Musicae Scientiae
http://msx.sagepub.com/

Musically puzzling I: Sensitivity to overall structure in the sonata form?

Roni Y. Granot and Nori Jacoby

Musicae Scientiae 2011 15: 365 originally published online 6 July 2011
DOI: 10.1177/1029864911409508

Article

MUSICÆ SCIENTIÆ

Musicae Scientiae
15(3) 365–386

DOI: 10.1177/1029864911409508
msx.sagepub.com
SAGE

Musically puzzling I: Sensitivity to overall structure in the sonata form?

Roni Y. Granot
The Hebrew University, Israel

Nori Jacoby
The Hebrew University, Israel

Abstract

Previous studies have suggested that listeners are not sensitive to the overall tonal structure of musical pieces. This assumption is reexamined in the current study in an active musical puzzle task, with no time constraints, focusing on the presumably most directional musical form – the sonata form. In our first study (reported here, and referred to as "the Mozart study"), participants with varying levels of musical training were presented with disordered sections of Mozart's piano sonata K. 570/I in B flat major and asked to rearrange the ten sections into a musically logical coherent whole. A second study (to be reported in *Musicae Scientiae* issue 16[1]) replicated the task in a different group of participants who listened to Haydn's piano sonata, Hob: XVI-34/I in E minor. In contrast with previous studies, we do not focus on listeners' ability to recover the original sonatas. Rather, we explore emergent patterns in their responses using new types of analysis. Our results indicate that listeners show: (1) Some sensitivity to the overall structure of A-B-A' around the non-stable B section; (2) Non- trivial sensitivity to overall "directionality" through a new type of analysis ("distance score"); (3) Correct grouping and placement of developmental sections possibly related to listener's sensitivity to musical tension; (4) Sensitivity to opening and closing gestures, thematic similarity and surface cues and; (5) No sensitivity to global harmonic structure.

Keywords

concatenationism, global structure, local processing, musical puzzle, order effects, sonata form, structural coherence

Introduction

One of the central concerns in music criticism and theory is the attempt to understand the ways in which the various elements in a piece of music – its pitches, rhythms, dynamics, timbres, motifs, sections – are organized so as to create a coherent whole. Many organizing principles underlying the musical structure at the tone-to-tone level, such as *gestalt* principles,

Corresponding author:
Dr Roni Y. Granot, Musicology Department, The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 91905, Israel.
Email: rgranot@huji.013.net.il

grouping, consonance vs. dissonance, and tonality, have been found to be cognitively valid (for a review see Bigand & Poulin-Charronnat, 2006; Justus & Bharucha, 2002). In contrast, there seems to be a large gap between notions of theorists and aestheticians regarding principles underlying the structure of entire movements, and the ability – or rather inability – of listeners to perceive these principles. In the current study we reexamine this gap using an active puzzle task, with no time constraints and minimum memory load, while focusing on the presumably most directional musical form – the sonata form.

Music theorists apply the terms "structure," "form" and "design" in dealing with large-scale structures such as entire musical movements. Aesthetical writings on musical pieces of the late eighteenth and nineteenth centuries stressed the notion that the emerging form should be unified as an organic whole, in such a way that parts cannot be removed, added, or rearranged, without destroying the integrity of the piece (Solie, 1980). Another common assumption underlying the notion of musical structure was that of flow, directionality, and motion from one musical event to another – up to the final arrival at closure. Form, according to Cone (1987), is "a process in time," a directed motion from beginning to end through important structural "stations" where concepts of "before," "after," preparation, arrival, departure, and return are pertinent. This presumed directed motion is supported on the harmonic level by progressions, preparation and realization of harmonic key changes, and cadences. On the melodic and rhythmic levels, it is driven by relationships of similarity and temporal order, such as those formulated by Cone (Cone, 1987; see also Ockelford, 2004).

Processing large-scale tonal harmonic structures

Previous studies have shown that Western listeners, even those with no formal musical training, have vast implicit knowledge about the harmonic structure of tonal music (for a review see Bigand & Poulin-Charronnat, 2006). However, this sensitivity seems to be limited to local harmonic relationships, as shown in Tillmann, Bigand, and Madurell (1998). Their study included three experiments based on the jigsaw puzzle methodology. In each experiment, participants were instructed to combine the two halves of a simple musical form, such as a Minuet, into a coherent piece, transposed into several keys. Correct responses consisted of choosing two parts of the piece in the same key, and ordering them in the traditional sequence. The important manipulation was the use of perfect authentic cadences on the dominant rather than on the tonic key. When processed locally, these cadences give the impression of full closure. But when processed globally, that is, taking into consideration the hierarchic level on which the cadence is positioned, they should be perceived as incomplete, since they do not provide closure on the main key. Results confirmed previous findings, demonstrating an understanding of tonal harmonic markers such as cadences and modulation, however only on a local level, without the ability to integrate these markers into an overall structure (see also Cook, 1987, 1990; Marvin & Brinkman, 1999). Moreover, although musically trained participants were more efficient in coping with the task (i.e., had shorter response times and compared fewer pairs), their responses show a similar pattern to that found in the untrained group, with more errors in pieces that contained authentic cadences on the dominant key. In summarizing the empirical data and their view on the relative importance of local and global structures, Tillmann and Bigand assert:

> Global structures seem to have only weak influences on perception, and local structures seem to be much more important. Independently of level of musical expertise, listeners have difficulties considering relations between events that are far apart in time. And yet, understanding such distant relations would be necessary to integrate events into an overall structural organization. (2004, p. 218)

This conclusion is consistent with Levinson's *concatenationism*, first presented in his *Music in the Moment* (1997).

Two more recent studies seem to temper somewhat these strong conclusions. Lalitte and Bigand (2006) showed that within contemporary and popular music, extreme segmentation and scrambling (28–29 short 6-second segments) elicit lower aesthetic and higher incoherence ratings as compared to the original pieces. This result was obtained despite the low rate of correct detection of wrong local transitions in the scrambled pieces. In contrast with previous studies, these findings suggest that listeners do show at least some sensitivity to the global structure. However, it is important to note that in addition to the extreme level of segmentation that differs from most previous studies, listeners heard both the original and the scrambled version during the same session (though not in succession). Moreover, in three of the six excerpts, the global structure was supported by a gradual increase in density, loudness, or tempo that can develop across long spans of time, but does so on a surface rather than a deep structural level. Lalitte et al. (2009) further explored the notion that global structure does not depend on tonality, but on other rhetorical cues related to the nature of the various themes and their arrangement. They compared the segmentation and arousal ratings of two Beethoven sonatas and their atonal devised counterparts. The similarity in segmentation and arousal responses to the tonal and atonal versions contrasted with explicit similarity measures that were based on tonal information. Moreover, arousal could only be explained partially by a set of tested psychoacoustic variables, leading the authors to suggest that rhetorical devices such as contrast, repetition and amplification underlie the global structure and are perceived by musically trained as well as untrained listeners.

Sensitivity to global thematic relationships

The central cognitive concept presumed to enable large-scale thematic relations is that of categorization, derivation, and prototypes (for a summary, see special issue of *Music Perception*, 2001, 18/3). Deliège (2001) proposed the "cue abstraction model," according to which listeners create a "mental line" based on a sequence of cues they abstract from the musical surface. These cues can be viewed as prototypes, representing in a compact way the various musical motifs and their variations. In contrast with hierarchic models, the "cue abstraction model" does not rely on deep level structures, but rather on surface cues such as change in register, dynamics, or texture (for empirical support for this model see Clarke & Krumhansl, 1990; Deliège, 1996; Koniari, Predazzer, & Mélen, 2001). The cue abstraction model was extended by Ockelford (2004), who suggested that the degree of "derivation" of one chunk from another depends on the salience of their common features. According to his "zygonic model," the context provides salience cues that are disrupted when the chunks are presented randomly and hence a number of juxtapositions are possible, as seen in Deliège, Mélen, Stammers, and Cross (1996). Although a number of studies have shown that categorization of themes often relies on surface cues (Lamont & Dibben, 2001; McAdams et al., 2004), listeners can identify "melodic invariants" with repeated exposure, even under the guise of changing surface features (Pollard-Gott, 1983).

