Cross-cultural representations of musical shape

GEORGE ATHANASOPOULOS
University of Edinburgh

NIKKI MORAN
University of Edinburgh

ABSTRACT: In cross-cultural research involving performers from distinct cultural backgrounds (U.K., Japan, Papua New Guinea), we examined 75 musicians' associations between musical sound and shape, and saw pronounced differences between groups. Participants heard short stimuli varying in pitch contour and were asked to represent these visually on paper, with the instruction that if another community member saw the marks they should be able to connect them with the sounds. Participants from the U.K. group produced consistent symbolic representations, which involved depicting the passage of time from left-to-right. Japanese participants unfamiliar with English language and western standard notation provided responses comparable to the U.K. group's. The majority opted to use a horizontal timeline, whilst a minority of traditional Japanese musicians produced unique responses with time represented vertically. The last group, a non-literate Papua New Guinean tribe known as BenaBena, produced a majority of iconic responses which did not follow the time versus pitch contour model, but highlighted musical qualities other than the parameters intentionally varied in the investigation, focusing on hue and loudness. The participants' responses point to profoundly different 'norms' of musical shape association, which may be linked to literacy and to the functional role of music in a community.

"Music confirms what is already there in society and culture, and adds nothing more than patterns of sound." Blacking (1973, p. 54).

INTRODUCTION

IN 1972, Deregowski posed an influential question when he asked: *"Do pictures offer us a lingua franca for inter-cultural communication?"* (Deregowski, 1972, p. 82). The explanations offered by his cross-cultural comparisons (and the studies whose evidence he reviewed) have since been critiqued for an ethnocentricism that was prevalent in contemporary academia (Layton, 1981). But at the very least, Deregowski demonstrated that the perception of three-dimensional objects is influenced by cultural learning: that the depiction of perspective as one culture knows it is *not* the aesthetic preference of *all* cultures. The theory of cultural relativity of shape runs counter to prototype theory, which holds that the perception of certain 'natural categories' (Rosch, 1973) follows universal principles. Rosch's pioneering research on the perception of shape proposed the psychological primacy of particular categories—namely, square, triangle and circle. However, results from a replication of Rosch's original study by Robertson, Davidoff and Shapiro (2002), with members of the North Namibian Himba tribe, demonstrated idiosyncratic shape categorization, while a further study by Nisbett (2003) also supports the view that shape perception is culturally contingent. Comparing responses from North American and Chinese participants, Nisbett (2003) found differences in the way that the two groups described pictures, with the former group attending to dominant foreground objects and the second group tending towards a more holistic description.

The cultural relativity of shape brings up fascinating questions with regard to musical representation. The relationship between musical sound and shape is a complex matter. Musicians who are enculturated in what can broadly be referred to as a western classical performance practice are so familiar with musical notation that it is common to refer to the score as "the music". In comparison, most culturally-distinct musical practices rely far less on textual representation, if they rely on it at all. And yet, the capacity to etch or otherwise make marks acting as signs is a defining characteristic of human culture. Given the strong relationship of symbolic representation with literacy in all cultures (Schmandt-Besserat, 1992), the effect of the acquisition of literacy is an obvious factor that seems likely to influence an individual's mode of visual representation of musical sounds (Athanasopoulos, Moran, & Frith, 2011). In this paper, however, we focus our discussion on the way in which culturally-specific perceptions of shape and time may affect the representation of musical sound.

Theoretical Background

REPRESENTATION OF TIME

The spatial representation of time is not culturally equivalent. Space-time metaphors are highly conventional, apparently dependent upon 'time-moving' or 'ego-moving' perceptions of the passing of time (Gentner, Imai, & Boroditsky, 2002). However, all written languages must follow a particular direction on the page, and this aspect seems to be linked to the representation of time in two dimensions (Fuhrman & Boroditsky, 2010; Mitchell, 2004; Zwaan, 1965). In fact, evidence suggests that culturally specific spatial representations inform judgements about time even in non-linguistic tasks (Boroditsky, 2001, 2011). For example, Mandarin speakers appear to think of the passage of time as movement in a vertical direction (top-to-bottom), while speakers of English as a first language seem more likely to conceive of a horizontal, left-to-right movement. Similarly, Fuhrman and Boroditsky (2010) reported that Hebrew participants' spatial representation of time was consistent with direction of writing, passing in a horizontal right-to-left manner, while Zwaan (1965) noted that Dutch and Israeli participants located the 'past' on the left-hand side and right-hand side of the page respectively. The visual representation of time in timelines also tends to follow the cultural convention for written language (Mitchell, 2004). For example, Japanese comics, in which the story traditionally flows in a vertical right-to-left manner, are mirrored horizontally before translations are printed for European and American markets (Farago, 2007).

In contrast to visual art, music takes place in and through time. In musical performance, time is apportioned and manipulated; time can be said to be the universal raw material of musical process (Blacking, 1973, pp. 105-111). The decisions taken by an individual in order to create a visual representation of musical sounds must, then, be informed not only by culturally specific ontologies of music, but also by culturally specific notions of time.

MUSIC AND SHAPE

Ethnomusicologist Steven Feld conducted extensive research with the Papua New Guinean Kaluli, documenting and interpreting the metaphorical relationships that underpin aesthetic concepts essential to Kaluli song and poetics. Feld argues that Kaluli music theory—in terms of "logical patterns of symbolic material" (Feld, 1982, p. 16)—is inextricable from its function, namely "to activate and bring forth meaningful social relations through structural expression" (Feld, 1982, p. 16). His work reveals the extent to which cultural organization and behaviour influence both the conceptualization of, and discourse on, the abstraction of musical sound into symbolic patterns. Feld's work provides a reminder that the concept of music (and its theorisation) common to most empirical music research is not, in fact, a universal given.

