ELSEVIER

Cognitive Science 28 (2004) 531–558

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Can musical transformations be implicitly learned?

Zoltán Dienes*, Christopher Longuet-Higgins

Department of Psychology, Sussex University, Brighton, East Sussex BN1 9QG, England, UK

Received 11 April 2003; received in revised form 4 March 2004; accepted 8 March 2004

Available online 18 May 2004

Abstract

The dominant theory of what people can learn implicitly is that they learn chunks of adjacent elements in sequences. A type of musical grammar that goes beyond specifying allowable chunks is provided by serialist or 12-tone music. The rules constitute operations over variables and could not be appreciated as such by a system that can only chunk elements together. A series of studies investigated the extent to which people could implicitly (or explicitly) learn the structures of serialist music. We found that people who had no background in atonal music did not learn the structures, but highly selected participants with an interest in atonal music could implicitly learn to detect melodies instantiating the structures. The results have implications for both theorists of implicit learning and composers who may wish to know which structures they put into a piece of music can be appreciated.

Keywords: Implicit learning; Music; Serialism; Chunks; Unconscious knowledge

1. Introduction

Back in 1967 Reber noted that people could incidentally learn to perceive or act appropriately in a structured domain without being able to say what it is that had been learned. This implicit learning is shown by, for example, people's ability to learn languages and also the contingencies of social life (e.g., Lewicki, 1986; Reber, 1989). Implicit learning also seems to be revealed by people's ability to learn musical styles and structures: We may be able to say that a piece of music comes from a certain style or composer without being able to justify our claim. Reber (1989) argued further that the implicit learning of the structure of an art form is the basis

* Corresponding author.
E-mail address: dienes@sussex.ac.uk (Z. Dienes).

0364-0213/$ – see front matter © 2004 Cognitive Science Society, Inc. All rights reserved.
doi:10.1016/j.cogsci.2004.03.003

for the aesthetic appreciation of that art form (a view anticipated by Longuet-Higgins, 1976), implying that implicit learning plays a fundamental role in learning to appreciate music.

Most research into implicit learning has concentrated on the learning of artificial grammars (using finite state grammars with almost entirely local dependencies) in the context of classification tasks or reaction time tasks (see Berry, 1997; Berry & Dienes, 1993; French & Cleeremans, 2002; Stadler & Frensch, 1998, for reviews). Based on this research, the dominant theory of what people can learn implicitly is that people primarily learn chunks of adjacent elements in sequences (i.e., *n*-gram structure), whether this is the explicit claim of the theory (e.g., Boucher & Dienes, 2003; Dienes & Fahey, 1995, 1998; Johnstone & Shanks, 1999; Perruchet & Pacteau, 1990; Redington & Chater, 1996; Servan-Schreiber & Anderson, 1990) or happens to be the natural consequence of the proposed learning mechanism (e.g., Cleeremans, 1993; Dienes, Altmann, & Gao, 1999). Reber (e.g., 1993) has always argued people can implicitly learn abstract rules without specified constraints, but to date the evidence available can be largely explained by assuming people implicitly learn chunks of adjacent elements in sequences.

The perhaps surprising consequence of the claim that we learn only local structure is that the theory cannot apply to the induction of natural grammars (as shown by Chomsky, 1957), the original inspiration for the investigation of implicit learning, nor the induction of musical grammars (Lerdahl & Jackendoff, 1983; Longuet-Higgins, 1978), which also involve non-local dependencies. Music embodying these types of non-local dependencies may thus be an ideal domain for finding evidence that implicit learning can go beyond the learning of chunks of adjacent elements. So far implicit learning researchers have scarcely addressed the domain of music (for studies using musical notes and an artificial finite state grammar, see Altmann, Dienes, & Goode, 1995; Bigand, Perruchet, & Boyer, 1998; Saffran, Johnson, Aslin, & Newport, 1999).

A type of well-specified musical grammar that does go beyond just specifying allowable *n*-grams is provided by serialist or 12-tone music (see e.g., Wuorinen, 1979). Serialism is a method of composition introduced by Arnold Schoenberg at the beginning of the century, his first fully serialist piece having been written in 1920 (the last of his Five Pianoforte Pieces, Op 23; see Schoenberg, 1941). Schoenberg wished to form a new system of music in which there was no tonal centre but a "democracy" of notes. A serialist composition is constrained by the following rules:

1. The serialist rules specify the ordering not of pitches but of pitch classes—sets of pitches separated by octaves (e.g., C is a pitch class, but middle C is a particular pitch). It is simply up to the composer's aesthetic intuition as to which pitch he chooses to instantiate a particular pitch class at a particular point in the composition.
2. A tone row consists of an ordered arrangement of the 12 pitch classes (notated 0..11). An example of each of the 12 pitch classes is provided by all the black and white keys on a piano from middle C to the B above: C, Db ("D flat"), D, Eb, E, F, F# ("F sharp"), G, Ab, A, Bb, and B. Let 0 indicate C. Then each number represents the number of semitones above C. That is, Db = 1, D = 2, ... B = 11. The tone row can be represented by some permutation of the numbers 0... 11. Each pitch class must occur once and only once (hence the "democracy" of notes). The composer constructs an initial tone row (the

prime tone row) to act as a theme for the piece; the exact ordering of pitch classes within the prime is up to the composer's musical intuition.

3. Once a prime tone row has been constructed for a composition, four transforms are possible: transpose, retrograde, inverse, and inverse retrograde. To form a transpose of a tone row, add a constant (modulo 12) to each note in turn. (Modulo 12 means clock face arithmetic; for example, $8 + 5 = 1$ in modulo 12 arithmetic: 5 h after 8 o'clock is 1 o'clock.) The constant can be any number 0. . . 11. For example (the prime is in fact the prime from the last of Schoenberg's Five Piano Pieces),

P:	1 9 11 7 8 6 10 2 4 3 0 5	(Db A B G Ab F# Bb D E Eb C F)
T(5)	6 2 4 0 1 11 3 7 9 8 5 10	(F# D E C Db B Eb G A Ab F Bb)

Any of these 12 transpositions can be played in reverse order to make the 12 retrogrades. The inverse to a prime can be formed by subtracting each prime note in turn (modulo 12) from a constant; the constant can be any number 0..11, giving 12 possible inverses. For example,

P:	1 9 11 7 8 6 10 2 4 3 0 5	(Db A B G Ab F# Bb D E Eb C F)
I:	6 10 8 0 11 1 9 5 3 4 7 2	(F# Bb Ab C B Db A F Eb E G D)

In this example, the constant is 7. Notice in the inverse, each successive interval is of the same magnitude but opposite sign compared to the prime. For example, the first interval in the prime is $(9 - 1) = +8$, or equivalently -4 in modulo 12. The first interval in the inverse is $(10 - 6) = +4$.

Any of these 12 inverses can be played in reverse order, to make the 12 retrograde inverses.

4. A serialist composition consists of successive or overlapping statements of these forms (with, in practice, the above rules violated to a greater or lesser degree; for example, immediate repetitions of a note are regarded as acceptable).

Serialism had an enormous impact on twentieth century classical music, perhaps the greatest impact of any well-specified system of composition (see DeLone et al., 1975) and for this reason alone seems worthy of the attention of psychologists. In addition, Schoenberg (1941) was very concerned that serialist music should be comprehensible and coherent.[1] Its comprehensibility as a serialist piece entails firstly that a tone row should be recognized as the same auditory object under the four transforms (transpose, retrograde, inverse, retrograde inverse). This would provide coherence or unity to a serialist composition: Listening to such a composition would involve beholding this auditory object as it was successively reflected, translated, and restated through the composition. The comprehensibility of a serialist piece entails also that a tone row is perceived as the same regardless of the octave in which each of the pitch classes is played (the assumption of octave equivalence; see Deutsch, 1982, 1999, for discussion of this assumption).

On the other hand, Lerdahl (1988) suggested that the "compositional grammar" employed by serialists in constructing a serialist piece does not correspond to any "listening grammar" that can be induced by a listener: The serialist grammar is not learnable by the listener. Smith and Witt (1989) and Thomson (1991) also raised the question of whether the syntax of serialist music is accessible to listeners. Imberty (1993) suggested the rules provide only a framework for composition and are not meant for the listener at all. So to what extent is the serialist grammar learnable by the listener? Francès (1958/1988; experiment VI) asked musically experienced participants to determine which of two tone rows had been presented each trial, as the rows were transformed across trials. He concluded that his participants had limited ability to perceive the tone row across different transformations. However, as one row occurred six times more frequently than the other, and no account was taken of response bias, it is in fact difficult to draw any conclusions from this interesting study. Without using inferential statistics, De Lannoy (1972) found some ability to perceive tone row transposes but not inverses. Dowling (1972) found that people could detect pitch transforms of atonal five-note sequences when the four transforms were fully explicitly explained to them and one transform was attempted at a time (*pitch* transform meaning that the transforms were not produced using modulo 12 arithmetic—as above, which specified *pitch class* transforms—but using normal arithmetic[2]). More impressively, Krumhansl, Sandell, and Sergeant (1987) found that with repeated testing using two primes (of 12 notes each) musically experienced participants fully aware of the nature of the task and the transformations used could accurately indicate which prime had been presented (transformed) on each trial. These studies provide an initial step in demonstrating that people can learn to deal with the serialist transformations at least to some degree when they are explicitly trying to discern the transformations. Balch (1981), using pitch rather than pitch class transforms, did not explain any of transforms to subjects. He found that people rated inverses and retrogrades of five and seven note sequences as better continuations of those sequences than unrelated sequences. The responses may or may not have been based to some degree on implicit knowledge, though Balch noted some subjects spontaneously described relevant explicit strategies. Bigand, D'Adamo, and Poulin (2003) argued that through repeated exposure to melodies instantiating a tone row, people simply learn the relative frequencies of the pitch class intervals in the tone row (that is, the order of the intervals may be irrelevant), and this allows identification of the tone row across transformations, providing some sense of unity (though, by implication, the same unity for any permutation of the intervals).