The relationship between global structure and aesthetic appreciation

Karno and Konèčni (1992) asked musically trained and untrained participants to rate the first movement of Mozart's 40th symphony (written in a sonata form) versus four other manipulated versions. Ratings were based on four different scales: like/dislike, wish to own/don't wish to own, interesting/uninteresting, and overall structure. Consistent with previous studies (Gotlieb & Konèčni, 1991), listeners did not prefer Mozart's original arrangement to the other versions. Instead, there was an order effect: the version heard first was the preferred one. Surprisingly, there were no differences between musically trained and untrained listeners.

Tillmann and Bigand (1996) radically changed the structure of three pieces from different styles (Bach, Mozart, Schoenberg). Half of the musically naive participants in their study heard the original pieces, whereas the other half heard the pieces rearranged so that the order of the phrases was reversed. Participants were asked to grade each piece on 29 bi-polar semantic scales, such as happy/melancholy, restless/relaxed, and so on. Although there were clear differences between the three pieces, there were no such differences between the original versions and the rearrangements. Furthermore, after the experimental trials, participants received an explanation of how the musical stimuli were created, and were asked to judge whether they had heard the original version or a manipulated one. Only 43% of those who heard a rearranged version identified it as such.

Eitan and Granot (2008) examined listeners' aesthetic judgments in an experimental design in which the inner form of masterworks was altered by creating hybrids from two unrelated works. Musically trained and untrained listeners unfamiliar with these sonatas rated the intact opening movements of Mozart's piano sonatas, K. 280 and K. 332 (sonata-form movements sharing meter, tempo, key and tonal structure, but considerably differing in thematic material) versus hybrids, which mixed sections from the two movements. Listeners performed the tasks twice, once after a single hearing, and then again after week-long exposure to both versions. Results in both experiments showed no significant preference for the original masterworks, even after repeated hearings. Moreover, musical training tended to enhance preference for the hybrid over Mozart's original, particularly after repeated exposure. These results concur with findings that large-scale structure in music is secondary to the local structure. It complements experimental results showing that not only structural convention, but also "inner" form and the supposedly organic unity it entails, may not take part in listeners' aesthetic evaluation of musical works.

Although these studies taken together present a relatively coherent picture stressing listeners' limited sensitivity to global structure, some methodological concerns may be raised. First, sensitivity to overall structure requires much more than a single hearing, thus limiting the validity of the results reported in Karno & Konèčni (1992) and in Marvin & Brinkman (1999). Second, sensitivity to overall structure is not necessarily consistent with aesthetic judgment or rating of interest and pleasantness (Karno & Konèčni, 1992; Tillmann & Bigand, 1996). Third, it may require substantial, and possibly explicit knowledge of Western musical syntax. In addition, such sensitivity may come to the fore only when all musical cues are present, including the harmonic, thematic, and even performance cues. Finally, it may be limited to musical forms and styles that set out explicitly to outline a large-scale structure such as the sonata form.

The sonata form

The sonata-allegro form represents to a large degree the culmination of synthesis between thematic material and tonal thinking (Rosen, 1980). Rosen viewed the sonata as a "process," stressing the variety of the procedures to be found in this form, which made it difficult, if not futile, to provide a clear plan for it. Nonetheless, more recent theoreticians such as Hepokoski and Darcy (2006) have taken up the challenge to provide some conventional pattern types for the sonata, stressing the composer's perspective. They describe the sonata form as a set of goals the composer chooses to articulate, creating for each sonata its unique path through a complex of compositional decisions, creating a dialogue with the web of a hypothesized selection of compositional norms. Although the degrees of freedom for each choice are varied, within the late eighteenth-century style some choices were more frequent or normative than others. This view has important cognitive ramifications since it reinstates the possibility, at least for those familiar with this music, of creating an abstract mental scheme that captures some of the common features underlying the different realizations of this form.

Webster (2001) defines the sonata-allegro form as a combination of a three-part sectional arrangement of ABA within a two-part tonal plan. The first tonal area includes the *exposition* in which the main thematic material is presented. Harmonically, the exposition establishes the tonic key, and then, through what Hepokoski and Darcy call a "series of energy gaining modules" (more traditionally known as the bridge), modulates to the secondary key. In late eighteenth-century sonatas this was most often to the dominant in the major mode sonatas, and to the major mediant III in the minor mode sonatas. The first secure perfect authentic cadence in the new key is taken as the important structural and perceptual goal of the exposition (Hepokoski & Darcy, p. 18). Rhetorically, the exposition lays out the themes and textures that serve as a reference point for subsequent re-interpretations of these materials in the development and in the recapitulation. The *development* does not usually present new materials, but rather modifies one or more ideas presented in the exposition in various ways. This is often accompanied by a heightened sense of tension, achieved through techniques of motivic elaboration including compression, contrapuntal texture, and truncation. The increase in complexity and disjunction is accompanied by harmonic instability brought about by fast modulations to relatively distant tonal areas.

The return to the tonic key at the beginning of the *recapitulation* is an important structural point, often heightened by a "double return" of both the tonic key and the main thematic material. The recapitulation's main function is to resolve the tonal opposition between the first theme or group of themes and the second theme presented in the exposition. A satisfactorily perfect authentic cadence on the tonic, paralleling the same structural point in the exposition but now articulated on the tonic key, is taken to be the structural goal towards which the whole movement has been driving.

As Kamien notes, (1988) the sonata form should not be conceived as a "rigid mold" into which the composer pours his musical ideas, but rather a set of principles that "serve to shape and unify contrasts of theme and key" (p. 219). Moreover, this form also creates, beyond these various sections, themes, and keys, an overarching curved contour of tension, which may hold the overall structure together and may turn to be cognitively as important as other structural considerations. Musical forms broadly comparable to the sonata in terms of some large-scale plan or structure can also be found in non-Western music (e.g., the Indian *kīrtanam* or the Japanese *jo-ha-kyū*). Nonetheless, the opposite notion of creating a whole piece by concatenating closed and largely independent musical segments is much more prevalent and typical of many musical forms. One such example is the "mosaic" form of composition in Arabic music (Cohen, 2006). This can be related to the overall "musical ideal" in many non-Western musical cultures, which stresses the importance of the immediate context, and the richness and complexity of the moment, rather than the overall form (Cohen & Granot, 1995; Sloboda, 1985).

Rationale and approach in the current study

In the present study we focused on two sonata-allegro movements – Mozart's piano sonata K. 570/I in B flat major and Haydn's piano sonata, Hob: XVI-34/I in E minor – which represent the structural principles of the sonata form and are relatively less well known within the classical literature (as compared to the sonata-allegro movement in Mozart's 40th symphony in the Karno & Konèčni 1992 study). In the first part of our two-part paper, we will present the results obtained in the group of listeners who responded to the Mozart sonata. The results obtained in a subsequent study with another group of subjects who were exposed to the Haydn piece will be presented in a follow-up paper. The minor-mode Haydn sonata provided an opportunity to examine whether contrasts of mode facilitate sensitivity to global structural relationships. Sections from these movements were presented to two groups of participants in a musical puzzle task, in which listeners had ample opportunity to listen to the different sections and to be actively involved. We used an expressive human recording rather than a deadpan Midi rendition so as to have participants perform the task under the most favorable conditions. Presenting rich musical materials performed expressively leads to gain in ecological validity with some compromise in experimental control.

We are aware that some concerns regarding the ecological validity of the puzzle task have been raised in the past. Clearly, this task does not mimic real listening conditions, although the idea of juxtaposed pieces, or excerpts from pieces that can have various temporal arrangements, is found in different musical genres from the eighteenth century to current music medleys. The question of whether or not the puzzle task is a valid methodology to reveal valuable information about global music processing is an empirical one. That is, if a null result is obtained, suggesting that the arrangement of the presented sections shows no constraints on proposed solutions, then one would be unable to conclude whether this null result is due to the invalid methodology or to the listeners' insensitivity to global structure cues. In contrast, any positive findings would suggest that this methodology could complement other methodologies, and does add some valuable information.