However, specific acoustic parameters are commonly associated with certain visual metaphors, and empirical research supports various cross-modal matches—for example, pitch height is commonly associated with relative height on the vertical axis (Casasanto, Philips, & Boroditsky, 2003) and fast tempo may be associated with fast movement (Eitan & Granot, 2006). In other investigations, loudness has also been empirically associated with verticality (Eitan, Schupak, & Marks, 2010) and length (Carello, Anderson, & Kunkler-Peck, 1998). However, as Walker (1987b) and Sadek (1987) indicate, these mappings are not found cross-culturally. Using a forced-choice design, Walker (1987b) asked participants to match various auditory stimuli to one of four visual metaphors, predicting associations between pitch-

height with vertical placement on an x–y axis; timbre with pattern-sign; loudness with size; and duration with length represented horizontally across the x–y axis. Reponses were collected from six different groups of participants, including four Canadian Indian groups, and two urban control groups comparing musically competent and musically naïve participants. Walker (1987b) – and a replication study with Egyptian participants by Sadek (1987) both indicated differences according to musical training and cultural background. A questionnaire study conducted by Prior (2010) also indicated that background culture may lead to different perceptions of musical shape. Küssner and Leech-Wilkinson (in press) found a link between participants' musical training and the visual representations that participants generated using a real-time drawing paradigm.

The latter study aside, little research to date has examined what happens when adult participants are given free rein to create visual representations of musical sounds. Undertaking just this task, Tan and Kelly (2004) used whole musical compositions as stimuli, requesting musically trained and untrained US college participants to "make any marks" in order to represent the sounds visually. The musically trained participants demonstrated a tendency to create representations aligned on an x–y axis (henceforth referred to as Cartesian), representing time in a horizontal, left-to-right fashion, and depicting aspects of the musical surface (primarily pitch) on the vertical axis, in accordance with findings described elsewhere (Athanasopoulos & Moran, 2012; Athanasopoulos, Moran, & Frith, 2011). However, Tan and Kelly (2004) also found that musically untrained participants tended to provide pictorial representations through images or pictures telling a story. Küssner (2013) obtained comparable results from British participants.

Similar to the method of Tan and Kelly (2004), Bamberger (2005) and Barrett (2005) designed studies that permitted their participants to create notational systems of their own "invention" without being restricted to the pitch versus time model which most organised notational systems follow. In these two investigations, participants (mainly children) were asked to provide a means of invented notation for a familiar tune so that one of their peers would be able to perform the tune, without further guidance from the researcher. This resulted in a number of participants mimicking familiar notational models, while others attempted to develop executive models of notation that fitted the task. These tended to prioritise aesthetic concerns: for example, participants included performance guidelines for playing techniques, but disregarded specific note duration values. In a separate free-drawing investigation with Japanese children as participants (Adachi, 1997), a considerable proportion of responses included onomatopoeia and linguistic script to depict information about attack rate.

Beyond the investigations mentioned above, other studies to have deployed free representation techniques have focused on questions related to cognitive development, examining children's responses (Reybrouck, Verschaffel, & Lauwerier 2009; Verschaffel, Reybrouck, Janssens, & Van Dooren, 2010). As far as we are aware, no existing cross-cultural research has used a free-drawing task with adult musicians. It seems inevitable that cultural difference will result in varied depictions of time and shape in response to musical sound stimuli, but in what ways? *What is the effect of cultural background on an individual's two-dimensional representation of musical sound?*

COMPARATIVE FREE-DRAWING STUDY

To respond to this question, this study compares the freely-drawn responses of performers from distinct cultural backgrounds to musical sounds whose pitch contours were systematically varied.[1] Nettl (1985) notes that western musical practices may transform traditional societies and may replace or modify existing norms. Therefore, this study draws on evidence collected through remote fieldwork, in order to meet with participants with the least exposure to western culture. In total, 75 individuals represented three culturally distinct groups: Edinburgh, U.K. (March/April 2011); Tokyo and Kyoto, Japan (May/June 2010); and the BenaBena villages in Papua New Guinea (August 2010).[2]

Japanese traditional music is distinctive from western musical culture in that it deploys various elaborate notational systems that do not follow the standard Western notation format. Rather, the systems are either alphabetic (where one symbol stands for one pitch) or executive (directions as to how to create a sound, similar to guitar tablatures). Meanwhile, the BenaBena of Papua New Guinea do not use any written system of communication—neither for verbal literacy, nor for musical notation—and are relatively secluded from urbanisation.[3]

For these participant groups, the existing literature and research evidence point to the following hypotheses:

- Individuals literate in western standard notation (WSN) or Japanese traditional notation (JTN) will depict variations in pitch taking a Cartesian approach, using orthogonal axes to show time versus pitch, thus representing sequential occurrence (Küssner & Leech-Wilkinson, in press; Tan & Kelly, 2004).
- A substantial number of literate participants may deploy written words to describe sound events (Adachi, 1997).
- In the absence of the common point of reference provided by general (or musical) literacy, non-literate individuals will depict variation in pitch idiosyncratically with regard to the sequential occurrence of events in time.
- As indicated by Feld (1982), symbolic representation of sound events may exist at a metaphorical level in non-literate cultures: this metaphorical aspect may be reflected in an internally consistent system of representation through visual responses.

Research Design

Using a quasi-experimental design, three participant groups heard twelve musical stimuli that varied in pitch contour (Up, Down, Peak or Valley), as detailed in the Method section below, and drew responses in which they depicted the presented sound events.

The groups were as well-matched as was possible in the circumstances of remote fieldwork, taking account of approximate number, experience as a performing musician, and gender split. The age of participants varied, but all participants were adults in this study and so acted as mature representatives of their particular musical tradition. The very concept of a dictation exercise—which in some ways describes the task—is very common when learning western music and WSN, but it does not exist in Japanese traditional music, nor for the BenaBena. The groups cannot be considered to be matched with their counterparts in terms of experience in representing sound visually. However, the open-ended nature of the task was designed to compensate for some of these limitations, the payback being the richness and diversity of the data.

DEVELOPMENT OF THE SOUND STIMULI

The sound stimuli were developed according to certain constraints, in order that the same stimuli could reasonably be used for all three groups. Auditory stimuli using unfamiliar or 'unmusical' timbres or patterns were likely to disrupt participants' responses. A synthesized flute sound was selected to represent pitch variations, as flutes in some form are common to all three musical cultures.

The musical examples consisted of very basic melodic patterns. Papua New Guinean music from the Highland provinces does not follow elaborate melodic patterns other than the pentatonic scale, favouring rhythmical complexity instead. Thus the pitch relations of the melodic examples were limited to fourths, fifths and octaves and used simple contours: in our terms, either rising (Up), rising-falling (Peak), falling-rising (Valley), or falling (Down), as illustrated in Figure 1.