Whether people can *implicitly* learn to perceive the four transformations and the unity of the underlying tone row above and beyond contingent *n*-gram statistics remains an open question. If the answer is YES, then people can learn more than current theories of implicit learning give them credit for, and it would be a challenge to existing models of implicit learning (e.g., Dienes et al., 1999). If the answer is NO, existing theories are corroborated. Answering this question also provides important information for composers. If the serialist structure of a piece of music is not learnable, then the piece would be appreciated in spite of its serialist structure, not because of it, and a composer may consider whether he or she wishes to abide by the constraints. Whatever aspects are learnable explicitly, are potentially available to the listener for appreciation at an intellectual level; whatever aspects are available implicitly may form the basis for direct aesthetic appeal. Determining the learnability of the transforms provides useful information for composers, whether they are strict serialists or not; the transforms can occur in some guise in many musical styles (e.g., Bach's fugues, which used pitch inverses and retrogrades, for example).

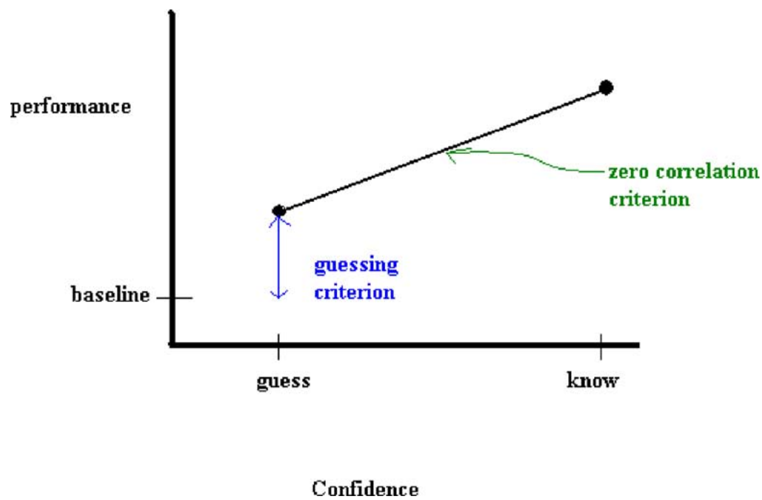

Fig. 1. Subjective measures of implicit knowledge.

In the following experiments, participants were asked to listen to a series of tone rows that follow one of the transforms, but participants were not told at that stage that there was any common regularity to the set of tone rows. Then they were shown some more tone rows, half of which followed the transform they were trained on and half of which did not. Participants were asked to classify them. The implicitness of the knowledge used for classification was assessed by confidence ratings given after each classification decision. Following Dienes and Berry (1997; also Dienes & Perner, 1999) we define unconscious or implicit knowledge as occurring when a person is in an occurrent mental state of knowing some content (e.g., that a tone row follows the rule) when the person does not know they are in that mental state (they do not know that they are occurrently representing and taking for true a particular content). Thus, if the person has implicit knowledge, they may classify above chance when they believe they are guessing (guessing criterion), and also be unable to distinguish between states of guessing and states of knowing (zero correlation criterion). Jointly, these criteria are called subjective measures of implicit knowledge because they are based on the person's awareness of the mental state they are in.[3]

These subjective measures of implicit knowledge are illustrated in Fig. 1, which shows a plot of classification accuracy against confidence. The classification performance given by the intercept when confidence is "guess" indicates how well people classify when they believe they do not know anything at all. If classification performance at this intercept is significantly above baseline then the guessing criterion of implicit knowledge is satisfied: People have knowledge they do not know that they have. If people cannot distinguish at all between when they know and when they are guessing over all trials, then the slope of the line will be zero. This is the zero correlation criterion of implicit knowledge: A zero confidence-accuracy slope indicates implicit knowledge. Notice that a significant guessing criterion indicates the presence of some implicit knowledge without ruling out the possibility of there being some explicit knowledge on other trials. Conversely, a significant slope indicates the presence of at least some explicit knowledge without ruling out the possibility of there being some implicit knowledge. The strongest evidence of knowledge being completely implicit is a significant guessing criterion and a non-significant zero correlation criterion (see Dienes, in press; Dienes & Perner, in press, for further discussion of these criteria).

2. Experiment 1

2.1. Method

2.1.1. Participants

The participants were students from Sussex University. Four groups of participants were used; each group was exposed to just one transform. There were 10 participants of unselected musical ability randomly assigned to each group.

2.1.2. Materials

In constructing the materials we took into account the fact that in actual serialist compositions, tone rows are usually deployed in smaller segments, the division of the row into two hexachords being common (Wittlich, 1975). (A hexachord is a sequence of six successive tones.) We incorporated this constraint by presenting people with tone rows (i.e., the 12 tones with each tone occurring once and only once), with the second hexachord being a transformation of the first. Constructing stimuli satisfying this description is a non-trivial task that we achieved in the following way.

Imagine the 12 numbers of the clock face and let these represent the 12 tones. The task is to divide the notes into two equal sets, one set being used for the first hexachord (call these the whites) and the second being used for the second hexachord (call these the blacks). Each white must be transposable into a black with the same constant of transposition. To satisfy this constraint we used 6 as the constant (a six semitone interval is known as the *diabolus in musica*, the diabolic interval, frequently used in serialist music). The constraint of "odd parity" is that each white reflects into a black through the centre of the clock face, and vice versa. For example, in Fig. 2, the white at 2 o'clock reflects through the centre into the black at 8 o'clock (and vice versa), so these two notes satisfy the constraint of odd parity. The one is a transpose of the other through six semi-tones. In the materials produced for experiment one, the odd-parity constraint was satisfied for each of six whites and six blacks of each tone row. Satisfaction of the odd-parity constraint allowed a transpose to be formed, defined in the following way. $\boldsymbol{h1}$ is a transpose of $\boldsymbol{h2}$ through six semi-tones, and vice versa, iff $\boldsymbol{h2} = \boldsymbol{h1} + \mathbf{6}$ (modulo 12), where $\boldsymbol{h1}$ is a 6-vector defining the first hexachord (i.e., an ordered list of the six notes in the first hexachord), $\boldsymbol{h2}$ is a 6-vector defining the second hexachord, and $\mathbf{6}$ is a 6-vector where all components are "6".

There must also be a constant of inversion that transforms each white into a unique black. Fig. 2 shows this geometrically. The diameter through half past one in the figure is an axis of anti-symmetry in the sense that the image of each black dot in this line is a white dot, and vice versa. The reader may care to verify that any dot and its image in the line add up to three, modulo 12. Thus, if each white reflects onto a black through a single axis of anti-symmetry,

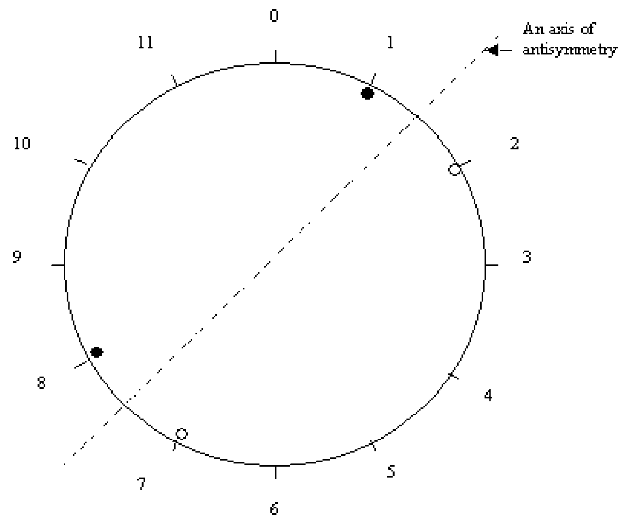

Fig. 2. The odd parity constraint and axis of anti-symmetry.

the whites will be invertible into the blacks (and vice versa). The inversion can be defined in the following way. $\boldsymbol{h1}$ is an inverse of $\boldsymbol{h2}$, and vice versa, iff $\boldsymbol{h1} + \boldsymbol{h2} = \boldsymbol{k}$ (modulo 12), where $\boldsymbol{k}$ is a 6-vector in which all components are k, $0 \leq k \leq 11$.

Materials were constructed according to the following algorithm, the application of which is illustrated in Figs. 2 and 3:

(i) Select randomly one of the six axes of antisymmetry. In Fig. 2 this is the dashed line.
(ii) Select randomly an available number, call it a white. In Fig. 2, this could be number 2.