Given the studies described above, we hypothesized that musically trained participants familiar with the sonata form would be able to reconstruct the sections into the original works. In contrast, participants with no formal training and no explicit acquaintance with the sonata form would be unable to do so. However, in contrast with previous studies we do not focus solely on the rate of success. Rather, we propose an exploratory analysis, in which we look for recurring patterns in participants' responses in order to outline what type of information actually is extracted in music listening, under the conditions of the puzzle task. The analysis is based on traditional statistical analysis of participants' solutions as compared to chance, as well as some new methodological and statistical tool-sets as described in the methods section.

Methods

Participants

Eighty-seven participants (45 males, 42 females; aged 17–64) with a wide range of formal and informal musical training participated in the study. Twenty-nine were "musically trained" (7 or more years of formal musical training $M = 11.5$, $SD = 4.3$), with an explicit familiarity with the sonata form. Thirty-five participants had very little or no musical training ($M = 0.54$, $SD = 0.88$, range 0–2), and 23 participants had learned to play a musical instrument for 3–6 years ($M = 4.04$, $SD = 1.06$) but had no formal theory studies, and no explicit familiarity with the sonata form. Pianists and composers, suspected of knowing the test piece well, were excluded a priori from this study. About half the participants performed the task as part of the requirement for their music cognition course, and the rest were paid approximately $12 for their participation.

Musical stimuli

Table 1 presents the structure of the 1st movement of the Sonata K.570 in B flat major by Mozart, written in a sonata-allegro form. The table indicates the sections according to traditional musicological analysis of form, which served as a main guideline in dividing the sonata into the sections presented to the participants. This sonata was chosen because it conforms to the traditional sonata form and it has mostly clear sections, separated from each other by rests, enabling us to perform a smooth dissection of the material into its various parts. The three sections in which the transition did not include a rest in the original work were the two transitions from the second theme to the closing theme in both the exposition and the recapitulation (segments 3 & 9), and the end of the development (segment 6). Note that in order to distinguish the first theme of the exposition from that of the recapitulation we cut the development on the three measures before its ending, providing what seemed to us a to be very easy cue as to which of the two sections should open the piece. The different sections provide a good example of "variety within unity," with various motifs shared between sections but quite dramatic changes in mood, register, texture and harmony – as seen in the right-hand "remarks" column of Table 1.

We used Sound-Forge5 to edit the recording of the piece as performed by Mitsuko Uchida (Philips 422–517–512 "Complete Mozart Edition"). The total duration of the recording, omitting the repeat of the exposition, is 4'07. No manipulations were made on the recording (including no use of fade-in or fade-out), except cutting the sections.

Task and procedure

Participants received a disc with 10 tracks, separated from each other by two seconds of silence. The two-second separation was inserted to discourage use of very local surface-feature puzzle-solving strategies (see also "Some methodological issues" in the "General discussion" of our second study). All discs had the same intentionally designed order. The first of these tracks was the modulatory section of the bridge of the exposition, beginning with a *forte* dramatic chordal gesture. We assumed that despite order effects already noted in the literature (Karno & Konèčni, 1992), the tonal and rhetoric nature of this section would hint at the inappropriateness of using it as the beginning of the reconstructed piece. Note that the "Distance score" that we used to compute a measure of distance between the composer's solution and subjects' proposed solutions (see "Distance score analysis" under "Results" below) indicates that the order of segments on the CD was as distant from the correct order as a random permutation. In the case of the group of trained musicians and the score of "Arrow of time" measure (see Figure 3b for details), even especially "distant." We re-address the question of order effects in the "General discussion" section of our second study). Listeners could either take the disc home (N = 48), or perform the task without time limit in the presence of the experimenter (N = 39). They were presented with the disc and a form which began with instructions as follows:

> In the disc you have received there are 10 sections from a musical piece. They are presented in a random order. Try to re-order them into a musically logical and coherent succession. Try to disregard the short silence between the tracks and listen to the order you have chosen as a whole piece.

They were then asked to fill in a form regarding their age, gender, musical training, absolute pitch abilities, and listening habits. Following the questionnaire, they received a table in which they were asked to insert the new order they had chosen. Finally, they were asked to describe freely how they solved the task and to report the time required for completing the task.

Table 1. Mozart K. 570: Structure of the sonata-allegro 1st movement

Function in Sonata Form (mm. = measure no.)	Track in disc (duration in seconds)	Initial harmony		Final harmony		Features possibly relevant to perception of local and global structure
		Local harmony	Global function	Local harmony	Global function	
Exposition: 1st group. mm. 1-20	7 (0:23)	B flat	I	I	I	Begins with undulating tonic chord in unison texture typical of opening gestures.
Exposition: bridge. mm. 21-40	1 (0:23)	Gm	Vi	C	V/V	Sudden change of harmony + dramatic gesture of repeated descending octaves.
Exposition: 2nd group. mm. 41-69	10 (0:32)	F	V	F	V	Undulating motive in the bass + a new counter- subject in the higher register. Ends on a gestural closing trill resolving on the 1st. beat of the closing section.
Exposition: closing. mm. 69-79	6 (0:14)	F	V	F	V	Typical V-I cadential gesture. Ends metrically open on the descending octave gesture.
Development : 1st part. mm. 80-100	4 (0:26)	D flat	III^b	D	III = V/VI	Octave gesture provides a dramatic shift to A flat. Thematically based on the bridge with more dramatic modulations.
Development: 2nd part. mm. 101-132	9 (0:34)	G	V/V	F	V	Thematically based on 2nd theme.
Recapitulation: 1st group. mm. 133-152	8 (0:28)	B flat	I	B flat	I	Repeats the first theme as in the exposition
Recapitulation: bridge. mm. 153-170	5 (0:21)	Gm	Vi	F	V	Repeats the materials of the Exposition with small necessary tonal modifications
Recapitulation: 2nd group. mm. 171-199	3 (0:32)	B flat	I	B flat	I	Repeats the materials of the Exposition with the necessary tonal modifications
Recapitulation: closing. mm. 199-209	2 (0:10)	B flat	I	B flat	I	Repeats the materials of the Exposition with the necessary tonal modifications

Statistical analysis methods

In this paper we use a number of different analysis methods. Each of these methods sheds a different light on the experiment. *Histogram* analysis and *Consecutive pair histogram* analysis are straightforward methods, revealing the basic structure of the subjects' solutions as well as the pattern of their errors. *Hyper-structure* analysis is a method of evaluating the coarse-grained behavior of the subjects' solutions. Finally, *Distance-score* analysis is a way to evaluate the distance between the original piece and subjects' proposed solutions on the one hand, and random permutations on the other. This measure also enabled us to examine the influence of musical experience. Note that the different statistical methods were required because the nature of the subjects' data is complex: each subject's solution is a permutation of length 10, one of 10! = 3,628,800 possible values. This high number of possible answers calls for more sophisticated analysis, and has fewer canonical methods in the statistical literature compared, for example, with questionnaires with a limited number of choices.

In this article, *p*-values were computed either analytically or using the Monte Carlo method. The Monte Carlo method (Grinstead & Snell, 1997) computes *p*-values numerically by estimating the empirical distribution of our statistical test. Therefore, the obtained *p*-values counterbalance any complicated dependencies and allow for an accurate estimation even when an analytic formula cannot be obtained.

Histogram Analysis. The motivation of this analysis is to look at the statistics of the "raw data," or in other words at subjects' permutations. We try to identify segments that tend to be placed in a certain position at a probability higher than chance. We also examine the entropies that measure the "concentration" of the possible segments. The entropy is a way to measure whether a given segment was positioned in specific places ("concentrated") or spread over the entire possible vector of positions ("spread"). A low entropy signals a "concentrated" result.