The stimuli were developed using Sibelius 6 software (Avid, 2009), exported as MIDI files at a tempo of 60 beats per minute, and produced using Digidesign Pro Tools 8 (DigiDesign, 2009). They were recorded in MP3 format and replayed to the participants with a Samsung K5 MP3 player through the in-built (slide-out) stereo speakers. Headphones were not used at all for any group, since they were likely to be an obstacle to participation for the BenaBena cohort, who view all music-related activities as communal and would not be familiar with using them. The participants were seated approximately one metre from the sound source, and the stimuli were reproduced at 52 dB SPL. The British and Japanese groups participated indoors, but by necessity the study took place outdoors with the BenaBena.
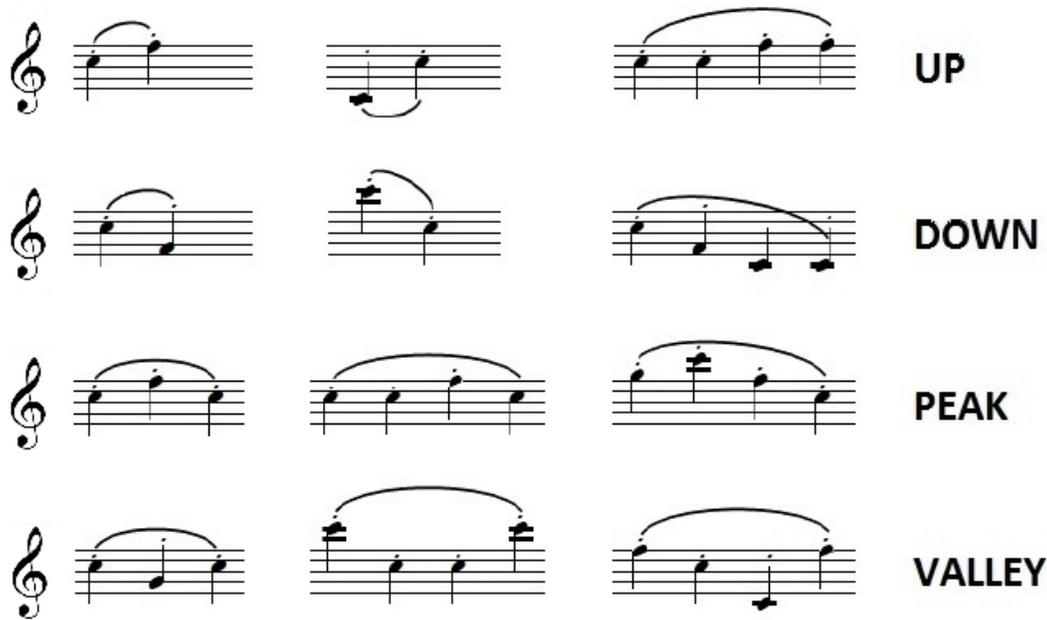
Fig. 1. Test stimuli consisted of 12 sound sequences representing four different pitch contours: Up, Down, Peak, and Valley. The sequences were presented in the same fixed, quasi-random order to participants in all three groups.

Method

PARTICIPANTS

The first group (Group A) consisted of 25 musicians of British nationality and cultural background (mean age = 23.5 years; *SD* = 4.2 years; range: 19-37 years, 10 males, 15 females; 21 right-handed, 4 left-handed). The mean age for starting a musical instrument was 6.7 years (*SD* = 1.7 years, r: 5-11 years), and mean duration of performing a musical instrument was 15.7 years (*SD* = 4.4 years, r: 11-32 years). All participants were acquainted with WSN. 28% were also acquainted with guitar tablature, 20% with jazz chord notation and 12% had performed or composed music using graphic scores. One participant (4%) also knew a traditional notational system used to transcribe Irish folk melodies.

The second group (Group B) consisted of 24 Japanese musicians with minimal or no knowledge of WSN (mean age = 47.6; *SD* = 26.4 years; r: 18-87 years; 11 males, 13 females; 23 right-handed, 1 left-handed). The mean age for starting a musical instrument was 18.6 years (*SD*=12.3, r: 3-40 years), while the mean duration of performing a musical instrument was 32.6 years (*SD* = 18.5 years, r: 8-63 years). All participants were acquainted with a form of JTN, while 28% claimed to be in a position to recognize WSN as a form of notation when they saw it but were unable to use it, and one participant (4%) was aware of the existence of graphic scores, though he had never used one in performance. It should be noted that it is practically impossible to find Japanese musicians with no exposure at all to WSN.

The third group (Group C) consisted of BenaBena tribesmen and women of the Eastern highlands region of Papua New Guinea, unfamiliar with any literary or notational script. 26 musicians (mean age by estimation = 57.1 years; *SD* = 10.5 years; r: 35-80 years; 15 males, 11 females; 26 right-handed, none left-handed) participated in the investigation. Handedness of the participants was established by asking them with which hand they used common farming tools amongst the Highlands of Papua New Guinea, such as machetes or a *cangkul* (hoe). The mean age for starting a musical instrument (from self-estimated responses) was 15.5 years (*SD* = 4.2 years, r: 10-25 years); while the mean age of performing a musical instrument was 42.2 years (*SD* = 10.5 years, r: 23-62 years). The numbers may not be accurate; performers might not have been actively performing music throughout this span. Participants provided responses for their engagement in music-making and participation in "sing-sings" and traditional community ceremonies. All but one participant reported having taken part in sing-sings since they were very young children. Recruited participants were from the BenaBena tribe and came from six hamlets (Keni, Logo, Sifu, Opeks, Siopeks, Moweto). It should be noted that music among the BenaBena tribe is a highly communal activity,

which does not separate active performers from a non-participating audience. Some individuals are regarded as the best singers or dancers, and may be called out to perform in group activities involving music and dance, but this does not exclude others from participating. Therefore, all BenaBena are considered musicians for the purposes of this study.

      The first author conducted this fieldwork and worked with Groups B and C with the assistance of local translators. For Group B, translators were recruited with the assistance of the Kyoto City University of the Arts and the Tokyo Geijutsu Daigaku. For Group C, the local schoolteacher, one tribe member who was a college student at the University of Goroka, and the first author's host, were all proficient in English and assisted with translations. For more detail, see Athanasopoulos (2013).