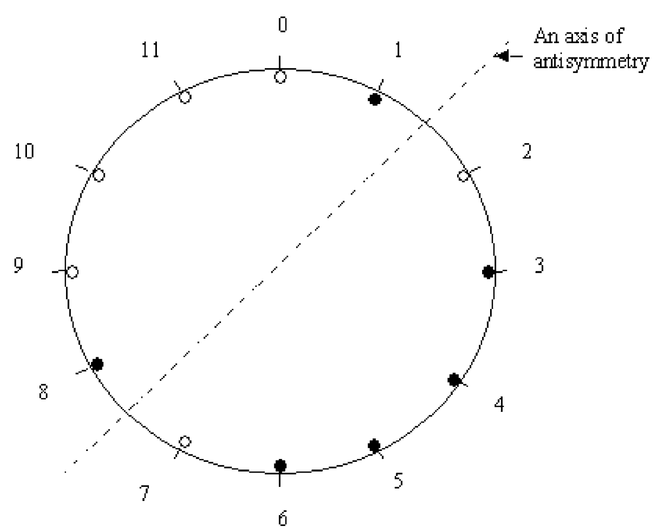

Fig. 3. An example of six white and six black notes having been chosen.

(iii) The constraints of odd parity and invertibility determine the white and black status of three other numbers. The 2 being a white means, by odd parity, the 8 must be a black. The 2 being white also means 1 must be black—this is reflection through the axis of antisymmetry. Similarly, if 8 is black, then 7 must be white by reflection through the axis of antisymmetry.

(iv) Repeat steps (ii) and (iii) until all numbers have their white or black status determined. For example, next the 0 could be randomly chosen to be white, which will determine the black/white status of three other numbers; there will be four numbers remaining, one of which is randomly chosen to be a white, fixing the other three. Fig. 3 illustrates a complete set of six whites and six blacks produced in this way.

All the whites were then randomly ordered to form the first hexachord; the second hexachord was determined by the transform used. For example, if the whites generated by steps (i)–(iv) above were the unordered set $\{0, 2, 7, 9, 10, 11\}$ then the blacks would be the complementary set $\{1, 3, 4, 5, 6, 8\}$. The whites could be randomly ordered to form the hexachord 11, 2, 0, 10, 9, 7.

We will now consider how the blacks would be ordered to form each of the transforms. If the transform were transpose, the second hexachord would be $(11 + 6)$, $(2 + 6)$, $(0 + 6)$, $(10 + 6)$, $(9 + 6)$, $(7 + 6)$, which is (modulo 12) 5, 8, 6, 4, 3, 1. Notice the second hexachord is composed entirely of blacks (this is a necessary consequence of the procedure (i)–(iv)). The two hexachords can be put together to form the tone row: 11, 2, 0, 10, 9, 7, 5, 8, 6, 4, 3, 1 (B, D, C, Bb, A, G, F, Ab, F#, E, Eb, Db).

If a second hexachord is a transpose of the first, it is a simple matter to turn it into a retrograde by running it backwards. For example, since 5, 8, 6, 4, 3, 1 is a transpose for the first hexachord in the above example, 1, 3, 4, 6, 8, 5 is a retrograde of the first hexachord, and only consists of blacks. Appending this retrograde to the first hexachord creates the tone row: 11, 2, 0, 10, 9, 7, 1, 3, 4, 6, 8, 5 (B, D, C, Bb, A, G, Db, Eb, E, F#, Ab, F).

Now consider the inverse transform. Once again, take the case where the first hexachord is 11, 2, 0, 10, 9, 7. These numbers were generated from the example in Fig. 1, where the axis of antisymmetry ensures that for every white there is a black such that the two sum to the number 3 (modulo 12). So, going through the first hexachord in order, we have $11 + 4 = 3$, $2 + 1 = 3$, $0 + 3 = 3$, $10 + 5 = 3$, $9 + 6 = 3$, and $7 + 8 = 3$. Hence, the second hexachord is 4, 1, 3, 5, 6, 8. Notice all numbers in the second hexachord are blacks, and the second hexachord is the inverse of the first. The first and second hexachords put together form the tone row: 11, 2, 0, 10, 9, 7, 4, 1, 3, 5, 6, 8 (B, D, C, Bb, A, G, E, Db, Eb, F, F#, Ab). The inverse retrograde is created by running this second hexachord backwards and appending to the first: 11, 2, 0, 10, 9, 7, 8, 6, 5, 3, 1, 4 (B, D, C, Bb, A, G, Ab, F#, F, Eb, Db, E).

For each transformation, 50 tone rows were independently generated for the training phase using the above algorithm, and 50 for the test phase. A random 25 of the test items had the order of pitch classes in the second hexachords randomly scrambled. The first 35 items of the training phase and of the test phase had pitches within the interval middle C to the B above. The last 15 were "octified", which meant that for each note (independently) $+1$, 0, or -1 octaves were added with equal probability (this manipulation was motivated by Schoenberg's assumption of octave equivalence; Krumhansl et al., 1987; Pederson, 1975). Each pitch was

Table 1
Percent correct classification in experiment 1 (standard errors in parentheses)

Transform	Octave displacement	
	Non-octified	Octified
Transpose	47.4 (2.4)	41.3 (7.0)
Inversion	50.1 (2.0)	57.4 (3.5)
Retrograde	50.8 (2.2)	44.5 (3.1)
Retrograde inversion	50.9 (2.5)	48.7 (2.0)

produced by the piano output of a Roland Piano 5500 keyboard, lasting 0.5 s. In the training phase there was a 4-s delay between each tone row, and in the test phase there was a 6-s delay between each tone row. Stimuli were recorded onto DAT tape, so for each transform the order of stimuli was the same for each subject.

The exact stimuli are shown at http://cognitivesciencesociety.org/supplements/.

2.1.3. Procedure

In the learning phase, participants rated how pleasant they found each tone row, on a scale ranging from 1 (very unpleasant) to 10 (very pleasant). In the test phase, they were told that the stimuli they had just heard obeyed some set of rules and that half the stimuli they were about to hear would obey the same rules and half would not; they were to classify which were which by circling yes or no. They were also asked to give a confidence rating for each response on a scale from 50% (complete guess) to 100% (complete certainty).

2.2. Results

The participants' scores were convincingly at chance for each type of material and for the materials as a whole. For the non-octified stimuli, the mean overall classification performance was 50% ($SE = 1.2\%$), which was not significantly different from chance $t(39) = 0.17$, $p = .86$ (upper limit of 95% CI = 52%); for the octified stimuli, the overall mean classification performance was 48% ($SD = 2.2\%$), also not significantly different from chance, $t(39) = 0.87$, $p = .39$ (upper limit of 95% CI = 53%).

Table 1 displays the mean percent correct classification (standard errors in parentheses) for each condition separately. A two-way (transform (transpose vs. inversion vs. retrograde vs. inverse retrograde) by octave displacement (octified vs. non-octified)) mixed model ANOVA indicated no significant effects, all p's > .10. Subjects convincingly failed to learn the materials in experiment 1.

The confidence ratings were not analysed because there was no learning.

3. Case study with PF

We ran a single person, Philip Fine, on the stimuli from experiment 1. In contrast to the participants in experiment 1, he was (a) highly experienced in atonal music as a singer of

Table 2
PF's percent correct classification of the materials from experiment 1

Transform	Octave displacement	
	Non-octified	Octified
Transpose	91	47
Inversion	77	73
Retrograde	83	80
Retrograde inversion	74	60

contemporary classical music (he had also previously composed a serialist piece); and (b) aware for each training and test set that the target relation between first and second hexachords was one of the serialist transforms (i.e., conditions favoured explicit learning). He was trained and tested on each transform in turn. Given our personal difficulty in perceiving any structure in our own materials, we were amazed by his performance, shown in Table 2. He performed well on both octified and nonoctified materials. (This is more impressive than the Krumhansl et al., 1987, results above because he had to deal with a different prime on every trial.) In Table 2, for each transform, for non-octified materials, there are 35 trials, so any performance above 64% is significantly above chance by the binomial, one tailed (PF performed significantly above chance for each transform), and for the octified materials, there are 15 trials and 80% is above chance, and 73% marginally so ($p = .06$ 1-tailed). No transform was significantly easier than any other (all p's > .10 by χ^2). Consistent with Krumhansl et al. (1987), PF found the non-octified material easier than the octified material, $\chi^2(1) = 6.31$, $p < .025$.

PF could explicitly describe various relevant strategies he used. He said he hummed the first three notes of each hexachord and compared them across hexachords. After each tape he could say what the target transform was. There is no doubt he could explicitly perceive the different transforms. Moreover, there was also evidence for implicit knowledge. Taking the 50 trials where he said he was just guessing (which were mainly the octified trials), he classified 62% correctly ($p = .03$, binomial 1-tailed), thereby satisfying the guessing criterion of implicit knowledge.

4. Experiment 2a

It may be that musically unselected people have difficulty dealing with the materials of experiment 1 partly because of the random ordering of the pitch classes within each hexachord: Such an auditory object may be difficult to hold in memory. Melodies that are smoother in contour, not randomly going up or down, may be easier to remember and process. We ran a follow up experiment in which the pitch class content of each hexachord was prepared in the same way as experiment 1, but was ordered in a different way. Namely, in the first hexachord the tones were presented in order, clockwise or anticlockwise, around the clock face from a random starting position. For example, consider the set of pitches $\{0, 2, 7, 9, 10, 11\}$. If the random starting note was 9, and the random direction was clockwise, the first hexachord would be: 9, 10, 11, 0, 2, 7 (A, Bb, B, C, D, G). Thus, for a transpose, the second hexachord would be 3, 4, 5, 6, 8, 1 (Eb, E, F, F#, Ab, Db). This procedure creates hexachords with fewer changes in pitch direction than those used in experiment 1. Twenty-eight musically unselected subjects were assigned randomly to be trained and tested on one of the four transforms; and 40 control subjects just received a test phase and were asked to sort tunes that seemed more rule governed from than those that did not. However, once again there was convincingly no learning (mean classification performance was 49% for experimental subjects and 53% for control subjects). (For details of this "smooth" experiment, see http://cognitivesciencesociety.org/supplements/.)