In this analysis we calculate a table $H_{i,j}$, where $H_{i,j}$ is the number of participants that positioned the ith segment in the jth position. If participants were placing the segments in a random manner, each bin in this histogram would be distributed Bin(N,1/M), where N is the total number of participants, M is the number of segments (M = 10 in our case), and Bin(N,p) is the binomial distribution. This type of analysis points to positions along the piece that may be privileged in terms of the segments that may, or may not, be appropriate to them (e.g., the beginning and end). Based on this, we can calculate the *p*-value $p_{i,j}$, which is the probability that the actual $H_{i,j}$ would be larger than a value obtained if participants were placing the segments at random. Extreme values that suggest that a result is significantly different from a random placing – either bigger or smaller – are highlighted in the following graphs and tables (*p*-value > 0.99 or *p*-value < 0.01). A high *p*-value means that a cell was significantly more populated relatively to a random permutation. Our proposed framework is an exploratory examination of what listeners *can* perceive, rather than what they cannot. Therefore, we will not consider the significance levels literally, but rather use them to identify the more interesting parts of the data (we will also not consider here the effect of multiple comparisons). Nonetheless, as we shall see, the order of segments that participants did choose was far from random (we will quantify that). This remark applies to *p*-values also in other parts of the analysis.

Another type of analysis examines the segments rather than the positions. In order to evaluate the degree to which each segment is positioned in a chance manner (i.e., equally across all positions), we calculate the entropy of the segment histogram. The entropy is a measure of "the concentration" of these histograms using the following formula:

$$ent(j) = -\Sigma_i (H_{i,j}/N) * \log_2(H_{i,j}/N).$$

By calculating the value of ent(j) for many random permutations in a Monte Carlo simulation, it is possible to test the probability that the entropy of these random permutations is smaller than ent(j). This probability is reflected in the *p*-value.

Hyper-Structure Analysis. The motivation for this analysis is to examine groups of segments or positions rather than limiting the investigation to a single position or segment. In order to do this, we use windows that include 5 segments in our analysis. This enables us to shift from a fine-grained examination to a more coarse-grained one. The main question this analysis aims to uncover is whether sections were organized around the axis of the developmental (B) sections retaining something of the general symmetry of A-B-A' typical of the sonata form. The analysis (hereafter hyper-structure analysis) examines how many segments pertaining to the exposition (sections 1–4) development (5–6) and recapitulation (7–10) were placed in three large overlapping "windows": First window (positions 1–5), Middle window (positions 4–8) and Last window (6–10), as shown in Figure 1.

As in the histogram analysis, it is possible to compare the obtained frequency table with random permutations, and to evaluate the probability of obtaining such a frequency table by chance alone by running the Monte Carlo simulation.

Distance score analysis. Here we describe a new statistical methodology that aims to quantify the correctness of participants' proposed solutions. This methodology is based on two different distance scores that we devised. These will be termed here *edit distance*, and *arrow of time distance*. Intuitively, the edit distance of two permutations evaluates the number of editing operations required to get from the first permutation to the second. The arrow of time distance measure evaluates the degree to which permutations retain something of the correct order of the sequence as described below. More directly, using these distance measures we test whether the distance between participants' permutations and the correct order differs significantly from the distance between participants' permutations and a random permutation, and between participants' permutations and the order of the tracks they were presented with.

Edit distance score. The Levenshtein (1966) edit distance between two lists of numbers defines the minimal number of editing operations one needs to perform in order to permute from the first list to the second list. Different editing distance measures differ in the editing operations allowed, and in the costs they attach to each operation. Because the results for different distances will be correlated, we arbitrarily choose a very simple and well-known edit distance score. We do not claim that the costs or the editing operations are optimal from a cognitive perspective, but, as presented below, some interesting results emerge even with this somehow arbitrary and simple score. The Levenshtein edit distance is based on three possible operations each with an equal "cost" (of 1):

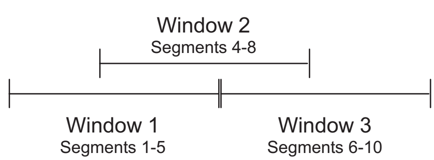

Figure 1. *Windows for hyper structure analysis.* A schematic example of the windows for hyper-structure analysis. We divided the piece into 3 overlapping "windows" of 5 segments each. The motivation for this analysis is to examine groups of segments or positions ("coarse grained" analysis) rather than limiting the investigation to a single position or segment ("fine grained" analysis).

Insertion of one number. For example: 1 2 3 4 -> 1 5 2 3 4; in this example, we inserted the number 5 in the second serial position.

Deletion of one number. For example: 1 4 3 2 -> 1 4 2; in this example, we deleted the number 3 in the third serial position of the original list.

Substitution of one number in the list. For example: 1 3 2 4 -> 1 3 3 4; in this example, we substituted the number 2 in the third serial position of the original list with the number 3.

The distance score $S_e(P_1,P_2)$ of two permutations, P_1 and P_2, is defined as the shortest number of editing operations required to obtain P_2 from P_1. Note that we only demand that P_1 and P_2 are permutations. In the intermediate steps we may pass through lists of numbers that are not permutations (as in the deletion example above, where 1 4 2 is not a permutation). For example, if P_1 = 10 1 2 3 4 5 6 7 8 9 and P_2 = 1 2 3 4 5 6 7 8 9 10, then $S_e(P_1,P_2) = 2$. The first operation is the deletion of the 10 at the beginning of P_1 and the second is the insertion of this number at the end of the sequence to obtain P_2. It is clear in this example that this is the shortest route from P_1 to P_2.

Arrow of time distance score. Another proposed measure of the correctness of a given permutation is the *arrow of time*. This measure embraces the notion that even if participants fail to place a given segment in the exactly correct serial position, they do perceive something of the correct linear order of the piece, and therefore receive some credit for semi-correct answers. For example, in the solution 1–3-4–7-8–2-5–6-10–19, segment 3 is placed incorrectly in the second serial position, yet it is placed before sections 4–5-6–7-8–9-10 – as indeed should be the case. This measure therefore may represent something of Cone's notion of "before" and "after."

To formally describe the distance score $S_a(P_1,P_2)$ of two permutations P_1 and P_2 we count the number of incorrect pairs, a pair of numbers (i,j) in which i appears *before* j in P_1 but not in P_2. By this definition, "before" does not mean sequentially, as we can see in the next example: 1 3 4 7 8 6 5 2 9 10; here, i = 3 (second position) comes before j = 5 (seventh position).

Formally speaking: $S_a(P_1,P_2) = |\{(i,j) \mid P_2^{-1}(P_1(i)) > P_2^{-1}(P_1(j)) \text{ and } i < j\}|$ where P(i) is the segment that the subject placed in the i's position, and $P^{-1}(i)$ is the inverse permutation: $P^{-1}(j)$ is the position of the j's segment in P, formally $P^{-1}(P(i)) = i$.

One can easily verify that the maximal score is M*(M–1)/2 for a permutation of length M. As in the edit distance score, a zero score means that two permutations are identical. Here is an example to clarify the way we calculate this measure: If P_1 = 4 1 3 2 and P_2 = 4 2 1 3 then, as seen in Table 2, $S_a(P_1,P_2) = 2$.

Table 2. Example of calculation of "arrow of time distance" in two hypothetical strings P_1= 4 1 3 2 and P_2 = 4 2 1 3. $S_a(P_1,P_2) = 2$

i before j in P_1	i before j in P_2
i = 4 j = 1	yes
i = 4 j = 3	yes
i = 4 j = 2	yes
i = 1 j = 3	yes
i = 1 j = 2	no
i = 3 j = 2	no

Monte Carlo simulation of edit distance scores. Assuming we have a group of participants, each having their own solution to the jigsaw problem (their own permutation), we can calculate the mean of the distance score of these permutations and the correct order (see Figure 2a). This distance score can be based on either the edit distance score or on the arrow of time score described above.