PROCEDURE

Participants were exposed to 20 trials, of which 12 varied according to the four pitch contour categories (Up, Down, Peak, and Valley: see Figure 1).[4] The remaining eight stimuli featured combinations of the four categories presented here. Each sound event had a total duration of four seconds, and was repeated after a four-second pause before the next trial. Participants could start drawing after the first stimulus presentation. The overall time-limit to provide a visual response to the sound event was 16 seconds. If participants were not able to provide a response, they proceeded to the next trial after this time elapsed and the next trial began. In preparation, the order of presentation of all 20 stimuli was initially randomised; this order was then held constant in its presentation to participants.

      Participants were asked to *represent the sound on paper in such a way that if another member of their community saw their marks they should be able to connect them with the sound*. Responses were drawn on A4 graph paper using ball-point pens. Before data collection began, participants were offered up to four trial runs using three sound stimuli drawn randomly from the database. On average, participants made use of two of the potential four trials.

      The BenaBena participants were not accustomed to the particular fine-motor skills associated with holding a pen and providing responses on paper. In order to communicate the concept behind the task it was necessary to draw a comparison between the acts of etching (with which all individuals were familiar) and drawing. As preparation for the study, the group were invited to draw any image from their surrounding environment on plain white paper using thick coloured markers, and then with ball-point pens. After the free-drawing investigation took place, all participants were interviewed about their experience of the task so as to record information that could contribute to the interpretation of their drawn responses.

      The participants' responses were classified according to three categories, based on their method of internal organisation and reference, and on the apparent representation of events in time (i.e. use of Cartesian pitch-versus-time axes) (Athanasopoulos, Moran, & Frith, 2011; Küssner & Leech-Wilkinson, in press; Tan & Kelly, 2004):

- *Symbolic Cartesian* (SC). Reference to sound events through abstract symbols. Pitch contour and time represented spatially on orthogonal axes. In this system, one symbol relates to another, and need not signify anything beyond this internal system. For example, note heads in WCN are abstract, symbolic representations of the sound event.
- *Iconic Cartesian* (IC). Reference to sound events through drawings that attempt to imitate the event in some respect as stand-alone icons. As opposed to abstract symbols, icons (in this classification) attempt to indicate directly and analogously some aspect of the sound event for which they stand. Icons and time represented spatially on orthogonal axes.
- *Iconic* (I). Reference to sound events through drawings that attempt to imitate the event in some respect as stand-alone icons.

     Further to these classifications, responses were separated according to their method of representing events in time:

- Left-to-Right (L-R), imitating WSN and script.
- Top-to-Bottom (T-B), imitating a majority of JTN systems and script.
- Neither L-R nor T-B.

     Classification was carried out by the first author in consultation with two independent raters. Where no majority decision was reached, the first author's decision was taken forward. The classification terms were not disclosed to participants, who drew their responses freely without such instruction.

Results

Responses were classified according to the apparent internal organisation as described above, as either SC, IC or I. Cartesian responses were further classified as either L-R (left to right), T-B (top to bottom), or as neither L-R nor T-B.

The results for each individual were examined to find the degree of reliability in their representational approach, both within contour-type and across all the stimuli. Results were then compared across groups. The classifications are illustrated below in Figures 2, 3 and 4, showing examples of each category.
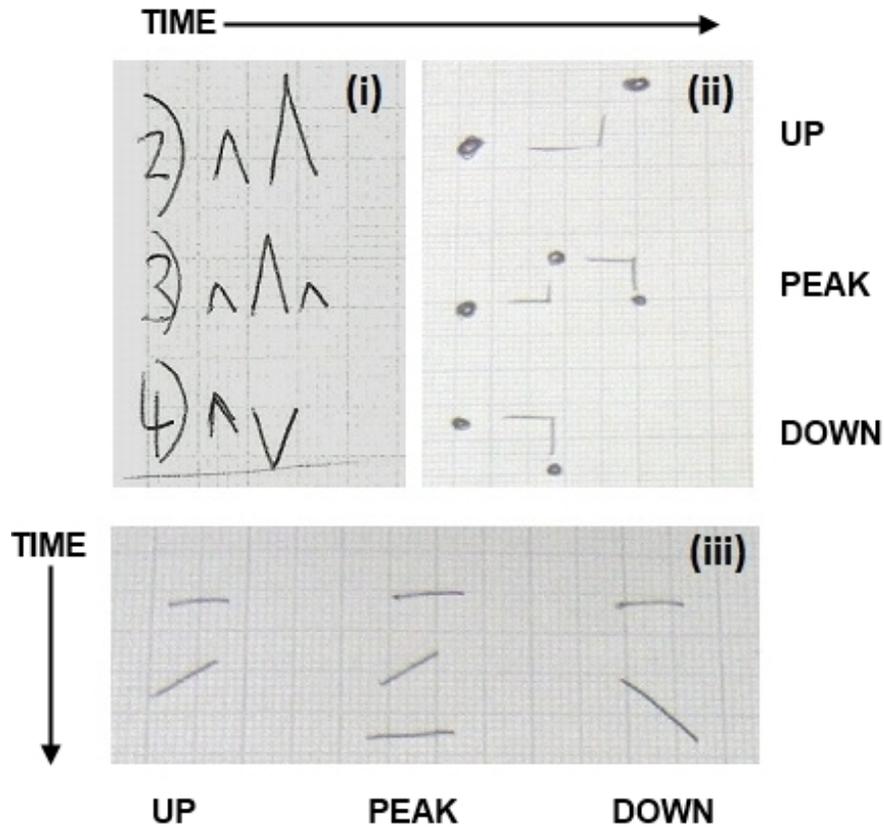

