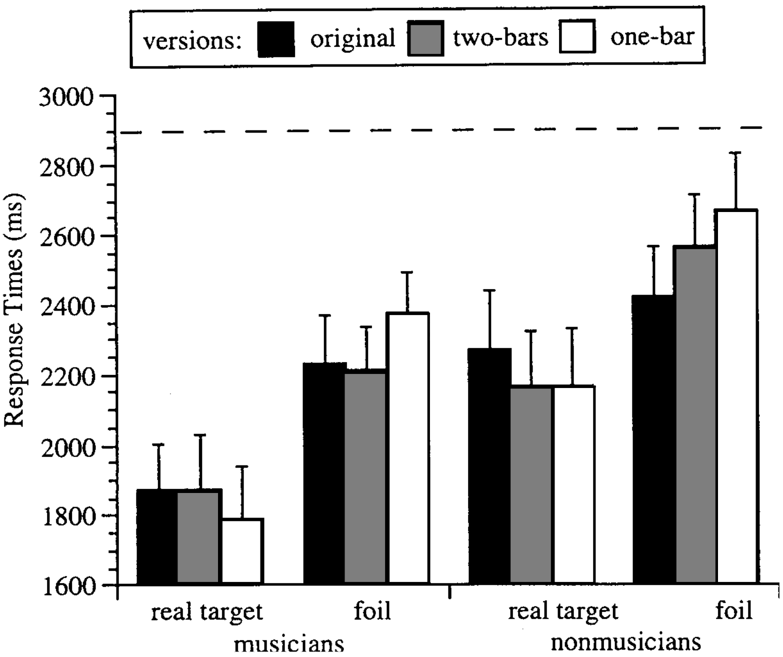

FIG 3. Average response times for the two target excerpts, the three versions, and the two levels of musical expertise (the dotted line represents the end of the target excerpt).

the different versions were used as a between-subject factor in Experiment 5.

EXPERIMENT 5

Method

Participants

Twenty-six students participated in this experiment: 13 musically untrained students (referred to below as nonmusicians) and 13 students at Dijon music conservatory (referred to below as musicians). None had participated in any of the previous experiments.

Material and Procedure

In order to pursue the analysis of the effect of context coherence, only the original versions and the two-bar versions of Experiment 3 were retained. Each minuet and each version was tested with both the foil and the target. Participants worked on the original versions of four minuets and on the modified versions of four other minuets. The way the minuets were assigned to the experimental conditions was counterbalanced across participants. The order of presentation was randomized, with the restriction that the same minuet was never presented twice in succession with its real target and its foil. The other aspects of the procedure were the same as those described in Experiment 4.

Results

Figure 4 shows that the different versions primarily influenced the correct response levels and response times for the foils. Subject analyses did not reveal a main effect of versions, although the interaction between versions and target excerpts was significant for correct responses, $F(1, 24) = 6.5$, $p < .05$, and marginally significant for response times, $F(1, 24) = 6.53$, $P = .06$. Planned comparisons indicate that for real targets the number of correct responses and the response times did not differ significantly between the original and the modified version. For foils, the number of correct foil rejections decreased for modified versions, $F(1, 24) = 8.1$, $P < .01$, for both groups. Response times increased significantly for musicians only, $F(1, 24) = 4.99$, $P < .05$, but not for nonmusicians, $F(1, 24) < 1$. Furthermore, for real targets fewer errors were committed, $F(1, 24) = 10.3$, $P < .01$, and response times were shorter than for foils, $F(1, 24) = 56.99$, $p < .0001$. Musicians gave more correct responses, $F(1, 24) = 11.4$, $P < .01$, and responded faster than nonmusicians, $F(1, 24) = 9.7$, $P < .01$. Item analysis for response times generally mirrored the pattern of significance of the subject analysis.

Discussion

In general, the results of Experiment 5 reveal much the same pattern as Experiment 4. Contextual coherence had no effect on the recognition of real targets, although it did affect the rejection of the foils. Performances decreased for the one-bar and for the two-bar versions. This is broadly similar to nonmusicians' results in Experiment 4. The most important new finding concerned musician listeners: in this experiment an increase in the response times and a decrease in the number of correct responses were also observed for the two-bar versions. Experiment 5 emphasized the differences between the original versions and the two-bar versions in the musicians' results.