Whatever constraints we may dream up to make the tone rows more learnable, it may be that serialist composers have already intuitively used appropriate constraints in the tone rows that they end up constructing for genuine compositions. Thus, experiments 2a and 2b looked at the learnability of real tone rows.

In addition, it may be that appreciation of material as difficult as serialist transforms depends on having extended exposure with the material, as in PF's case. Thus, in experiments 2a and 2b, in addition to normal undergraduates, aficionados of serialist music were also selected. If anyone can implicitly learn to perceive the transforms, it must be people who regularly listen to and enjoy serialist music. PF was explicitly trying to determine the transform used; could aficionados detect the transform under more incidental conditions? We wanted to select aficionados of serialist music without them knowing they were being selected because of their interest in serialism.

People with considerable experience of atonal music may have superior perceptual skills which might allow them to induce a transform based on exposure to examples instantiating the transform. Alternatively, experienced people may already have explicitly or implicitly learned to detect the transforms of serialist music and a short exposure to examples of a transform only functions to trigger the knowledge that a particular transform is relevant to the current situation, rather than to produce learning of the transform per se (cf. Hadley, 2001). Either way, the knowledge that the transform is relevant to the current situation could be implicit or explicit. If the training phase leads to implicit knowledge, as shown by the guessing or zero correlation criteria, we will conclude that implicit learning has occurred at least in a broad sense, i.e., at least in the sense that people have acquired implicit knowledge that a transform is relevant in this context. Relevant pre-existing knowledge may include implicit knowledge of the transform as a rule structure or even as a large look up table, in which the person has merely rote learned that a set of tone rows happen to instantiate a given transform. We will not empirically distinguish the possibility of triggering pre-existing knowledge from inducing the transform from scratch, nor the type of prior knowledge people might have. The issue of what exactly experienced people may have learnt from their prior experience will be considered further in the discussion.

We introduced three other changes. First, participants were asked to focus their attention on the relationship between the second hexachord and the first, but we tried to do this in a way that would not encourage explicit hypothesis testing. It may be that implicit learning, as much as explicit learning, is strongly influenced by what is attended to (Jiménez, 2003; Perruchet & Vinter, 2002; Whittlesea & Dorken, 1993). Second, in the test phase participants had to distinguish between different transforms (i.e., the transform they had been trained on, and one other). To appreciate fully the structure of a serialist piece, the listener has to distinguish the transforms from one another. Third, none of the stimuli were octified. PF found the non-octified material easier than the octified; Krumhansl et al. (1987) also found their subjects had more difficulty dealing with tone rows in which the pitches could occur in one of a number of octave ranges rather than being fixed. Hence, participants may do better if none of the material is octified.

In experiment 2a, participants were trained on transposes and then in the test phase had to select transposes from a set of transposes and inverse retrogrades. In experiment 2b, participants were trained on inverse retrogrades, and then performed the same test phase as participants in experiment 2a, but this time, they had to select the inverse retrogrades.

4.1. Method

4.1.1. Participants

Thirty-two participants from the University of Sussex participated in experiment 2a. There were two groups: one group inexperienced with serialist music and one group that was experienced. The 22 participants in the group not experienced with serialist music were all undergraduate students not selected on the basis of any musical criteria. In fact, no one in this condition had any knowledge of serialist music.

The 10 participants in the experienced condition were recruited via e-mail. The participants approached were either currently undertaking or had finished or were teaching a degree in Contemporary Music at Sussex University. The e-mail read "If you could create your own perfect radio station, how much time per day would you dedicate to playing each of these music styles?

(Pre-1900)	1. Baroque
	2. Classical
	3. Romantic
	4. Nationalist
(1900=>)	5. Minimalist
	6. Serialist
	7. Stochastic
	8. Textural
	9. Electronic"

The e-mail asked recipients to give a percentage for each style they would like to listen to, in all comprising 100%. They were told they did not have to include all the styles in the list, only the ones they would prefer to listen to and they were encouraged to include other styles not in the list. A person was chosen to be a serialist aficionado if they gave serialism a 10% rating or above; these people were re-approached and asked to participate in the second part of the experiment. Four of the participants were undergraduates, one was a PhD student, and five had finished their PhD (three of whom were currently lecturing in music).

All participants were unaware of the purpose of the experiment. There was one between-participants factor, musical experience (experience with serialism vs. no such experience). For this experiment, there was no manipulation of transform: All participants were trained on transposes.

4.1.2. Materials

We culled 35 prime tone rows from textbooks and reviews of serialism (see http://cogsci.psy.utexas.edu/supplements/ for exact sources and hexachords). The 70 hexachords were screened for odd parity (24 satisfied this constraint); invertibility (54 satisfied this constraint); non-equivalence to the other hexachords under the four transformations (42); and non-identity of transpose with inverse retrograde (66). There were 11 hexachords that satisfied all constraints. These hexachords came from Webern's Variations for Piano, Cantata no. 2, and Symphony Opus 21; Luigi Nono's Il Canto Sospeso; Berg's Lyric Suite, and Lulu; and Krenek's String Quartet, Opus 78.

From these 11 hexachords we needed to construct enough hexachords to make a training set and a test set. So two transpose versions (each a transpose through a random number of semitones 0..11) of each of the prime, retrograde, inverse, and inverse retrograde were constructed to make 88 hexachords (11 initial hexachords $\times$ 4 transforms $\times$ 2 transpose versions). Note so far that we are just dealing with hexachords, no tone rows have been constructed. Of the 88 hexachords, 40 hexachords were randomly assigned to the training set, 48 to the test set. Transposes (through $+6$ semitones) of all 40 hexachords in the training set were constructed to make the 40 tone rows of the training set, in each tone row the second hexachord being a transpose of the first. The training set consisted only of such tone rows instantiating transposes. The items in the test set consisted of an equal mixture of tone rows in which the second hexachord was a transpose of the first and in which the second hexachord was an inverse retrograde of the first hexachord. The tones rows within the training set and within the test set were presented in the same random order for all participants. All pitches were in the interval from middle C to the B above. None of the stimuli were octified. The timings were the same as experiment 1 (e.g., each tone lasted 0.5 s, with no gaps between tones within a hexachord), except the two hexachords were separated by a delay of one second to help draw attention to the hexachord structure of the tone row.

The exact stimuli are shown at http://cognitivesciencesociety.org/supplements/. An example of a transpose tone row is: 7, 8, 4, 6, 3, 5, 1, 2, 10, 0, 9, 11 (G, Ab, E, F#, Eb, F, Db, D, Bb, C, A, B) (the first—and second—hexachord, transformed in the ways allowed by serialism, was used by Webern in his "Variations for Piano").

4.1.3. Procedure

For training, participants were told, "students of contemporary music had been asked to compose six note themes, followed by a reply to the theme of the same length. They were not told to follow any particular rules in doing this, but just to follow their artistic intuitions. We will present you with 40 examples of the students' work. We have removed rhythm information and just present the tones as the reply was to the theme. It is important that you go by your first impressions, you have only 4 s to make your decision before the next theme/reply appears. We are interested in your immediate impression. Please rate each pair by circling very bad,

bad, good, or very good". The instructions were designed to keep participants unaware of the purpose of the experiment, simply making participants listen to the relation between the hexachords without explicitly looking for any rules underlying the musical structure.

At test, participants were told that for half the pairs the theme/replies have been swapped, and the other half were further genuine compositions from the students, and they had to classify which was which. In fact, "swapped" referred to the inverse retrogrades and "genuine" referred to the transposes. After each classification, participants gave a confidence rating on a scale from 50% (pure guess) to 100% (completely certain).

4.2. Results

The 22 inexperienced participants classified 50% ($SE = 1.5\%$) of the test items correctly, exactly on chance expectation. In contrast, the experienced participants classified 56.3% ($SE = 1.7\%$) correctly, which is significantly greater than the inexperienced participants, $t(30) = 2.51$, $p = .018$, and also significantly greater than chance, $t(9) = 3.61$, $p = .006$.

The experienced participants could perceive the difference between the transforms. But did they do this implicitly or explicitly? One way of measuring implicit knowledge is by looking at the plot of participants' classification performance against their confidence ratings. According to the guessing criterion (Dienes & Berry, 1997) people have implicit knowledge if they perform significantly above chance even when they believe they are literally guessing. According to the zero-correlation criterion, the knowledge is implicit if accuracy and confidence are unrelated. The guessing criterion can be measured by the intercept of the plot (at a confidence of 50%), and the zero-correlation criterion by the slope. Fig. 4 shows the plot of accuracy against confidence for just the experienced participants.