Fig. 2. Examples of three different Symbolic Cartesian (SC) systems: i. Drawn by a participant from Group A (British); ii, iii. Drawn by participants from Group B (Japanese). For i and ii, time is represented horizontally L-R, and pitch variation is shown through vertical variation. No elements of WSN have been used. Example iii represents the passing of time using a vertical axis, T-B. Pitch variations are represented by the inclinations of the strokes.
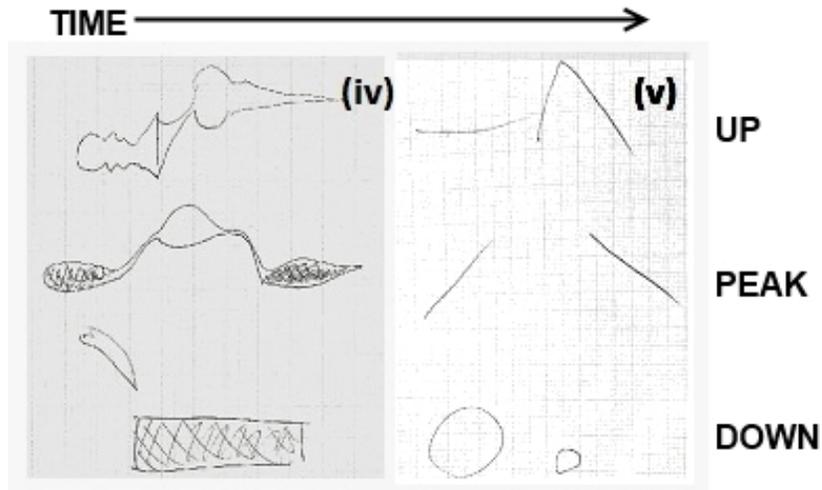
Fig. 3. Examples of Iconic Cartesian (IC) representation by a British participant from Group A (iv), and by a Japanese participant (Group B) (v). Although the drawings follow Cartesian representation (L-R), they are not internally consistent, and are therefore classed here as Iconic Cartesian (IC), not Symbolic Cartesian (SC).
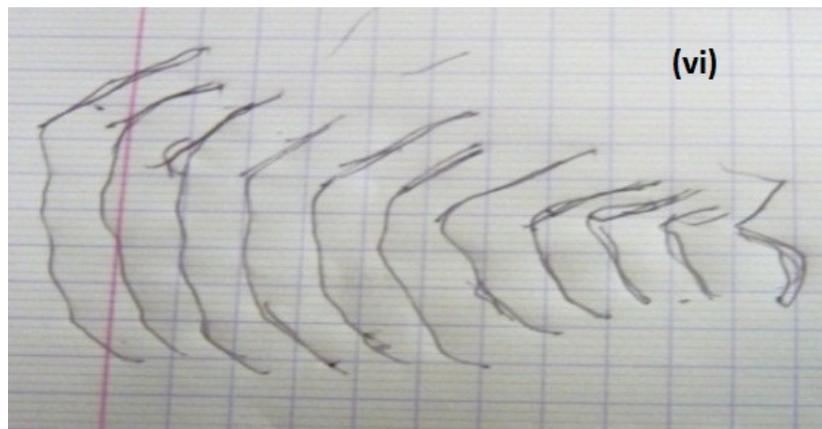

Fig. 4. This example (vi) illustrates the category of Iconic (I). The BenaBena participant has used this single image to represent all pitch stimuli (Up, Down, Peak and Valley). The depiction of the sound events does not involve a timeline. According to the participant at the follow-up interview, all of the flute sounds were represented by the lines drawn above.

RELIABILITY OF RESPONSES

According to the prescribed categories, individuals were entirely consistent in their own approach to the depiction of the sound events, both within and between contour types. For all groups (British, Japanese and BenaBena), individual participants maintained one method of representation throughout: see Table 1.

Table 1. Individual participants' preferred style of representation for the four different types of pitch contour: Symbolic Cartesian (SC), Iconic Cartesian (IC) or Iconic (I). NB: Responses were entirely congruent for the three different stimuli used within each contour group.

Contour	Group A (n=25)			Group B (n=24)			Group C (n=26)		
	SC	IC	I	SC	IC	I	SC	IC	I
Up	24	1	0	19	4	1	4	3	19
Down	24	1	0	19	4	1	4	3	19
Peak	24	1	0	19	4	1	4	3	19
Valley	24	1	0	19	4	1	4	3	19

All but one of the British participants (24 out of 25) and the majority of Japanese participants (19 out of 24) used a consistent symbolic mode of representation of the pitch stimuli, presenting abstract symbols on a timeline. Of the British participants, one-quarter employed WSN in some way: three participants provided responses entirely through WSN, and a further three incorporated some WSN elements. Two of the British participants used written text directions in addition to their drawn response.

One Japanese participant apparently experienced a dilemma with a small number of his responses, which he subsequently crossed out. Had these been included in the analysis, these would have been classed as iconic. The reason he gave for changing his mind was that since the directives were to *represent the sound so that if another member of their community saw them they should be able to connect them with the sound*, his responses would not be clear.

The BenaBena participants provided the largest variety across the three categories of response. The majority of responses (19 out of 26) were classified as iconic, using stand-alone icons depicting sound events and no obvious (linear) portrayal of time. However, four responses were classed as consistently symbolic and used a timeline, and three were classed as consistently iconic and also used a timeline. Drawing on the post-test interview data, the nineteen iconic responses could be further separated into two categories. Twelve responses (slightly less than half of all the BenaBena responses) showed some internal consistency as a system: in terms of organisation, participants attempted to draw iconic representations of the sound event's source (a flute), rather than the sound event itself. They deemed pitch variations to be unimportant; rather, according to self-reports from the post-task interviews, they focused on elements of timbre or perceived variations in loudness (though we attempted to minimize the latter by keeping the amplitude constant (Athanasopoulos, 2013, p. 238). The remaining seven iconic respondents (approximately one-quarter of all BenaBena participants) did not appear to demonstrate this level of organisation.

DIRECTION OF TIME IN THE RESPONSES

All Cartesian (SC or IC) responses are set out in Table 2 below. Of the British participants (Group A), all used an SC approach to represent events in time. The unanimously preferred timeline was horizontal L-R. Regarding the traditional Japanese group (Group B), two participants who provided SC responses opted for a vertical representation of time. One Japanese participant provided responses that did not follow Cartesian representation.

Table 2. Directionality of time as represented by SC and IC responses by group. Direction is either Left-to-Right (L-R) or Top-to-Bottom (T-B).