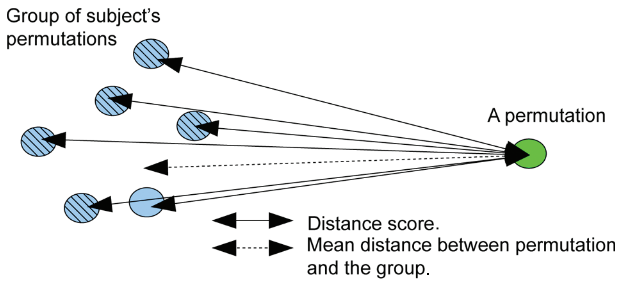

Figure 2a. *Mean distance between group and a permutation.* A schematic representation of the mean of the distance score as calculated between a given permutation and the permutations proposed by the group of subjects. For each given permutation (on the right) we calculate the distance to each of the subject's permutations (on the left), based on the mean of those distances for this specific group of subjects. Note that "a permutation" stands here for any permutation: this may be a random permutation, the correct order, or the original order on the CD.

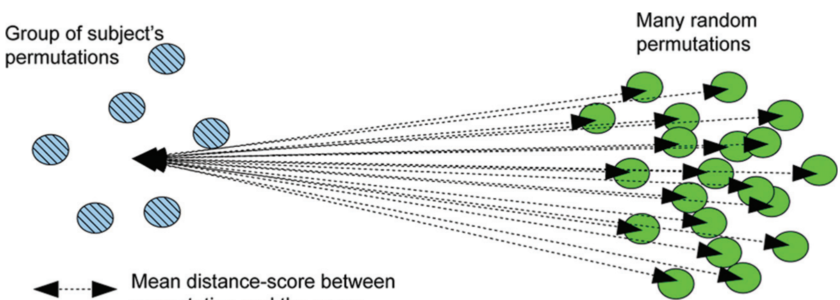

Figure 2b. *Monte Carlo simulation of mean distance score.* A schematic representation of the mean of the distance score as calculated between many random permutations created by a Monte Carlo simulation and the permutations proposed by the group of subjects. For each random permutation we calculate the mean distance to the subject's permutations.

Similarly, we can calculate the mean of the distance score between the participants' permutations and the original order on the CD. This can be compared with a random permutation in order to evaluate whether the mean distance between the actual permutations and the correct order (or the original order) is significantly different from a random permutation. Rather than using a single random permutation, we create a large number of random permutations using the Monte Carlo technique. For each of these random permutations we calculate the mean distance to the group (see Figure 2b). Now we have the random permutation mean distance distribution, and we can calculate the mean, standard deviations and *p*-values.

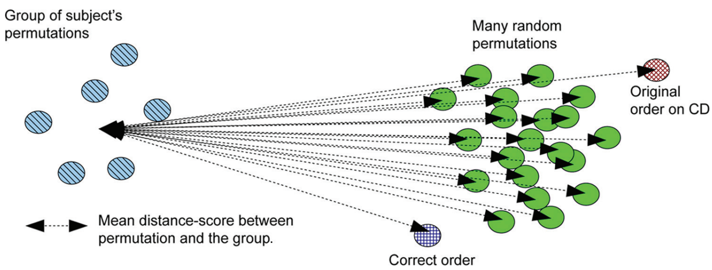

Figure 2c. *Comparing random permutations and the original and the correct order.* A schematic representation of the mean of the distance score as calculated between many random permutations created by a Monte Carlo simulation, the permutations proposed by the group of subjects, and the correct order (i.e. as it appears in the original piece).

By comparing the mean distance of random permutation to the mean distance of the correct (or the original) order on the CD, we calculate significance values (see Figure 2c).

Consecutive pair analysis. In this analysis, we simply count the number of participants that positioned two segments consecutively. These actual probabilities are compared to the frequency table obtained by running the Monte Carlo simulation on 10,000 random permutations.

Results

Participants who performed the task in the presence of the experimenter took 1–2 hours to complete the task. Participants who took the disc home returned it within 7–10 days, and reported investing a similar span of time. The results described below relate to the data pooled over all participants, regardless of whether they performed the task at home or in the presence of the experimenter, based on the observation that there were no differences between the two sets of data.

Overall, only two out of 87 participants provided the full puzzle solution for the Mozart sonata. Both are highly trained musicians, one of whom has absolute pitch. An additional musically trained participant provided a nearly complete solution for the Mozart piece with six correct successive items, dislocating only the bridge of the exposition (E): Theme 1 (E) – Theme 2 (E) – Closing (E) – Bridge (E) – Development – Recapitulation (R). Most participants provided concatenations of one or two correct pairs of segments.

Histogram analysis. Table 3 presents the frequency of the segments in each of the serial positions as proposed by all 87 participants. For example, 26 participants (29.9%) positioned the correct first segment in the first serial position (the beginning of their constructed piece). No participant selected the last (closing) segment for this location. As seen across the columns in Table 3, the first and last positions are indeed privileged in terms of the distribution of proposed segments appropriate for these positions. In the last position, 44.8% of the participants correctly positioned segment 10. A somewhat lower proportion of 29.9% of the participants positioned the first segment correctly in the first position. However, an identical 29.9% positioned the incorrect bridge of the exposition in this position (segment 2). There are two other interesting data points: the opening of the development (segment 5) in position 7, and the second part of the development (segment 6) in position 9.

As explained in the “Statistical Analysis Methods” section, we can calculate the entropy ent(j) of a segment in the jth position in order to evaluate the sparseness of the segment distribution. In general, it is evident that segments were not placed randomly. This extended well beyond the nonrandom placement of segment 1 and segment 10 – the beginning and ending sections (entropy p-value < .0001). The bridge of the exposition (segment 2, entropy p-value < .0001), as well as the two developmental segments (5, 6, entropy p-value < .0001) and the bridge of the recapitulation (8, entropy p-value < .01) were all placed in a nonrandom manner, albeit not necessarily in their correct position. Some possible explanations for these nonrandom placements are suggested in the discussion. Here we only point out that all of these sections (beyond the first and last) are relatively unstable, with some dramatic motivic gestures, and include either harmonic modulations or minor-mode inflections.

Hyper-structure analysis. The previous analysis suggests that, in general, participants sensed that the development sections should be placed in the middle of the piece. We verified this by the hyper-structure analysis explicated in the Statistical analysis methods section.

As seen in Table 4, correct placements of the Recapitulation sections in the last window were significantly higher than chance ($N = 194$, $p > .99$) whereas wrong placements of these segments in the first and second windows, were significantly low ($p < .01$). Note that a high p-value determines that the window was populated significantly more than a random permutation, whereas a small p-value determines that the window was significantly less populated than a random permutation. A similar pattern can be observed in the placement of the Exposition segments. The placement of the Development sections in the correct window was also significantly higher than chance. However, these sections were also relatively highly misplaced in the last third of the piece (locations 7–8-9–10). In sum, participants were sensitive to the general symmetry of A-B-A’, and placed segments on a coarse resolution in the right window significantly above chance.

Distance scores analysis. We used the tool set described under Statistical analysis methods in order to test whether musical training has any effect on the degree to which the proposed solutions resemble Mozart’s compositional choices. More specifically, we tested whether the higher the number of years of musical training, the closer the proposed solution to Mozart’s piece, and the further from a random solution or from the original order in which the sections were presented on the CD. We first divided our subject group into four sub-groups of roughly equivalent size based on their musical training. Group 1 consisted of 25 participants with no training at all; Group 2 consisted of 26 participants with four or less years of training; Group 3 consisted of 16 participants with more than four and less or equal to nine years of training, and Group 4 included 20 highly trained musicians with over nine years of training.

Based on the method described above, we estimated:

1. The mean distance between each of the four groups and the correct order.
2. The mean distance between each of the four groups and the original order on the CD (constant for all participants). We compared this distance with the distance to the correct order, and examined whether there were any order effects.
3. The mean distance and standard deviation of each of the four groups and many random permutations.
4. The p-value of the probability of the scores in (1) and (2) as compared with the scores obtained using many random permutations as in (3). We defined significance values here to be (p-value > .95 or p-value < 0.05).