To summarize, both Experiments 4 and 5 showed that contextual coherence did not influence target recognition, but that it impaired the rejection of the foils. Response accuracy and response times were influenced by a reorganization on a two-bar level and on a one-bar level. Furthermore, the rejection of a foil required longer response times and caused more incorrect responses than the recognition of a real target. These findings are consistent with the interaction between coherence and response category observed by Biederman et al. (1973): response times were longer for foils than for real targets and the effect of jumbling was essentially shown for the foils.

GENERAL DISCUSSION

This set of experiments adapted an approach used in visual perception for the investigation of the effect of global musical structure on target detection and recognition. Contextual coherence was modified by a chunk-jumbling procedure and by transposing chunks to different keys. The latter modification did not affect performance, suggesting that tonal unity is not sufficiently important to facilitate target detection for the listener. However, jumbling the order of the chunks had a weak influence on target detection and recognition. Despite this, in the case of target detection a context effect was observed only for the most incoherent versions, and in the case of target

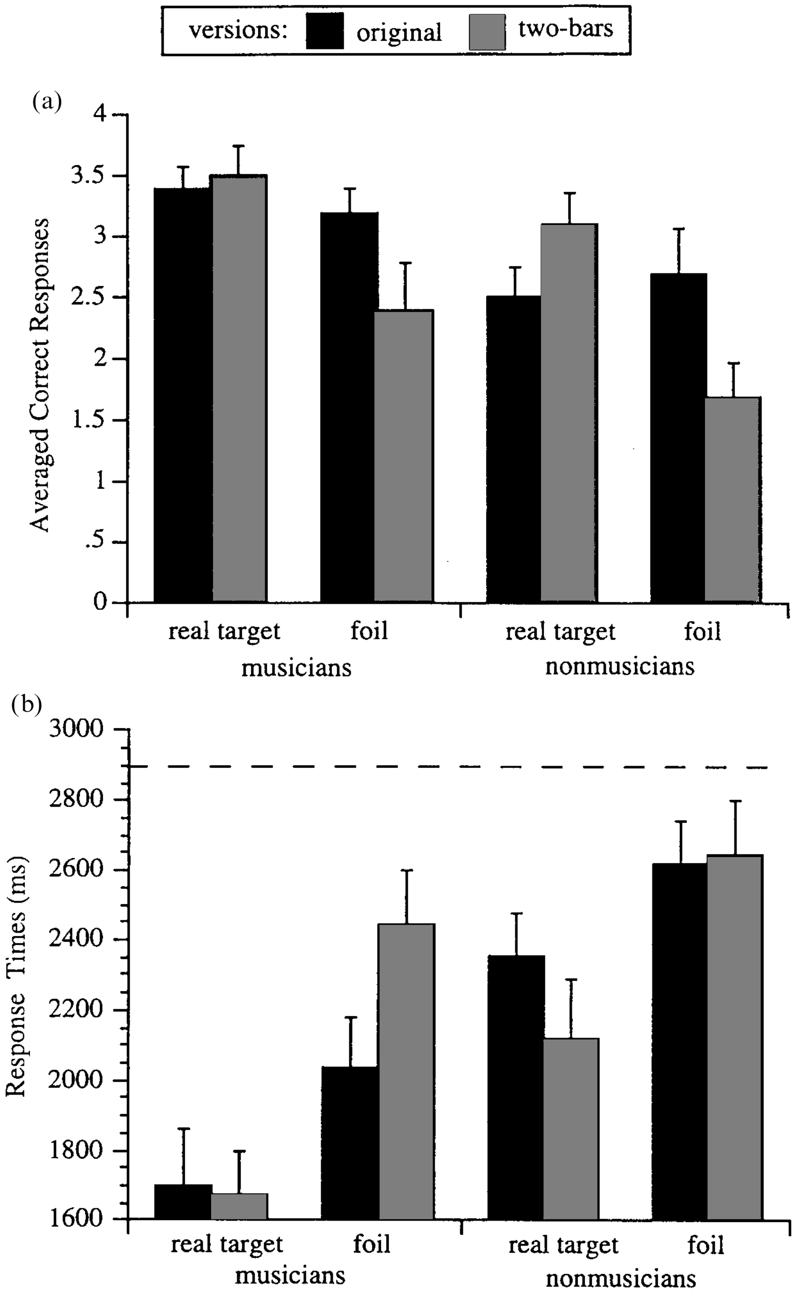

FIG. 4. Averaged correct responses (a) and average response times (b) for the two target excerpts, the three versions, and the two levels of musical expertise (the dotted line represents the end of the target excerpt).

recognition the performance on the one-bar and two-bar versions was impaired for the foils only.