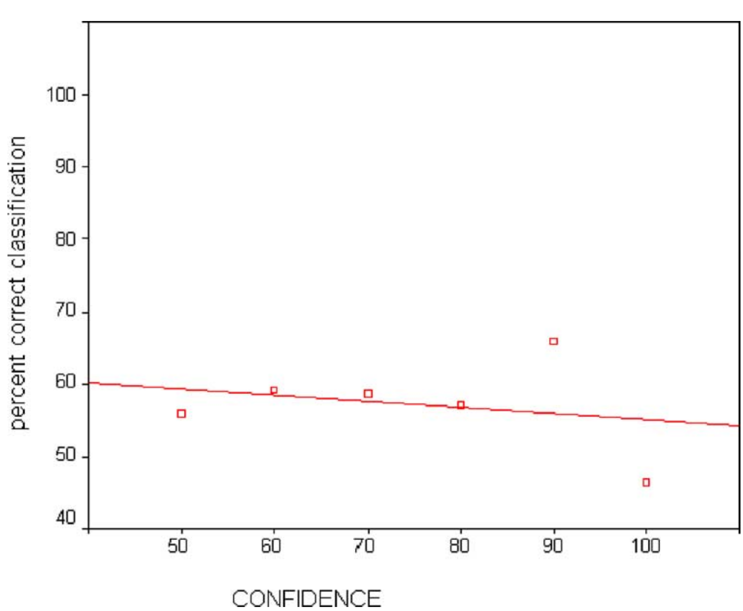

Fig. 4. Plot of classification percent correct against confidence for the experienced participants of experiment 2a.

The regression of classification performance against accuracy was calculated separately for each experienced participant. The average intercept (60%) was significantly above chance, $t(9) = 1.81$, $p = .05$, 1-tailed, and the slope (−0.2%) was not significantly different from chance, $t(9) < 1$.[4,5] That is, participants appeared to perform the task implicitly, without knowing that they knew or when they knew. Consistently, in free report at the end of the experiment, no participant mentioned transposition as a basis of their classification. Participants generally gave vague reports, for example, that their responses were based on intuition.

5. Experiment 2b

Experiment 2a showed that atonally experienced participants after training could classify the test set better than atonally inexperienced participants after training. Experienced participants could perceive the difference between the transposes and the inverse retrogrades (without knowing what the distinction was that they were responding to). However, the design does not allow the conclusion that the training phase was important in guiding experienced participants' test choices; maybe atonally experienced people just liked the transposes more than the inverse retrogrades anyway, regardless of the training phase. Experiment 2b completed the design by training experienced participants on inverse retrogrades and then testing them on the same test set as experiment 2a. If these participants no longer prefer the transposes then participants are learning specifically from the training phase how to choose items in the test phase.

5.1. Method

5.1.1. Participants

Fourteen participants were recruited from the composition department of the Birmingham Conservatoire by use of the same e-mail questionnaire used in experiment 2a. Those who said they would like to listen to serialism 10% or more of the time were selected for the experiment. Nine were third year undergraduates and five were postgraduate and faculty. All participants had received a minimum of one semester training in serialist theory.

A different music school was used for experiment 2b as compared to experiment 2a because the participants in 2a had been debriefed and recruiting serialist aficionados is not an easy task.

5.1.2. Materials

The training set used in experiment 2a was adapted by taking the second hexachord of each tone row and rearranging the notes so that they formed an inverse retrograde rather than a transpose of the first hexachord. The same test set was used as experiment 2a.

5.1.3. Procedure

The same procedure and exact instructions were used as in experiment 2a.

5.2. Results

The number of transposes endorsed in the test phase was 48.1% ($SE = 1.8\%$). This is significantly less than the number endorsed in experiment 2a by the experienced group (56.3%),

$t(22) = 4.00$, $p < .001$. That is, the results provide evidence that it was the training phase that biased participants to endorse test items of the same transform as appeared in the training phase.

One can further ask if one of the transforms was easier than the other. The endorsement of transposes in experiment 2a (56.3%) was significantly higher than the endorsement of inverse retrogrades (51.9%) in experiment 2b, $t(22) = 2.15$, $p < .05$.

For the participants in experiment 2b, overall endorsements of inverse retrogrades (51.9%) was not significantly above a baseline of 50%, $t(13) = 1.10$, $p = .29$. The participants were split into two further groups of different levels of experience: undergraduates ($n = 9$) and postgraduates and faculty ($n = 5$). The undergraduates endorsed 49.3% ($SE = 1.9\%$) of the inverse retrogrades; the postgraduates and faculty endorsed 57% ($SE = 2.1\%$). The difference between these two groups is significant, $t(12) = 2.54$, $p = .026$. The postgraduates and faculty scored significantly above 50%, $t(4) = 3.31$, $p < .05$.

For the postgraduates and faculty, the percentage endorsements of inverse retrogrades when they believed they were literally guessing (i.e., gave a confidence rating of 50%) was 64.3% ($SE = 5.5\%$), significantly above 50%, $t(4) = 2.58$, $p = .03$, 1-tailed. The plot of accuracy against confidence for these five participants is shown in Fig. 5.[4,6]

The regression of classification performance against accuracy was calculated separately for each experienced participant. The average intercept (59%) was marginally significantly above chance, $t(4) = 1.97$, $p = .06$, 1-tailed, and the slope (−0.2%) was not significantly different from chance, $t(4) < 1$.

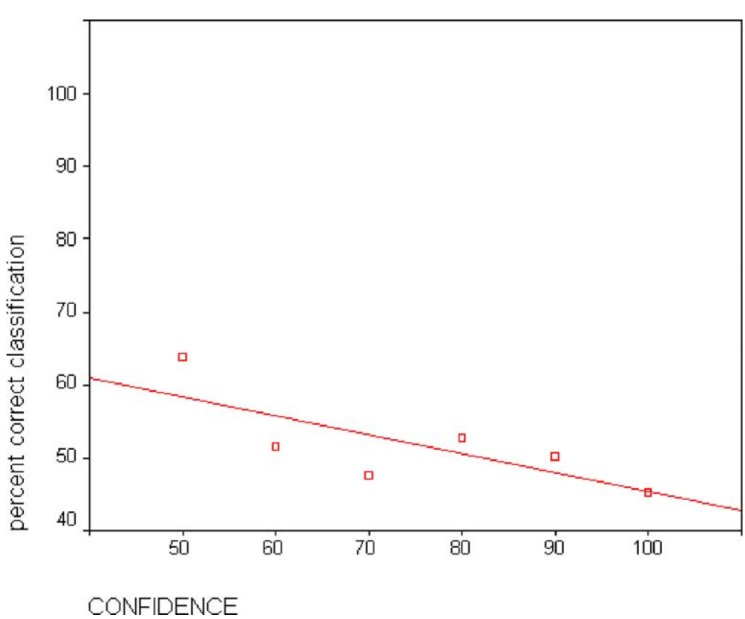

Fig. 5. Plot of classification percent correct against confidence for the postgraduates and faculty in experiment 2b.

6. What was implicitly learnt?

Experiments 2a and 2b and the case study with PF provide evidence of implicit learning in the training phase allowing discrimination of test items. But had participants really learnt to perceive implicitly the transforms per se? Based on existing implicit learning literature, the first hypothesis to be considered must be that participants had implicitly learnt *n*-gram structures. Saffran et al. (1999) found that people rapidly learned the local statistical structure of tone sequences, and Krumhansl (2000) found that amount of exposure to different styles of music predicted knowledge of the style's *n*-gram statistics. At least a majority of learning in the published implicit learning literature can be accounted for by presuming subjects learn only *n*-grams. Boucher and Dienes (2003) review the evidence for this claim in the artificial grammar learning literature. Similarly, Dienes and Fahey (1995, 1998) showed that learning in the dynamic control task paradigm could be modelled by the learning of associations between successive events, and Cleeremans (1993) modelled learning in the sequential reaction time paradigm as progressive sensitivity to the immediately preceding context.

If subjects discriminated test tone rows due only to the learning of *n*-grams, then the results of this paper present no challenge to existing models of implicit learning. However, if they go beyond the learning of *n*-grams, they present an interesting theoretical challenge. The possible role of *n*-gram learning will be explored for the materials of experiments 1, 2a and 2b in turn.

6.1. PF and the materials of experiment 1

For the materials of experiment 1 we calculated for each test tone row the mean number of times each element in it matched an element in the training tone rows, the mean number of times each bigram in it matched a bigram in the training tone rows, and the same for trigrams and tetragrams. For the elements considered as pitch classes or pitch class intervals, the means (and standard deviations over tone rows) for the grammatical and non-grammatical tone rows are shown in Table 3 (just transposes shown). Note that by virtue of being serialist tone rows, the grammatical and non-grammatical items have identical first order frequencies of each pitch class.

t-tests were performed on each corresponding difference between grammatical and non-grammatical test items; in two cases it was significant, indicating an imbalance in (pitch

Table 3
Mean *n*-gram match between test and training items for the transpose materials of experiment 1

		Unigrams	Bigrams	Trigrams	Tetragrams
Pitch classes	Grammatical		4.7 (0.9)	0.6 (0.3)	0.6 (0.4)
	Non-grammatical		4.5 (0.7)	0.5 (0.2)	0.6 (0.3)
Pitch class intervals	Grammatical	55.3* (4.9)	6.5* (1.7)	1.0 (0.9)	0.2 (0.4)
	Non-grammatical	42.8 (3.2)	5.4 (1.1)	0.7 (0.3)	0.1 (0.1)

Standard deviations in parentheses.

* The grammatical and non-grammatical items differ at the .05 level, $t(48) = 10.59$ for the unigrams and $t(48) = 2.47$ for the bigrams.