Table 3. Mozart: Histogram of segment placements and entropy values

Position	1	2	3	4	5	6	7	8	9	10	Entropy
Theme 1(E)	29.9%**	16.1%	6.9%	8.0%	8.0%	10.3%	2.3%*	10.3%	5.7%	2.3%*	2.9602**
Bridge (E)	29.9%**	17.2%*	6.9%	14.9%	5.7%	11.5%	4.6%	3.4%	2.3%*	3.4%	2.8939**
Theme 2(E)	5.7%	5.7%	13.8%	6.9%	9.2%	9.2%	10.3%	13.8%	13.8%	11.5%	3.2529
Closing (E)	3.4%	2.3%*	17.2%*	11.5%	11.5%	11.5%	11.5%	9.2%	8.0%	13.8%	3.1682
Development 1	2.3%*	3.4%	9.2%	14.9%	17.2%*	12.6%	21.8%**	12.6%	3.4%	2.3%*	2.9827**
Development 2	2.3%*	5.7%	8.0%	12.6%	16.1%	10.3%	8.0%	14.9%	21.8%**	0.0%**	2.9761**
Theme1 (R)	5.8%	10.3%	9.2%	10.3%	4.6%	8.0%	13.8%	8.0%	16.1%	13.8%	3.2324
Bridge (R)	9.2%	9.2%	3.4%	4.6%	13.8%	13.8%	16.1%	16.1%	11.5%	2.3%*	3.1255*
Theme 2(R)	11.5%	17.2%*	14.9%	8.0%	8.0%	4.6%	9.2%	9.2%	11.5%	5.7%	3.2239
Closing (R)	0.0%**	12.6%	10.3%	8.0%	5.7%	8.0%	2.3%*	2.3%*	5.7%	44.8%**	2.5437**

*$p < .01$ or $p > .99$; **$p < .0001$ or $p > .9999$

Table 4. Mozart: Hyper-structure across 3 overlapping windows

	Exposition	Development	Recapitulation
Window 1[a]	**201[b] (0.9997)***	80 (0.1465)	154 (0.0060)*
Window 2	160 (0.0364**)**	**123 (1.0000)***	152 (0.0024)*
Window 3	147 (0.0003)*	94 (0.8868)	**194 (0.9964)***

[a]Window 1 = segments 1-5; Window 2 = segments 4-8; Window 3 = segments 6-10
[b]*N* segments found in the window & (*p*-value)

Table 5a. Mozart: Edit distance score as a function of musical training

Group (years of training)	N	Mean distance to Mozart	Mean distance to original CD order	Mean distance to random permutations
1 (0)	25	8.24	8.44	8.29
2 (4 ≥ Y > 0)	26	7.62*	7.96	8.28
3 (9 ≥ Y > 4)	16	7.50*	7.81	8.28
4 (Y > 9)	20	5.45**	8.00	8.28

*$p < .05$; **$p < .0001$

Table 5b. Mozart: Arrow of time distance score as a function of musical training

Group (years of training)	N	Mean distance to Mozart	Mean distance to original CD order	Mean distance to random permutations
1 (0)	25	21.56	23.92	22.51
2 (4 ≥ Y > 0)	26	21.04*	22.42	22.51
3 (9 ≥ Y > 4)	16	17.06*	23.69	22.52
4 (Y > 9)	20	12.85**	27.25	22.52

*$p < .05$; **$p < .0001$

As seen in Figures 3a, 3b and Tables 5a, 5b the distance score (whether "edit distance" or "arrow of time") to the correct order diminishes with musical experience. This is equally true in both distance measures used. However, in the edit distance score this relationship is consistent across all groups of participants, so that even a small amount of training (> 0 and ≤ 4 years of training) makes a difference, whereas in the arrow of time score, participants perform better only if they have more than 4 years of training. Group 1 (0 training) was not significantly better than random, at least when using these simple distance scores. It is also noticeable that the distance score of the original order did not change in any consistent way as a function of training, although it was slightly larger (*p*-value 0.925) in the case of the highly trained participants (Group 4). Note also the non-significant difference between the random permutations and the original order of the CD (for almost all cases), suggesting this order was a good base line. Figure 3b also shows the intrinsic big distance schematically between the order on the CD and the correct order.

In order to test whether the results we obtained are meaningful, we also ran the same analysis using a different type of subdivision of our participants' group, in which there was no reason to assume a priori any differences in performance: males ($N = 39$) versus females ($N = 45$).

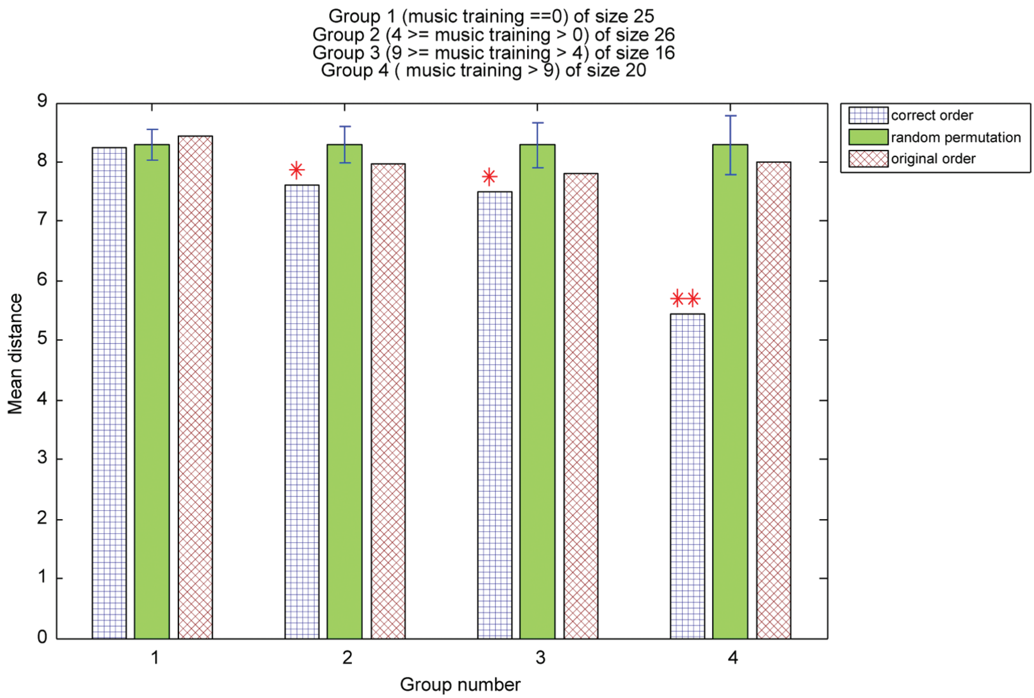

Figure 3a. *Edit distance score grouped by music training*. Edit distance score between subjects' proposed solutions and (1) the correct order of the Mozart sonata, (2) a random solution, and (3) the order of the tracks as presented on the CD ("original order"). The edit distance score is compared across four groups of subjects grouped by musical training (see Figures 2a–2c).

The distance arrow score for both males and females is significantly different from a random permutation (males: mean = 18.26 *p*-value = 0.006, females: mean = 18.69 *p*-value = 0.032) and, more importantly, no different from each other.

One immediate concern that can be raised with regard to these proposed new measures is whether the effects we see are entirely driven by the sensitivity we have already noted to the first and last segments of the piece, or whether they capture some additional structure in the listeners' responses. In order to test this, we compared the mean arrow of time distance between the solutions proposed by each of the four sub-groups (divided on the basis of years of training as in Table 5b and Figure 3b) and random permutations that had the first and last segment positioned correctly *with the same probability* as the probability of the group. Using this manipulation, we canceled out the effect of these specific segments on the score. Trends seen in the previous analyses were maintained with an increased markedness of the effects of training. The distance to the original is significantly different from randomness only for the most highly trained group ($p = .0003$) and marginally so for the group with 4 or more years of training ($p = .077$).

To conclude, using distance scores, a simple and new statistical and methodological approach, we could show that our participants were better than random in positioning the segments. Moreover, musical experience positively affected their performance.