Direction of time	Group A: WSN	Group B: JTN	Group C: None
Left-to-Right (L-R)	25	21	7
Top-to-Bottom (T-B)	0	2	0

Finally, the BenaBena (Group C) presented the majority of their responses through iconicity and not symbolism, and without apparently depicting the sound events as occurring in time. However, seven respondents did deploy a Cartesian system and all chose to use a horizontal timeline.

DISCUSSION

Several points arise in relation to the initial hypotheses. These are addressed in turn, followed by a more expansive discussion of the results.

- Individuals literate in WSN or JTN will depict variations in pitch taking a 'Cartesian' approach, using orthogonal axes to show time versus pitch, thus representing sequential occurrence (Tan & Kelly, 2004; Küssner & Leech-Wilkinson, in press).

Regardless of cultural background, almost all literate participants in this study provided systems of representation characterised by a Cartesian (orthogonal) depiction of time versus pitch in a free-drawing paradigm, supporting this hypothesis. The one literate participant who did not do so is a Noh/Kabuki singer-actor. In the subsequent interview, she reported that she felt the stimuli had a kinetic quality, and she held that her resulting representations aimed at demonstrating this, regardless of whether they lacked clarity. Within the categories used for classification in this study, her responses were deemed *iconic,* but it is arguable that since she consciously deviated from the explicit directions for the task, her response should be disregarded.

- A substantial number of literate participants may deploy written words to describe sound events (Adachi, 1997).

A minority of the literate participants—two out of forty-nine—provided responses using text (words). They only used the words to create directives that accompanied their (SC) drawings, contradicting this hypothesis.

- In the absence of the common point of reference provided by general (or musical) literacy, non-literate individuals will depict variation in pitch idiosyncratically with regard to the sequential occurrence of events in time.

As expected, the majority of participants unfamiliar with notational systems provided iconic responses and without a timeline, supporting this hypothesis, in accordance with Tan and Kelly's (2004) findings regarding participants unfamiliar with musical notation. Tempered with the reports from the participant interviews, however, the findings from this study suggest that nearly half of those participants who created idiosyncratic systems did aim to create internally consistent representations.

- Since symbolic representation of sound events may exist at a metaphorical level in non-literate cultures as indicated by Feld (1982), this may be reflected in an internally consistent method of representation through visual responses.

Following the discussion of the previous point, non-literate individuals did not deploy internally consistent symbolic systems in the manner of their literate counterparts, but still produced a considerable proportion of internally consistent iconic responses (12 out of 26 participants). We discuss this further below.

Further findings

Other findings beyond these predictions revealed that all participants, regardless of cultural background, literacy or type of musical training, were able to provide coherent responses with varying levels of organisation in a free-drawing investigation; and regardless of group, all participants who completed these tasks were highly consistent in their manner of depicting the sound events and variations.

All musicians were able to provide invented notational systems in order to depict sound. This suggests that everyone has the ability to associate musical sounds with some form of visual representation. Although many participants were sceptical about the value of the task presented to them, all (with three exceptions due to mistranslation) were able to provide graphic representations of sound, even where the idea to represent sound visually was an alien concept, as was the case with the BenaBena.

The resulting representations are surprisingly similar between groups of different cultural backgrounds, due to the wide adoption of the SC method of representation. A large number of the entire participant population (regardless of origin or type of musical training) tended to represent sound in a linear, left-to-right axial representation resembling analogue notational systems, with time located horizontally. The linear representation of music may have its roots in literacy, since the latter provides participants with a timeline of reference and an axis on which to put responses. However, this does not account for the small number of BenaBena who used the IC mode of representation, depicting the passage of time horizontally. One could argue that this minority may have been responding to what they thought

that the Western investigator perceived as 'correct', mimicking a script that they may have seen elsewhere (perhaps even by watching the hands of the investigator as he took occasional notes). An alternative theory may examine the idea of cross-cultural tendencies towards the linear representation of musical sounds. Further investigation is required in this area. From the results of this study, then, the conclusion should be drawn that the visual representation of music is surprisingly consistent across cultures, and that Cartesian representation is the dominant mode.

However, we return to the notable proportion of responses which were iconic in nature and did not make use of a linear timeline. The majority of participants who did not follow the left-to-right horizontal path of representation were the non-literate BenaBena, who opted for an abstract pictorial method which did not depict time. Additionally, we have the responses of the two Japanese master musicians unfamiliar with WSN, whose interviews suggest that they are ideologically opposed to what they consider a threatening expansion of western culture (Athanasopoulos, 2013, pp. 167-169). This suggests that sociological factors may well influence—or overrule—participants' first responses to the task. The responses and interview with the BenaBena thus provide a unique insight into the opposing perspective, with the example of a non-literate community's first encounter with a novel form of musical communication. (This was a benign encounter whose potential consequences the researcher considered and attempted to mitigate (see *Ethical Considerations* in Athanasopoulos, 2013)).

Responses from the United Kingdom and Japan classified as iconic were idiosyncratic in nature and represent a minority in relation to the number of symbolic responses within the same groups. The BenaBena's pictorial approach was more predictable. We attribute the consistency of their responses to their approach to the task. First, after the initial introduction of the idea to the BenaBena, the participants would discuss the notion of sound representation among themselves trying to reach some sort of consensus before the task took place. Though they were not permitted to look at each other's responses, participants would openly discuss their answers and debate the appropriateness of their responses. This mode of collaborative engagement was as open to Groups A and B as it was to C, but it only took place—it only seemed to be required—for the BenaBena.

Subsequent interviews revealed that the participants were not attempting to indicate variations within the stimuli, but rather attempted to adhere strictly to the instruction given to them: "Produce responses so that if another member of the community saw them they would be able to link them to the sound events". Thus, they completed the task for its communicative function. Usually after discussion with their peers, they invented–collaboratively–an iconic method of representation appropriate for the task. Participants attempted to indicate that a flute was playing through iconic representation of the sound, without attempting to indicate specific variations between the sound events (see Figure 4). When questioned if all the sound events were the same (regardless of contour variation), the participant responded, "Yes, it is still a flute playing" (Athanasopoulos, 2013, p. 239). When asked about perceived differences regarding the values of specific sound events that followed a pattern (up, down, peak or valley) the participant replied by asking why these differences should matter, since it is always a flute playing, regardless of these variations (Athanasopoulos, 2013, p. 239). This particular style of response was, in fact, replicated throughout further sections of the task (attack rate and duration), whose results are not discussed in this particular article. For example, in order to depict variations in attack rate, participants deployed circles to stand for drums, while lines next to or on the circles indicated the action of hitting the drum, or the number of strokes (Athanasopoulos & Moran, 2012). Small variations in the length of lines next to the circles indicated perceived variations in loudness.