Table 4
Mean *n*-gram match between test and training items for the materials of experiment 2a

		Unigrams	Bigrams	Trigrams	Tetragrams
Pitch classes	Grammatical		3.7 (0.8)	0.4 (0.3)	0.5 (0.3)
	Non-grammatical		3.6 (0.7)	0.4 (0.2)	0.4 (0.3)
Pitch class intervals	Grammatical	43.9 (4.7)	5.7 (1.8)	1.9 (1.2)	1.0 (1.0)
	Non-grammatical	44.3 (4.8)	5.6 (1.8)	2.0 (0.9)	1.0 (0.5)

Standard deviations in parentheses.

class interval) unigrams and bigrams in experiment 1. The other transforms in experiment 1 also showed imbalances in unigrams, and also marginal differences in bigrams (see http://cognitivesciencesociety.org/supplements/for tables). Could this have been the basis of PF's implicit knowledge? To check this possibility, we considered every response that PF regarded as a guess (for all transforms) and determined whether (interval) bigrams or first order frequencies of intervals (unigrams) could predict PF's responses. The mean unigram match for each test item given a "guess" confidence rating was compared to the average unigram match for all items in that test set. If PF was responding on the basis of unigrams he should say "yes" more often to items with high unigram matches than low unigram matches. In fact, unigram match predicted PF's responses 60% of the time. (For the same items, the transform that the item instantiated predicted 62% of PF's answers.) Bigram matches predicted his responses 52% of the time. It is possible therefore that PF's responses were based on first order frequencies of intervals; the data do not distinguish this possibility from the idea that PF could use implicit knowledge of the transforms themselves. A multiple logistic regression with both unigrams and transform (coded 0 for no transform and 1 for any of the transforms being instantiated) as predictors of PF's responses showed non-significant effects of both variables when both were in the equation; thus, we do not know if each variable contributed independently to PF's responses.

6.2. Experiments 2a and 2b

In experiments 2a and 2b, the materials were extremely well balanced in terms of their *n*-gram structure, at least up to tetragrams, as seen in Table 4 for experiment 2a. Even though participants in experiments 2a and 2b could not have been using *n*-gram statistics where *n*-grams are defined independently of position, maybe participants were responsive to *n*-grams in particular positions (e.g., Johnstone & Shanks, 1999; Knowlton & Squire, 1994). Table 5 shows how well the 48 test items could have been classified just using *n*-grams in each position for experiment 2a (the corresponding tables for experiment 2b are displayed at http://cognitivesciencesociety.org/supplements/, and show the same pattern as for experiment 2a). Relying on *n*-grams in a particular position would actually lead to below baseline classification approximately as often as above baseline classification. However, if the *n*-gram spans the central position of the tone row, very good classification can be achieved. For each *n*-gram, the position that allows best classification was chosen and the corresponding *n*-gram match was entered into a multiple regression to predict each participant's responses. Some of the

Table 5
Classification (/48) using *n*-gram intervals starting at position *X*

X	Unigram	Bigram	Trigram
1	28	19	22
2	20	22	22
3	25	22	22
4	25	24	32
5	20	36	40
6	37	34	32
7	23	15	19
8	19	17	20
9	25	20	20
10	22	21	
11	23		
Mean	23.9 (49.8%)	23.0 (47.9%)	25.4 (53.0%)

Materials for experiment 2a.

test items also had the same interval sequence as one of the training items. A binary variable coding whether there was an interval match with a training item was also entered. Whether the test item was a transpose or inverse retrograde was also entered. The dependent variable was a 1 or a 0, depending on whether the participant endorsed that item or not. Fig. 6 shows the regression slopes for the experienced participants of experiment 2a (confidence intervals calculated over participants).

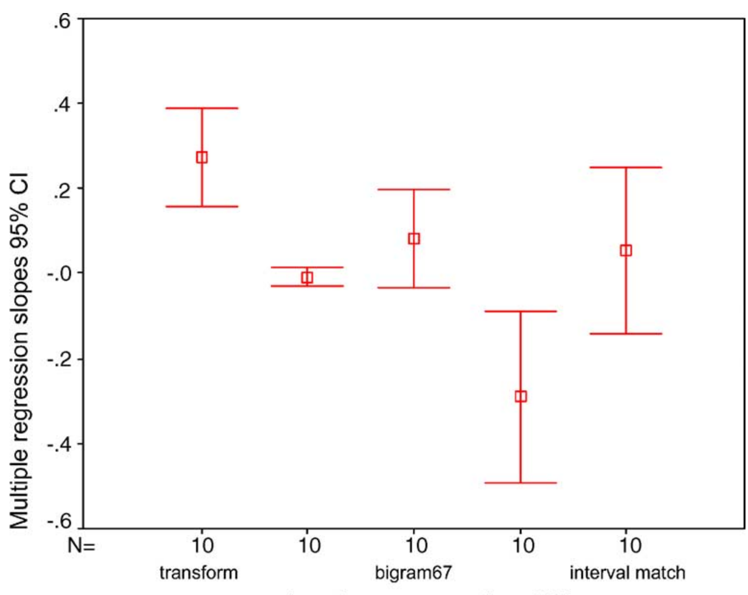

Fig. 6. Multiple regression coefficients for predicting the responses of the experienced participants in experiment 2a.

Fig. 6 shows that in experiment 2a the transform remained a reliable predictor of participants' responses, partialing out all the other variables. Participants seemed to be sensitive to the transform independently of the *n*-gram variables. The results for experiment 2b were inconclusive; the effect of transform was non-significant, but the confidence interval was also wide enough to include the size of slope found in experiment 2a. With only five participants in experiment 2b, the data do not allow us to say exactly what the participants were responding to.[7]

7. Discussion

The aim of this research was to determine the learnability of the transforms of serialist music. We found that participants who had no specialized training in atonal music could not (in a short period of time) come to perceive the transforms of serialist music implicitly. However, participants with routine exposure to atonal music could come to perceive the distinction between pitch class transpose and inverse retrograde implicitly; further PF, a participant with considerable prior atonal music experience, could recognize all the transforms explicitly and, while attempting to apply explicit strategies, also be accurately guided by implicit knowledge. We will first consider what exactly was learned, and then the implicit nature of the acquired knowledge.

The extent to which participants may have been responding on the basis of *n*-gram statistics was explored. PF may have implicitly learnt about first order frequencies rather than the transform; the data do not discriminate these possibilities. However, the use of *n*-grams independent of position could not explain the learning of participants in experiments 2a and 2b. Further, the use of *n*-grams in particular positions was plausibly ruled out for participants in experiment 2a. This is an important conclusion because, as reviewed above, at least a majority of learning in the published implicit learning literature can be accounted for by presuming subjects learn only *n*-grams (cf. Perruchet, 1994). What else then could people have been learning? The current data leave open a number of possibilities to be explored. These possibilities include, in order of abstractness (the top level corresponding to the musical transforms per se): look-up tables; associative value–value mappings across non-contiguous positions; variable–variable mappings between sequences of fixed length; variable–variable mappings between sequences of any length. We will consider each possibility in turn.

Musically experienced people, like PF could simply have formed a look-up table so vast it contains, for example, at least some of the specific sequences in our study embodying each transform. In the course of his life, PF may have simply rote learned some of the test stimuli as being examples of the transforms they instantiate. (Out of interest, PF himself does not regard this as plausible of himself.) This would pose no problem in principle to current models of implicit learning because the suggested learning process amounts to a case of rote memory of items belonging to set categories.

Further research could try training inexperienced participants over an extended period of time with training stimuli different from the eventual test stimuli. If the participants could eventually implicitly classify the test stimuli, this would show that people could classify items not in any potential look up table. In an initial exploration in that direction, we trained inexperienced participants on the training material of experiment 2a, but for five times as long as experiment

2a. Twelve participants took a tape home and were asked to play it themselves once each day for five days, or until they had heard it five times. When they returned for the test phase, their classification performance was 51.7% ($SE = 1.6\%$), not significantly above chance, $t(11) = 1.06$. Clearly, more extensive training or simpler stimuli are needed for learning to be observed.

People need not learn a vast look-up table to learn to classify the test stimuli. People could instead learn the association of a particular value of, for example, a pitch interval in one position (e.g., the first position in the first hexachord) with a particular value in another position (e.g., the first position in the second hexachord). Such value–value association implies learning an arbitrary mapping between intervals, like a third maps onto a fourth, would be just as easy as learning a regular mapping, like the identity mapping (e.g., a third maps onto a third). At the most arbitrary, the mapping could be learnt separately for each position. This is the sort of mapping achieved by many connectionist networks (e.g., Dienes, 1992). In terms of existing paradigms in the published implicit learning literature, it appears people cannot learn a straightforward statistical association between sequence elements that are six positions apart in the space of a few minutes. In the standard artificial grammar learning paradigm, for example, people learn *n*-grams of a length up to tetragrams with extended training (Johnstone & Shanks, 1999). In the sequential reaction time task, people also learn *n*-grams no higher than tetragrams, even after 60,000 trials of training (Cleeremans, 1993). Consistently, Mathews et al. (1989) and Johnstone and Shanks (2001) found that people did not implicitly learn biconditional grammars with arbitrary mappings between letters that are four letters apart. However, the instructions in experiment 2 and the musical nature of the stimuli may have focussed attention on corresponding positions in the two hexachords, allowing associations to form between corresponding positions (cf. Whittlesea & Dorken, 1993). Alternatively, it may be that the musical stimuli in experiments 2a and 2b could be learnt precisely because the mapping was not a mapping between arbitrary values, but a definite and simple operation, and indeed the same operation for all positions.