In the recognition task, the listener was supposed to form memory traces of the whole minuet and then to match the presented excerpt with the stored information. This processing demands longer response times for foils than for real targets and the contextual coherence of the material only affected the rejection of the foils. Differences between positive and negative responses were also observed in a "two choice-matching" paradigm. Same or positive responses are faster than different or negative responses and more false different responses are made than false same responses (Bamber, 1969; Proctor & Rao, 1983). Ratcliff (1985) has proposed a model in terms of criteria setting to account for this "fast same effect".

The results of the recognition task presented here can be more comfortably integrated within the framework of a memory model such as the "dual process model" defined by Mandler (1980). Two processes are invoked to account for long-term memory: when a subject is asked to make a judgement concerning the prior occurrence of an object, the first process tests the familiarity value of the presented event. The second, slower, mechanism engages in a search and retrieval process that attempts to determine whether the target was originally present (i.e. elaboration). Although the testing of the dual process was primarily based on a long-term recognition paradigm, this theory may be of relevance for the interpretation of the present data. Responding to real targets could be based on a fast familiarity judgement. For foils, the subjective familiarity test does not permit a response and a search or retrievel process is activated. A memory list must be checked before responding, and this results in slower responses (Juola, Fischer, Wood, & Atkinson, 1971). The search and retrieval process, which is necessary for the rejection of the foils, is influenced by the organization of the material. This influence on the retrieval process is generally observed in memory studies (Bower, 1970; Thorndyke, 1977). In contrast to the foils, responses for real targets were not influenced by contextual coherence. Atkinson, Hermann, and Wescourt (1974) have observed that familiarity seems to affect recognition independently of the context in which familiarity is incremented. This independence and the fact that familiarity-based judgements require little processing capacity (Jacoby, 1991) could result in the present observation. Therefore, it is interesting to ask whether context would also influence the recognition of the real target when the whole task is made more difficult. Two possibilities can be considered: the use of shorter target excerpts that do not form a unit or the use of more complex and longer musical pieces. The global structure might become more important for the processing of long pieces and incoherence may exert a stronger effect on target recognition.

The present findings suggest that global structures in music are not very important for target identification. Despite the difference in modality (auditory v. visual) and in task (recognition v. detection), the results of the present experiments mirrored Biederman et al.'s result (1973), with a context effect primarily being observed for foils and in really disorganized scenes. Biederman et al. cut the visual scenes into sixths, a procedure that is quite comparable to the seven jumbled chunks of the two-bar versions.

Furthermore, the findings of the present experiments can be compared with results for word and form identification. In these studies, the influence of context on target identification was also observed with extremely modified structures. Miller (1962) juxtaposed the context of a well-formed sentence with a random string of words. In the case of form identification, the influence of context was observed when comparing a coherent object with a completely incoherent configuration (Weisstein & Harris, 1974). The data of the present study raise the question of what would happen when less drastic modifications are employed in these fields of perception.

Manuscript first received October 1996
Revised manuscript accepted August 1997

REFERENCES

Atkinson, R.C., Hermann, D.J., & Wescourt, K.T. (1974). Search processes in recognition memory. In R.L. Solso (Ed.), *Theories in cognitive psychology: The Loyola Symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Bamber, D. (1969). Reaction times and error rates for "same"–"different" judgements of multidimensional stimuli. *Perception and Psychophysics*, *6*, 169–174.

Biederman, I., Glass, A.L., & Stacy, E.W. (1973). Searching for objects in real world scenes. *Journal of Experimental Psychology*, *97*, 22–27.

Bigand, E. (1990). *Perception et compréhension des phrases musicales*, Unpublished doctoral dissertation, Université de Nanterre (Paris X), Nanterre.

Bigand, E. (1997). Perceiving musical stability: The effect of tonal structure, rhythm and musical expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 808–822.

Boltz, M. (1989). Perceiving the end: Effect of tonal relationships on melodic completion. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 749–761.

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics*, *53*, 585–600.