CONCLUSION

At the moment of its creation, the iconic representation of music for a communicative function must follow society's norms. Conceivably, this may in certain circumstances gradually give rise to symbolic systems. But in order for such musical communication to take place in the first instance, members of a community must reach agreement on the salient aspects of what is being perceived. The BenaBena participants' debates made the task possible by creating a relevant context for the functional (communicative) task in hand. They produced responses which attempted to signal common points of reference without recourse to an existing bank of symbols or through the structure of Cartesian organisation. The BenaBena's purely functional approach to the task highlights the specific nature of WSN, whose role in the established canon of Western art music–and relationship with the concept of the musical work–embodies both ideological and aesthetic

concerns (Goehr, 2007). Whilst WSN is ubiquitous, we are reminded that the underlying principles with which it is associated are not universal.

On the topic of universal expression and comprehension of musical meaning, there is little agreement amongst scholars in ethnomusicology, music psychology or in semiotics. In this study, three culturally diverse groups of participants provided empirical data with which to contemplate one small aspect of the topic, using a free-drawing paradigm. The results point to literacy as the most significant factor in predicting the type of representation used to depict musical sounds. Literate participants unfamiliar with music notation still tended to deploy a timeline to mark the progression of sounds in time. The responses of the non-literate BenaBena participants, who were also non-music-notational, were most striking in their distinctive approach to musical shape association. Despite the unfamiliarity of the task, the participants quickly defined appropriate iconic references to meet the communication criteria of the task. These culturally appropriate references focused on musical parameters that seemed either to matter more to participants, or were deemed least ambiguous. As a result, variations between sound events within a specific category of musical stimuli (such as pitch variations) were put aside, in order to maintain the clarity of the symbolic (icon) reference.

NOTES

[1] A note on "culture". In a recent special issue of this journal (Meaning and Entrainment in Language and Music, Volume 12), Widdess draws on cultural anthropologist Maurice Bloch's (1998) definition of culture: "That which needs to be known in order to operate reasonably effectively in a specific human environment" (Bloch, 1998, p. 4). This definition is consonant with our own usage, and we appreciate Widdess' elaboration that much of what gives a musical *culture* its identity depends on knowledge that is acquired and deployed in non-linguistic ways (Widdess, 2012, p. 88). Meanwhile, Cross reminds the reader that academic discourse on musical *culture* typically refers only to constructs which are continually "refashioned" in relation to one another (Cross, 2012, p. 95); culture, therefore, is not a reliable concept for the faithful description of permanently or objectively recognisable "domains" (Cross, 2012, p. 95). We acknowledge this argument too, recognising the reductive and desensitising effect of over-use of the term. This project, however, uses the concept of "culture" to make comparisons that can highlight particular elements of difference and diversity in individuals' responses to specific elements of musical sound.

[2] The study and fieldwork were carried out as part of the first author's doctoral research, involving more than 120 participants from 5 groups from three distinct cultural backgrounds (U.K., Japan, Papua New Guinea) in Edinburgh (U.K.), Tokyo & Kyoto (Japan), and Port Moresby and the BenaBena villages in Papua New Guinea. The preliminary results of this study are published as conference proceedings (Athanasopoulos, 2013; Athanasopoulos & Moran, 2012; Athanasopoulos, Moran, & Frith, 2011). The research and fieldwork were funded by the Onassis Foundation, Greece, the Great Britain Sasakawa Foundation, and the University of Edinburgh.

[3] This is relative. Missionaries have now reached even the most remote tribes in Papua New Guinea. The world that Feld (1982) described no longer exists; within two generations Papua New Guinean societies have been radically transformed.

[4] The doctoral study from which these findings are drawn examined participant responses to three sets of 20-trial auditory stimuli varying on pitch, duration and attack rate for the free-drawing investigation. Participants also responded to a further 12 auditory stimuli (varying on pitch and attack rate) for an additional forced-choice study.

REFERENCES

Adachi, M. (1997). Japanese Children's Use of Linguistic Symbols in Depicting Rhythm Patterns. In: *Proceedings of the 4th International Conference on Music Perception and Cognition*. Montreal, Canada: McGill University, pp. 413-418.

Athanasopoulos, G. (2013). *Scoring Sounds: the Visual Representation of Music in Cross-Cultural Perspective.* PhD dissertation, Reid School of Music, University of Edinburgh.

Athanasopoulos, G., & Moran, N. (2012). Pictorial notations of pitch, duration and tempo: A musical approach to the cultural relativity of shape. In: *Proceedings of the 12th International Conference on Music Perception and Cognition.* Thessaloniki, Greece, p. 69.

Athanasopoulos, G., Moran, N., & Frith, S. (2011). Literacy makes a difference: A cross-cultural study on the graphic representation of music by communities in the United Kingdom, Japan and Papua New Guinea. Paper presented at *the 2011 Biennial Meeting of the Society for Music Perception and Cognition (SMPC).* Rochester, NY, USA.

Avid Technology. (2009). *Sibelius 6 Software*. http://www.sibelius.com/home/index_flash.html

Bamberger, J. (2005). How the conventions of music notation shape musical perception and performance. In: D. Miell, R. MacDonald, & D. J. Hargreaves (Eds.), *Musical communication*. New York: Oxford University Press, pp. 143-170.

Barrett, M. (2005). Representation, cognition and communication: Invented notation in children's musical communication. In: D. Miell, R. MacDonald, & D.J. Hargreaves (Eds.), *Musical communication*. New York: Oxford University Press, pp. 117-142.

Blacking J. (1973). *How Musical is Man?* University of Washington Press (February 1, 1990) reprint from 1973.

Bloch, M. (1998). *How We Think They Think: Anthropological Studies in Cognition, Memory and Literacy*. Boulder: Westview Press.

Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, Vol. 43, No. 1, pp. 1-22.