The serialist transforms are what Marcus (2001) calls “operations over variables”. In terms of pitch classes, the variables are the hexachord sequence of pitch classes and the operations are the arithmetic operations specified in the introduction. In terms of pitch class intervals, the variables are the ordered sequence of intervals in each hexachord. In the transpose and inverse retrograde, the intervals are mapped by the identity operation. In the inverse and retrograde, the intervals are mapped by negation. An operation over variables involves a mapping between variables regardless of their values, rather than mappings from certain values of a variable to certain values of another variable. If people had learnt the transform per se, they could be trained on, for example, transposes involving only certain pitch classes (or intervals), but still classify test stimuli with new pitch classes or intervals. This would test if people had learnt the variable mapping rather than a mapping between particular values.[8] If people had learnt just mappings between values they would not generalise to test stimuli instantiating the transform with new pitch classes or pitch class intervals (at least they would show a generalisation gradient). The current experiments were not designed to tease these two possibilities apart (i.e., value–value mappings vs. variable–variable mappings), but future experiments could do so.

The transforms of transpose, inverse, retrograde and inverse retrograde are most generally operations over variables of any length. That is, while we, for example, transposed sequences of only one length, namely six notes, a transpose in general can apply to a sequence of any length. It is an interesting question whether people can implicitly learn a musical transform for, in principle, any length. If people only generalise to sequences of the same length they are trained on, the learning can be modelled by a finite state device. If people can extrapolate to sequences with lengths greater than those trained on, there would be evidence for learning a grammar more powerful than finite state. The transforms are structurally like certain recursive constructions in natural languages that have been used to argue that natural languages are more than finite state. Retrogrades and inverse retrogrades are structurally like centre embedding in English; transposes and inverses are structurally like the "respectively" construction in English or cross dependency in other languages. According to the Chomsky hierarchy of grammars, centre embedding requires a context free grammar and cross dependency constructions require a context sensitive grammar (or, in any case, more than context free) (e.g., Shieber, 1985).

Interestingly, existing data give clues to the order of difficulty of learning the different transforms that contrasts with the Chomsky hierarchy. Dowling (1972) found that in terms of explicitly detecting the transform, inverse (context sensitive) was easier than either retrograde or inverse retrograde (context free). Further, everyday experience in listening to music tells us that transpose (context sensitive) is the easiest of the transforms to detect explicitly; relatedly, tonal music makes extensive use of transposition. Consistently, the performance in experiment 2a (transposes: context sensitive) was significantly higher than performance in experiment 2b (inverse retrogrades: context free). Sensitivity to which positions map onto which others may arise from the way sequences are stored and retrieved. If the first hexachord is put in a buffer, a first-in-first out buffer will facilitate detecting that the *x*th position in the first hexachord predicts the *x*th position in the second (like transpose or inverse, and in natural language, like the respectively and cross dependency constructions: context sensitive). On the other hand, a last in-first out buffer will facilitate detecting retrogrades and inverse retrogrades (like centre embedding in natural language: context free). If an implicit learning mechanism uses a first-in-first out buffer, it may be able to learn constructions regarded as more than context free, and be unable to learn constructions regarded as merely context free.

In summary, we have established that participants learnt more than *n*-grams, and indicated a hierarchy of abstractness on which future research could locate implicit learning of musical transforms. Clearly, locating implicit learning on that hierarchy will have important implications for how it should be modelled.

The experiments relied on a particular approach to measuring the implicitness of the knowledge, namely using subjective measures (guessing criterion and zero correlation criterion). These measures have their assumptions, which are discussed in detail in Dienes (in press) and Dienes and Perner (in press). For example, if participants deterministically applied a set of rules so that they always gave the same answer to the same string, a null relation between confidence and accuracy does not necessarily mean the knowledge is unconscious. In the standard artificial grammar learning paradigm, we know participants do not respond deterministically to strings (Dienes, Kurz, Bernhaupt, & Perner, 1997; Reber, 1989). However, it is unlikely that the null relation between confidence and accuracy in the current paper was due to the participants applying incorrect or partially correct explicit rules; the above chance performance when subjects thought they were literally guessing still needs to be explained. Participants were told that "guessing" means their answer had no basis or justification whatsoever. Given the significant guessing criterion, the zero-correlation criterion indicated that subjects were no more correct when they thought that they knew than when they thought that they were guessing; they literally could not discriminate states of guessing from states of knowledge. Whatever it was that people learnt in experiment 2, we argue that the knowledge was implicit.

The experienced participants were selected by us to maximize the likelihood they could perceive the transforms, and so they were selected to be people who like to listen to atonal music. They likely differed from the inexperienced participants along a number of dimensions though, not just in their experience of atonal music. They may have had superior perceptual skills in various ways. Possessing perfect pitch is not logically relevant to performing the tasks because the transforms are defined over intervals; the absolute label given any particular pitch does not matter. Conversely, being able to perceive intervals precisely would make a major difference to the ability to perform the task. For current purposes, it is enough to know the transforms can be implicitly learnt. Future research can investigate exactly which characteristics were important in allowing experienced rather than inexperienced participants to perform well.

A key musical question regarding serialist music is whether the same tone row can be recognized as the same under the transformations; this is the question of whether a tone row can provide the unity that Schoenberg argued for. PF's ability to classify the materials of experiment one, in which a hexachord was either followed by a transform of itself, or by a random rearrangement of the transform, may reflect the ability to hear a tone row as a unified object despite transformation. Conversely, experiment 2 investigated the ability to perceive the different transformations as different transformations. In that experiment, each hexachord was followed by a transform of itself, and the task was to distinguish the transforms. Simply being able to see the unity of a hexachord under transformation would not enable above chance classification in experiment 2; it is the distinction between transforms rather than the unity of the hexachord that must be perceived. Jointly, the results of PF and experiment 2 indicate the ability of experienced subjects to have some perception of the unity of a tone row, and also to distinguish the various transforms of it. Presumably, both abilities contribute to musical experience.

Notes

1. We take it that it is desirable for the piece to be comprehensible and coherent to the listener, but Schoenberg did not explicitly specify this.
2. For example, in modulo 12 arithmetic (clock face arithmetic), $11 + 2 = 1$. But in "normal arithmetic" (non-modulo arithmetic), $11 + 2 = 13$. Hence, in a pitch class transpose, the Bb above middle C could be transposed by +2 to the Db just above middle C, actually dropping 10 semitones. In a pitch transpose, a transposition of +2 semitones means the pitch must increase by literally +2 semitones. A pitch transpose is just one instantiation of a pitch class transpose.
3. Subjective measures have a long history in psychology. In subliminal perception, they have been used by Peirce and Jastrow (1884), and more recently by, for example, Merikle (1992), Cheesman and Merikle (1984, 1986), Kolb and Braun (1995), and Kunimoto, Pashler, and Miller (2001). In implicit learning, they have been used by, for example, Dienes, Altmann, Kwan, and Goode (1995), Dienes and Altmann (1997), Allwood,

Granhag, and Johansson (2000), Kelley, Burton, Kato, and Akamatsu, (2001), Tunney and Altmann (2001), Channon et al. (2002), Dienes and Perner (2003), and Tunney and Shanks (2003).

4. There was no quadratic component in the plot of accuracy against confidence, $p > .10$.
5. Collapsing over participants, the intercept in Fig. 2 at a confidence of 50% was 59.8%, significantly above chance $t(4) = 3.50$, $p = .025$. The slope (−0.2%) was non-significantly different from zero, $t(4) = 1.83$, $p = .14$.
6. Collapsing over participants, the intercept in Fig. 3 at 50% confidence is significantly above 50% (58.5%), $t(4) = 2.50$, $p = .034$, 1-tailed. The slope (−0.3%) is not significantly different from zero, $t(4) = 2.33$, $p = .08$.
7. Brooks and Vokey (1991) argued that participants could be sensitive to matches in the repetition structure between test and training items. For example, the letter string MMXMTX has the repetition structure 112132 indicating that the first element is repeated in the second and third positions and the second novel element is repeated in the final position (see also e.g., Tunney & Altmann, 2001). For the stimuli in experiment 2a, no pitch class is repeated, so the repetition structure over pitch classes is identical for all tone rows in the test phase. The repetition structure over pitch class intervals was determined for the training and test tone rows. For each test tone row, the number of times its repetition structure matched that of a training tone row was determined. In fact, this allowed perfect discrimination between the transposes and inverse retrogrades in the test set. For the inverse retrogrades, each tone row had zero matches with transpose training tone rows. Transposes had between 1 and 20 matches. When the number of matches was entered into a multiple regression, together with transform and the *n*-gram statistics entered in the regression reported in the text, with participant's response as the dependent variable, transform remained a highly significant predictor (its regression coefficient was positive for all 10 experienced participants, so $p = (.5)^{10}$ by binomial, 1-tailed); repetition structure match had in fact an overall negative regression slope. In sum, transform remains a significant predictor of participants' responses when repetition structure is partialed out.
8. Grammatical stimuli were constructed so as to be entirely regular with respect to operations on pitch classes or pitch class intervals. Nonetheless, the stimuli would also have instantiated operations over pitches or pitch intervals in a noisy way. For example, in any of the transpose stimuli, a fragment of the second hexachord would have been a pitch transpose of the corresponding fragment of the first hexachord (i.e. a transpose with normal arithmetic rather than modulo 12 arithmetic). Participants may have implicitly learnt a noisy pitch transpose rather than a pitch class transpose.

Acknowledgments

Sadly, Christopher's health began to deteriorate during this work, and he died on 27 March 2004. This is the last paper by the person who coined the term "cognitive science".