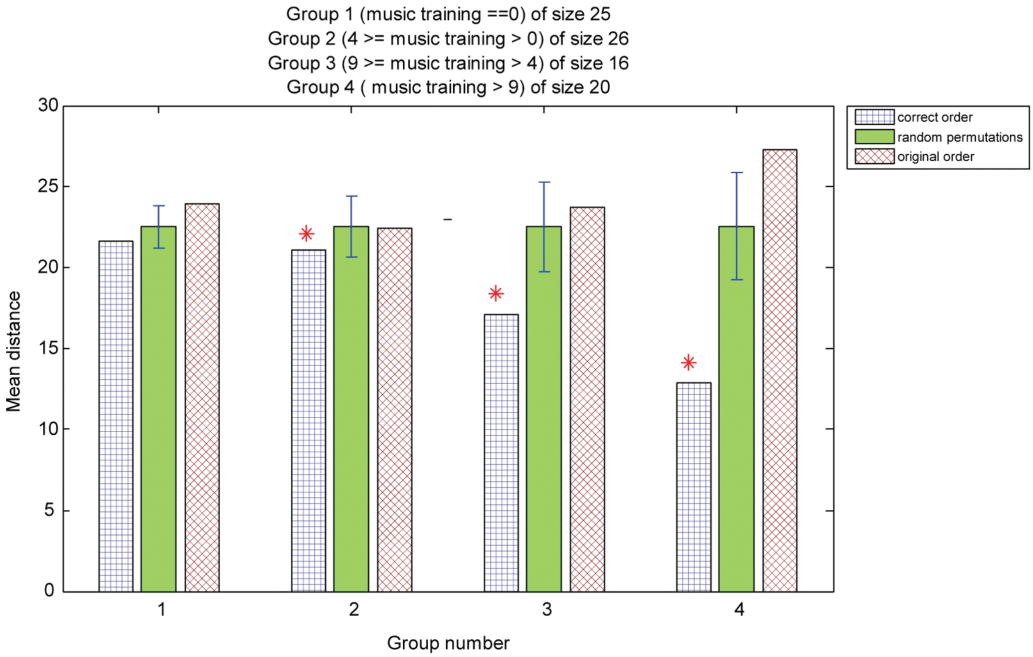

Figure 3b. *Arrow distance score grouped by music training.* Arrow distance score between subjects' proposed solutions and (1) the correct order of the Mozart sonata, (2) a random solution, and (3) the order of the tracks as presented on the CD ("original order"). The arrow distance score is compared across four groups of subjects grouped by musical training.

Consecutive pair analysis. The next analysis examines sensitivity to local rather than global cues. Table 6 and Figure 4 present the frequencies of positioning segments consecutively. These actual probabilities are compared to the frequency table obtained by running the Monte Carlo simulation on 10,000 random permutations. The probabilities of obtaining frequencies significantly different from the frequency values obtained by this simulation are also presented (p-value < .01 or p-value > .99).

The pairing of the frequent pairs can be explained by three different principles. Pairs 1–9 and 6–3 exhibit pairing on the basis of thematic derivation. A second principle that seems to drive the selection of consecutive pairs is sensitivity to conjunction cues in those sections that did not end on a pause. This includes the sections of the second part of the exposition leading to the recapitulation (pairs 6–7) and the incorrect pairing of 9–7. Finally, an especially interesting correct pairing is that of the two parts of the development (5–6), possibly indicating sensitivity to the heightened tension common to these two sections. Despite the plausibility of these explanations, we are aware that they are post hoc and require further corroborations. Furthermore, the explanations are not parsimonious, inasmuch as a number of mechanisms are proposed in order to explain the various results.

In contrast, all rare pairs can be explained on the basis of a single principle: all include at least one segment that did not end or begin with a pause, suggesting that these surface cues did serve to constrain the possible solutions – as we indeed hypothesized.

Table 6. Mozart: Extremely rare (p-value < 0.01) and extremely frequent (p-value > .99) pairs.

After→	1	2	3	4	5	6	7	8	9	10
Before→1		13.8%	6.9%	4.6%	9.2%	6.9%	10.3%	11.5%	24.1%$^{+}$	10.3%
Before→2	17.2%		11.5%	16.1%	11.5%	5.7%	3.4%*	11.5%	8.0%	11.5%
Before→3	6.9%	3.4%*		18.4%	10.3%	16.1%	8.0%	8.0%	4.6%	12.6%
Before→4	11.5%	9.2%	6.9%		13.8%	5.7%	10.3%	12.6%	8.0%	8.0%
Before→5	4.6%	9.2%	16.1%	6.9%		29.9%$^{+}$	4.6%	10.3%	13.8%	2.3%*
Before→6	5.7%	3.4%*	19.5%$^{+}$	10.3%	11.5%		26.4%$^{+}$	8.0%	2.3%*	12.6%
Before→7	5.7%	10.3%	8.0%	6.9%	6.9%	6.9%		13.8%	8.0%	19.5%$^{+}$
Before→8	10.3%	5.7%	12.6%	11.5%	18.4%	11.5%	9.2%		13.8%	4.6%
Before→9	3.4%*	6.9%	5.7%	11.5%	10.3%	10.3%	19.5%$^{+}$	8.0%		18.4%
Before→10	4.6%	8.0%	6.9%	10.3%	5.7%	4.6%	2.3%*	6.9%	5.7%	

*$p < .01$; $^{+}p > .99$

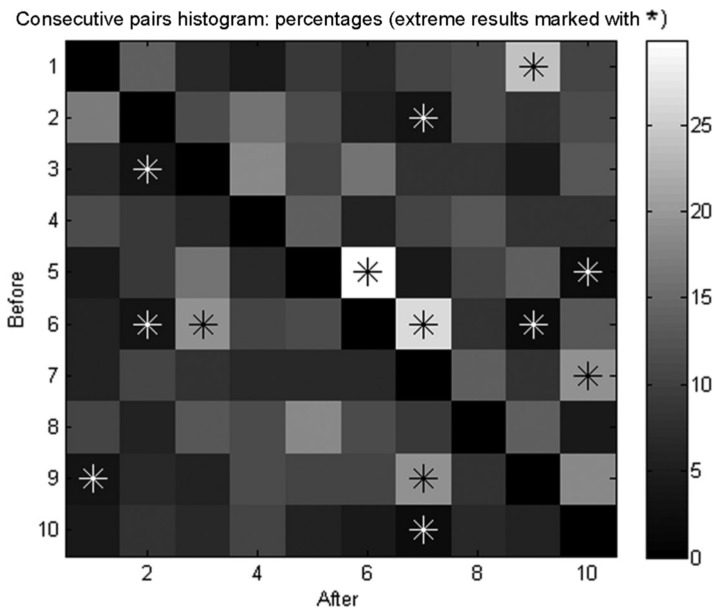

Figure 4. *Mozart consecutive pairs analysis.* Consecutive pairs analysis in Experiment 1, indicating the percentage of placing segments consecutively. For example, the percentage of the correct placement of segment 6 after segment 5 is relatively high (bright along the percentage scale). Extreme values—either relatively frequent ($p > .99$) or relatively rare ($p < .01$)—are indicated by asterisks.

Discussion

As expected from the literature, participants performed poorly on the task overall, with only two musically trained participants solving the entire puzzle. However, as described in the Introduction, using exploratory methods of analysis we can point to recurring patterns in participants' proposed solutions, which seem to provide insights about local as well as global music processing. We first present an overview of these results, and then expand on interesting or problematic points raised in the various analyses.

First (see entropy segment's distribution in the Results section), participants tended to position certain segments in a nonrandom manner. These segments included, as expected, the opening and the ending of the piece, but they also included the non-stable segments of the bridge sections, and the developmental sections. Moreover, we were able to show that participants had a general notion of the overall musical form: they positioned the exposition segments at the beginning of their solution to the puzzle, the developmental sections in the middle, and the recapitulation segments at the end significantly better than chance. Thus, participants were sensitive to the general symmetry A-B-A' of the sonata form – showing some awareness of the overall global form. Finally, consecutive-pair histograms of the participants' permutations revealed a high sensitivity to thematic similarity and derivation, as well as to surface cues as revealed in sections ending with no pause. As seen in previous studies, sensitivity to harmonic cues seemed to be local rather than global.