Boroditsky, L. (2011). How language shapes thought. *Scientific American,* Vol. 304, pp. 62-65.

Casasanto, D., Philips, W., & Boroditsky, L. (2003). Do we think about music in terms of space? Metaphoric representation of musical pitch. In: R. Alterman & D. Kirsch (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society.* Boston, MA: Cognitive Science Society, p. 1323.

Cross, I. (2012). Musics, cultures and meanings: Music as communication. *Empirical Musicology Review*, Vol. 7, Nos. 1-2, pp. 95-97.

Carello, C., Anderson, K.L., & Kunkler-Peck, A.J. (1998). Perception of object length by sound. *Psychological Science*, Vol. 9, pp. 211-214.

Deregowski, J.B. (1972). Pictorial perception and culture. *Scientific American,* Vol. 227, No. 5, pp. 82-88.

DigiDesign (2009). *Digidesign Pro Tools 8 Software*. Avid Technology. http://www.avid.com/US/products/family/Pro-Tools

Eitan, Z., & Granot, R.Y. (2006). How music moves: Musical parameters and listeners' images of motion. *Music Perception*, Vol. 23, No. 3, pp. 221-248.

Eitan, Z., Schupak, A., & Marks, L.E. (2010). Louder is Higher: Cross-Modal Interaction of Loudness Change and Vertical Motion in Speeded Classification. In: K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMP10)*, Causal Productions, Adelaide (2008), (10 pp).

http://www2.tau.ac.il/InternetFiles/Segel/Art/UserFiles/file/Proceedings%2010.pdf

Farago, A. (2007). Interview: Jason Thompson. *The Comics Journal*, September 30, 2007.

Feld, S. (1982). *Sound and Sentiment: Birds, Weeping, Poetics and Song in Kaluli Expression*. Philadelphia, PA: University of Pennsylvania Press.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, Vol. 19, No. 7, pp. 573-576.

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit non-linguistic task. *Cognitive Science*, Vol. 34, No. 8, pp. 1430-1451.

Gentner, D., Imai, M., & Boroditsky, L. (2002). As time goes by: Evidence for two systems in processing space-time metaphors. *Language and Cognitive Processes*, Vol. 27, No. 5, pp. 537-565.

Goehr, L. (2007). *The Imaginary Museum of Musical Works*, Oxford: Oxford University Press.

Küssner, M.B. (2013). Music and shape. *Literary and Linguistic Computing.* Advance Access published January 15, 2013: 10.1093/llc/fqs071.

Küssner, M.B., & Leech-Wilkinson, D. (in press). Investigating the Influence of Musical Training on Cross-Modal Correspondences and Sensorimotor Skills in a Real-Time Drawing Paradigm. *Psychology of Music*.

Küssner, M.B., Prior, H.M., Gold, N., & Leech-Wilkinson, D. (2012). Getting the shapes 'right' at the expense of creativity? How musicians' and non-musicians' visualizations of sound differ. *Proceedings of the 12th international conference on Music Perception and Cognition*, Thessaloniki, Greece, p. 121.

Layton, R. (1981). *The Anthropology of Art (*2nd Edition). Cambridge: Cambridge University Press.

Mitchell, M. (2004). The Visual Representation of Time in Timelines, Graphs and Charts. In: *Australian & New Zealand Communication Association Conference*. http://epublications.bond.edu.au/hss_pubs/107

Nettl, B. (1985). *Western Impact on World Music: Change, Adaptation, and Survival.* New York: Schirmer Books.

Nisbett, R. (2003). *The Geography of Thought: How Asians and Westerners Think Differently - And Why*. New York: Free Press.

Prior, H.M. (2010). Links between music and shape: Style-specific, language-specific, or universal? *1st International Colloquium on Universals in Music: Data, issues, perspectives*, Université de Provence, Aix, France.

Reybrouck, M., Verschaffel, L., & Lauwerier, S. (2009). Children's graphical notations as representational tools for musical sense-making in a music-listening task. *British Journal of Music Education*, Vol. 26, No. 2, pp. 189-211.

Roberson, D., Davidoff, J., & Shapiro, L. (2002). Squaring the circle: The cultural relativity of good shape. *Journal of Cognition and Culture*, Vol. 2, No. 1, pp. 29-53.

Rosch, E.H. (1973). Natural categories. *Cognitive Psychology*, Vol. 4, No. 3, pp. 328-350.

Sadek, A.A.M. (1987). Visualization of musical concepts. *Council for Research in Music Education*, Bulletin No. 91, pp. 149-154.

Schmandt-Besserat, D. (1992). *How Writing Came About*. Austin: University of Texas Press.

Tagg, P. (1993). Universal music and the case of death. *Critical Quarterly*, Vol. 35, No. 2, pp. 54-98.

Tan, S., & Kelly, M. (2004). Graphic representations of short musical compositions. *Psychology of Music*, Vol. 32, No. 2, pp. 191-212.

Tarasti, E. (2002). *Signs of Music: A Guide to Musical Semiotics*. New York: Mouton de Gruyter.

Verschaffel, L., Reybrouck, M., Janssens, M., & Van Dooren, W. (2010). Using graphical notations to assess children's experiencing of simple and complex musical fragments. *Psychology of Music*, Vol. 38, No. 3, pp. 259-284.

Walker, A.R. (1987a). Some differences between pitch perception by children of different cultural and musical backgrounds. *Council for Research in Music Education,* Bulletin No. 91, pp. 166-170.

Walker, A.R. (1987b). The effects of culture, environment, age and musical training on choices of visual metaphors for sound. *Perception and Psychophysics*, Vol. 42, No. 5, pp. 491-502.

Widdess, R. (2012). Music, meaning and culture. *Empirical Musicology Review*, Vol. 7, Nos. 1-2, pp. 88-94.

Zwaan, E.W.J. (1965). *Links en rechts in waarneming en beleving* (Left and right in visual perception as a function of the direction of writing). Doctoral Thesis, Rijksuniversiteit Utrecht, The Netherlands. Cited in: Winn, W. (1994), Contributions of perceptual and cognitive processes to the comprehension of graphics. In: Schnotz, W., & Kulliavy, R.W. (Eds.), *Comprehension of Graphics*. Amsterdam: North-Holland, pp. 3-27.