Thanks to Georgina Jones, Cecilie Gudim, and James Willis for running participants in experiments 1 and 2, and Alastair Goode for other help. The research was partly funded by the Economic and Social Research Council.

Appendix A. Material for annexe

http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/music%20materials.html contains the training and test materials for experiments 1 and 2 with audio samples. It also contains the original 35 tone rows culled from serialist compositions and the 11 hexachords selected for experiment 2.

http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/music%20ngrams.htm contains tables, figures and further analyses of the *n*-gram structure of the materials in experiments 1 and 2.

http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/smooth.html contains a description of the "smooth experiment" described in the introduction to experiment 2, and all materials (with audio samples) and results.

References

Allwood, C. M., Granhag, P. A., & Johansson, H. (2000). Realism in confidence judgements of performance based on implicit learning. *European Journal of Cognitive Psychology*, *12*, 165–188.

Altmann, G., Dienes, Z., & Goode, A. (1995). On the modality independence of implicitly learned grammatical knowledge. *Journal of Experimental Psychology: Learning, Memory,* & *Cognition*, *21*, 899–912.

Balch, W. R. (1981). The role of symmetry in the good continuation ratings of two-part tonal melodies. *Perception* & *Psychophysics*, *29*, 47–55.

Berry, D. C. (1997). *How implicit is implicit learning?* Oxford: Oxford University Press.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hove: Lawrence Erlbaum.

Bigand, E., D'Adamo, D., & Poulin, B. (2003). *The implicit learning of a contemporary musical grammar*. Paper presented at the at the thirteenth meeting of the European Society for Cognitive Psychology, Granada, Spain, 17–20 September 2003.

Bigand, E., Perruchet, P., & Boyer, M. (1998). Implicit learning of an artificial grammar of musical timbres. *Current Psychology of Cognition*, *17*, 577–600.

Boucher, L., & Dienes, Z. (2003). Two ways of learning associations. *Cognitive Science*, *27*, 807–842.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General*, *120*, 316–323.

Channon, S., Shanks, D., Johnstone, T., Vakili, K., Chin, J., & Sinclair, E. (2002). Is implicit learning spared in amnesia? Rule abstraction and item familiarity in artificial grammar learning. *Neuropsychologia*, *40*, 2185–2197.

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception* & *Psychophysics*, *36*, 387–395.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, *40*, 343–367.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton & Co.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.

De Lannoy, C. (1972). Detection and discrimination of dodecaphonic series. *Interface*, *1*, 13–27.

DeLone, R., Kliewer, V., Reisberg, H., Wennerstrom, M., Winold, A., & Wittlich, G. (1975). *Aspects of 20th century music*. New Jersey: Prentice-Hall.

Deutsch, D. (1982). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (pp. 271–316). London: Academic Press.

Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 349–411). London: Academic Press.

Dienes, Z. (1992). Connectionist and memory array models of artificial grammar learning. *Cognitive Science*, *16*, 41–79.

Dienes, Z. (in press). Assumptions of subjective measures of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies*.

Dienes, Z., & Altmann, G. (1997). Transfer of implicit knowledge across domains? How implicit and how abstract? In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 107–123). Oxford: Oxford University Press.

Dienes, Z., Altmann, G., & Gao, S.-J. (1999). Mapping across domains without feedback: A neural network model of implicit learning. *Cognitive Science*, *23*, 53–82.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory* & *Cognition*, *21*, 1322–1338.

Dienes, Z., & Berry, D. C. (1997). Implicit synthesis. *Psychonomic Bulletin and Review*, *4*, 68–72.

Dienes, Z., & Fahey, R. (1995). The role of specific instances in controlling a dynamic system. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 848–862.

Dienes, Z., & Fahey, R. (1998). The role of implicit memory in controlling a dynamic system. *Quarterly Journal of Experimental Psychology*, *51A*, 593–614.

Dienes, Z., Kurz, A., Bernhaupt, R., & Perner, J. (1997). Application of implicit knowledge: Deterministic or probabilistic? *Psychologica Belgica*, *37*, 89–112.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, *22*, 735–808.

Dienes, Z., & Perner, J. (2003). Unifying consciousness with explicit knowledge. In A. Cleeremans (Ed.), *The unity of consciousness: Binding, integration, and dissociation*. Oxford: Oxford University Press.

Dienes, Z., & Perner, J. (in press). Assumptions of a subjective measure of consciousness: Three mappings. In R. Gennaro (Ed.), *Higher order theories of consciousness*. John Benjamins Publishers.

Dowling, W. J. (1972). Recognition of melodic tranformations: Inversion, retrograde and retrograde inversion. *Perception and Psychophysics*, *12*, 417–421.

Francès, R. (1958/1988). *The perception of music* (W. J. Dowling, Trans.). Hove: Erlbaum.

French, R. M., & Cleeremans, A. (2002). *Implicit learning and consciousness: An empirical, philosophical and computational consensus in the making?* Hove, England: Psychology Press.

Hadley, R. (2001). *Very rapid induction of general patterns*. Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society, August 1–4, Edinburgh, Scotland (pp. 387–392). Mahwah, NJ: Lawrence Erlbaum Associates.

Imberty, M. (1993). How do we perceive atonal music? Suggestions for a theoretical approach. *Contemporary Music Review*, *9*, 325–337.

Jiménez, L. (2003). Intention, attention, and consciousness in probabilistic sequence learning. In L. Jiménez (Ed.), *Attention and implicit learning* (pp. 43–70). Amsterdam: John Banjamins.

Johnstone, T., & Shanks, D. (1999). Two mechanisms in implicit artificial grammar learning? Comment on Meulemans and van der Linden (1997). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 524–531.

Johnstone, T., & Shanks, D. R. (2001). Abstractionist and processing accounts of implicit learning. *Cognitive Psychology*, *42*, 61–112.

Kelley, S. W., Burton, A. M., Kato, T., & Akamatsu, S. (2001). Incidental learning of real world regularities in Britain and Japan. *Psychological Science*, *12*, 86–89.

Knowlton, B., & Squire, L. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 79–91.

Kolb, F. C., & Braun, J. (1995). Blindsight in normal observers. *Nature*, *377*, 336–338.

Krumhansl, C. L. (2000). Tonality induction: A statistical approach applied cross-culturally. *Music Perception*, *17*, 461–479.

Krumhansl, C. L., Sandell, G. J., & Sergeant, D. C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, *5*, 31–78.

Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness and Cognition*, *10*, 294–340.

Lerdahl, F. (1988). Cognitive constraints on compositional systems. In J. Sloboda (Ed.), *Generative processes in music: The psychology of performance, improvisation, and composition* (pp. 231–295). Oxford: Oxford University Press.

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

Lewicki, P. (1986). *Nonconscious social information processing*. New York: Academic Press.

Longuet-Higgins, H. C. (1976). The perception of melodies. *Nature*, *263*, 646–653.

Longuet-Higgins, H. C. (1978). The perception of music. *Interdisciplinary Science Reviews*, *3*, 148–156.

Marcus, G. (2001). *The algebraic mind*. Cambridge, MA: MIT Press.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchard-Fields, F., Cho, J. R., & Druhan, B. (1989). The role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1083–1100.

Merikle, P. M. (1992). Perception without awareness: Critical issues. *American Psychologist*, *47*, 792–795.

Pederson, P. (1975). The perception of octave equivalence in twelve-tone rows. In D. Sergent (Ed.), *Psychology of music* (pp. 3–8). Society for Research in Psychology of Music and Music Education.

Peirce, C. S., & Jastrow, J. (1884). On small differences in sensation. *Memoires of the National Academy of Sciences*, *3*, 73–83.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks (1992). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 223–228.

Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, *119*, 264–275.

Perruchet, P., & Vinter, A. (2002). The self-organising consciousness: A framework for implicit learning. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness: An empirical, philosophical and computational consensus in the making?* (pp. 41–67). Hove, England: Psychology Press.

Reber, A. S. (1989). Implicit learning and tactic knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Redington, M., & Chater, N. (1996). Transfer in artificial grammar learning: A re-evaluation. *Journal of Experimental Psychology: General*, *125*, 123–138.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.

Schoenberg, A. (1941). *Style and idea: Selected writings of Arnold Schoenberg*. London: Faber and Faber.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 592–608.

Shieber, S. (1985). Evidence against the context free-ness of natural language. *Linguistics and Philosophy*, *8*, 333–343.

Smith, J. D., & Witt, J. N. (1989). Spun steel and stardust: The rejection of contemporary compositions. *Music Perception*, *7*, 169–186.

Stadler, M. A., & Frensch, R. M. (1998). *Handbook of implicit learning*. Thousand Oaks, CA: Sage Publications.

Thomson, W. (1991). *Schoenberg's error*. Philadelphia: University of Pennsylvania Press.

Tunney, R. J., & Altmann, G. T. A. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 614–639.

Tunney, R. J., & Shanks, D. R. (2003). Subjective measures of awareness and implicit cognition. *Memory and Cognition*, *31*, 1060–1071.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic processing account of implicit learning. *Journal of Experimental Psychology: General*, *122*, 227–248.

Wittlich, G. (1975). Sets and ordering procedures in twentieth century music. In R. DeLone, V. Kliewer, H. Reisberg, M. Wennerstrom, A. Winold, & G. Wittlich (Eds.), *Aspects of 20th century music* (pp. 388–476). New Jersey: Prentice-Hall.

Wuorinen, C. (1979). *Simple composition*. London: Longman.