We now turn to a more detailed look at the results and their interpretation. One interesting aspect of the results was that the nonrandom placement of segments was clustered not only around the first and last segments as expected, but also around the nonstable modulatory segments of the bridge of the exposition and the two developmental sections. The tendency to place the bridge of the exposition at the beginning of the piece is most probably an order effect, since this section was presented first on the disc the participants were given. This is especially significant since the bridge section opens in a gesture of octave-jumps in *forte* on a minor chord and a dramatic pause on the third beat. These, along with the fast harmonic motion following this gesture, are more typical of a transitional section than of an opening one. Nonetheless, it was placed at the beginning of participants' proposed solutions as frequently as the correct first theme, stressing the importance of the temporal order in which musical materials are presented to listeners.

Yet this bias of the original order had only a limited effect. The mean distance score of the participants' permutations and the correct order was always smaller than the mean distance score of the participants' permutations and the original CD order. That is, participants were not completely misled by the order of presentation. Although seen in all the graphs presented in the distance score section, it is especially evident in the case of the highly trained group (> 9 years) and in the arrow distance score. In this case, the distance score was significantly *small* in the correct order and significantly *high* in the original order.

Other segments showing low entropy (low randomness in placement) were the two parts of the development. Participants rarely placed the first and most unstable part of the development (Segment 5) – near the beginning (positions 1–3) or end (positions 9–10). Rather, they tended to place it in location 7, possibly signaling the highpoint of the piece around its golden section (Konèčni, 2005). The low entropy of the second part of the development (segment 6) seems to be driven by the relatively large number of participants erroneously placing it in the next to last position (location 9). This section, like the correct next-to-last 9th segment, ends on an incomplete gesture of a cadence leading to the tonic, B flat. Moreover, it ends on a clear ritardando, heightening the expectation for resolution and stressing the structural importance of this point. It is also significant that the two parts of the development (5–6) were paired together on a higher than chance level (Table 6). This pairing seems to point to participants' sensitivity to the heightened tension common to these two sections, and occurs despite the local cadence and the pause at the end of the first part of the development. This heightened tension is manifested in the unstable harmony and minor mode inflections of the structure of the piece itself.

However, it is also evident if one listens attentively to certain performance cues such as the tempo and dynamics used by the performer to highlight the peak tension of the piece.

The above mentioned sensitivity to tonal instability contrasts with the insensitivity to the global harmonic structure. The most significant structural point for examining this is at the end of the bridge (on C major as V/V), which leads to the second theme on the dominant key of F major. The cadence at the end of the bridge sounds complete if processed locally, but incomplete and strongly pointing to the second theme if processed globally. As in Tillmann, Bigand, & Madurell (1998), listeners seem to have processed this cadence locally, with only 11.5% of the participants correctly pairing the end of the bridge with its ensuing second theme. This percentage is not significantly different from a chance pairing ($p > 0.3$, using the exact Binomial test), and this was equally true of the musically trained participants (7 or more years of training) as of the untrained group ($p = 0.37$ using Fisher's exact test).

To conclude the first study, listeners' ability to categorize sections as stable versus unstable, their sensitivity to rhetoric cues of opening and closing gestures, and their sensitivity to melodic relationships seem to underlie their performance in the current study. Importantly, these three factors together were sufficient to lead participants to "compose" structures which, although different in detail, share the general A-B-A' structure in an above-chance manner. As in previous studies, we found no evidence for integration of harmonic information into a global structure. In order to further test the validity of these results and the analysis methods developed in this study we ran a second study, the details of which appear in the second part of our paper. In this second report we will also further discuss some methodological issues common to both studies.

Acknowledgements

This study was supported by the Hebrew University's Authority for Research and Development grant to the first author.

References

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*, 100–130.

Clarke, E. F., & Krumhansl, C. L. (1990). Perceiving musical time. *Music Perception*, *7*(*3*), 213–252.

Cohen, D. (2006). *East and west in music* (2nd ed.). Jerusalem: The Hebrew University. Magnes Press. [In Hebrew.]

Cohen, D., & Granot, R. (1995). Constant and variable influences on stages of musical activities. *Journal of New Music Research*, *24*, 197–229.

Cone, E. (1987). On derivation: Syntax & rhetoric. *Music Analysis*, *6*(*3*), 237–255.

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

Cook, N. (1990). *Music, imagination and culture*. Oxford: Clarendon Press.

Deliège, I. (1996). Cue abstraction as a component of categorization processes in music listening. *Psychology of Music*, *24*(*2*), 131–156.

Deliège, I. (2001). Prototype effects in music listening: An empirical approach to the notion of imprint. *Music Perception*, *18*(*3*), 371–407.

Deliège, I., Mélen, M., Stammers, D., & Cross, I. (1996). Musical schemata in real-time listening to a piece of music. *Music Perception*, *14*(*2*), 117–160.

Eitan, Z., & Granot, R. Y. (2008). Growing oranges on Mozart's apple tree: "Inner form" and aesthetic judgment. *Music Perception*, *25*(5), 397–418.

Gotlieb, H., & Konèčni, V. J. (1991). The effects of instrumentation, playing style, and structure in the Goldberg Variations by Johann Sebastian Bach. *Music Perception*, *3*, 87–102.

Grinstead, C. M., & Snell, J. L. (1997). Introduction to probability (2nd ed). *American Mathmatical Society*.

Hepokoski, J., & Darcy, W. (2006). *Elements of sonata theory: Norms, types, and deformations in the late-eighteenth-century sonata*. Oxford: Oxford University Press.

Justus, T., & Bharucha, J. (2002). Music perception and cognition. In S. Yantis (Vol. Ed.) and H. Pashler (Series Ed.), *Steven's handbook of experimental psychology: Vol. 1. Sensation and perception* (3rd ed., pp. 453–492). New York: Wiley.

Kamien, R. (1988). *Music: An appreciation* (4th ed.). New York: McGraw-Hill.

Karno, M., & Konečni, 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.

Konečni, V. J. (2005). On the "golden section." *Visual Arts Research*, *31*, 75–87.

Koniari, D., Predazzer, S., & Mélen, M. (2001). Categorization and schematization processes used in music perception by 10- to 11-year-old children. *Music Perception*, *18*(*3*), 297–324.

Lalitte, P., & Bigand, E. (2006). Music in the moment? Revisiting the effect of large scale structures. *Perceptual and Motor Skills*, *103*, 811–828.

Lalitte, P., Bigand, E., Kantor-Martynuska, J., & Delbe, C. (2009). On listening to atonal variations of two piano sonatas by Beethoven. *Music Perception*, *26*(*3*), 223–234.

Lamont, A., & Dibben, N. (2001). Motivic structure and the perception of similarity. *Music Perception*, *18*(*3*), 245–274.

Levenshtein (1966). Bindary code capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*, 707–710.

Levinson, J. (1997). *Music in the moment.* Ithaca, NY: Cornell University Press.

Marvin, E. W., & Brinkman, A. (1999). The effect of modulation and formal manipulation on perception of tonic closure by expert listeners. *Music Perception*, *16*(*4*), 389–408.

McAdams, S., Vieillard, S., Houix, O., & Reynolds, R. (2004). Perception of musical similarity among contemporary thematic materials in two instrumentations. *Music Perception*, *22*(*2*), 207–237.

Ockelford, A. (2004). On similarity, derivation and the cognition of musical structure. *Psychology of Music*, *32*(*1*), 23–74.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology*, *15*, 66–94.

Rosen, C. (1980). *Sonata forms*. New York: Norton.

Sloboda, J. (1985). *The musical mind.* Oxford: Clarendon Press.

Solie, R. A. (1980). The living work: Organicism and musical analysis. *19th-Century Music*, *4*, 147–156.

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

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research*, *61*(*3*), 15–174.

Tillmann, B., & Bigand, E. (2004). The relative importance of local and global structures in music perception. *The Journal of Aesthetics and Art Criticism*, *62*(*2*), 211–222.

Webster, J. (2001). Sonata form. In the *new Grove dictionary of music and musicians* (S. Sadie, Ed.), Vol. 23, pp. 687–701. (2nd ed.).