Bower, G.H. (1970). Organizational factors in memory. *Cognitive Psychology*, *1*, 18–46.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers*, *25*, 257–271.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception*, *5*, 197–206.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, *88*, 503–522.

Dibben, N. (1994). The cognitive reality of hierarchic structure in tonal and atonal music. *Music Perception*, *12*, 3–17.

D'Indy, V. (1987). Cours de composition musicale. Paris: Durand et fils.

Dowling, W.J., & Harwood, D.L. (1986). *Music cognition*. Orlando, FL: Academic Press.

Francès, R. (1958). *La perception de la musique* (2nd edn., 1984). Paris: Vrin.

Jacoby, L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*, 513–541.

Juola, J.F., Fischer, I., Wood, C.T., & Atkinson, R.C. (1971). Recognition time for information stored in long-term memory. *Perception and Psychophysics*, *10*, 8–14.

Kallman, H.J., & Massaro, D.W. (1983). Backward masking, the suffix effect and perceptual storage. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*, 312–327.

Karno, M., & Konecni, V.J. (1992). The effects of structural interventions in the first movement of Mozart's Symphony in G-Minor, K. 550, on aesthetic preference. *Music Perception*, *10*, 63–72.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Mandler, G. (1980). Recognizing: The judgement of previous occurrence. *Psychological Review*, *87*, 252–271.

Meyer, L.B. (1973). *Explaining music: Essay and exploration*. Chicago, IL: University of Chicago Press.

Miller, G.A. (1962). Decision units in the perception of speech. *IRE Transaction on Information Theory*, *8*, 81–83.

Palmer, C., & Krumhansl, C. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 116–126.

Palmer, C., & Krumhansl, C. (1987b). Pitch and temporal contributions to musical phrase perception: Effect of harmony, performance timing and familiarity. *Perception and Psychophysics*, *41*, 505–518.

Proctor, R.W., & Rao, K.V. (1983). Evidence that the same-different disparity is not attributable to response bias. *Psychological Review*, *88*, 291–326.

Ratcliff, R. (1985). Theoretical implications of the speed and accuracy of positive and negative responses. *Psychological Review*, *92*, 212–225.

Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 275–280.

Rosen, C. (1972). *The classical style*. New York: Norton.

Schenker, H. (1935). *Der Freie Satz* (Trans. E. Oster [1979]). *Free composition*. New York: Longman.

Schmuckler, M.A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, *7*, 109–150.

Schmuckler, M.A. (1990). The performance of global expectations. *Psychomusicology*, *9*, 122–147.

Schmuckler, M.A., & Boltz, M. (1994). Harmonic and rhythmic influences on musical expectancy. *Perception and Psychophysics*, *56*, 313–325.

Serafine, M., Glassman, N., & Overbeeke, C. (1989). The cognitive reality of hierarchic structure in music. *Music Perception*, *6*, 397–430.

Thorndyke, P. (1977). Cognitive structures in comprehension and memory in narrative discourse. *Cognitive Psychology*, *9*, 77–110.

Tillmann, B., & Bigand, E. (1996). Does formal musical structure affect perception of musical expressiveness? *Psychology of Music*, *24*, 3–17.

Tillmann, B., Bigand, E., & Madurell, F. (in press). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, *167*, 392–393.

Weisstein, N., & Harris, C.S. (1974). Visual detection of line segments: An object-superiority effect. *Science*, *186*, 752–755.

APPENDIX 1

Example of One Original Piece and Its Three Modified Versions

Original version

Two-bar version

Four-bar version

One-bar version

APPENDIX 2

Composer	*Reference*	*Page*
Target in 9th and 10th bars		
Mozart	*Early Dances*, Ed. Lakos, A., Köneman: Budapest.	125
Mozart	*Early Dances*, Ed. Lakos, A., Köneman: Budapest.	127
Mozart	*Klavierstücke*, Ed. Lakos, A., Köneman: Budapest.	70
Turk	*Early Dances*, Ed. Lakos, A., Köneman: Budapest.	130
Target in 13th and 14th bars		
Haydn	*Early Dances*, Ed. Lakos, A., Köneman: Budapest.	133
Haydn	*6 Sonatinen*. Ed. Woehl, W., Schott: Mainz, 2333.	12
J. Ch. Bach	*Piano Progress*, Book 2, Faber Music: London, 1985.	5
Bach	*Le Petit Livre d'A.M. Bach*, Ed. H. Lemoine: Paris	29