Goldsmiths, University of London

Candidate number: 33230950

Supervisors: Professor Adam Ockelford (University of Roehampton, London) and Professor Linda Pring (Goldsmiths)

Title: Musical memory and expectation in an individual with savant syndrome

Submitted: 10th September 2012

Word count: 9767[1]

Thesis submitted in partial fulfilment of an MSc in Music Mind and Brain

This project was completed in collaboration with the University of Roehampton

[1] **Excluding abstract, acknowledgements, headers, footnotes, headings, captions, figures, tables, references and appendices**

Table of contents

Lists of tables and figures iii

Abstract iv

Acknowledgements v

Note on collaboration v

Introduction 1

Expectation in music 4

The model 9

Method 11

Participant 11

Design 11

Materials 12

Stimulus piece 12

Isolation of segments 12

Equipment 14

Procedure 14

Procedure for each session 15

Scoring system 15

Imitation threshold task 16

Materials 16

Procedure 17

Scoring 17

Results 17

Results 18

Analysis of whole dataset 18

Item analysis 20

Discussion 21

Expectancy and memory **25**

Implications for the model **27**

Implications for music and expectancy research **27**

Comparison with previous *Fragments of Genius* study **28**

Further implications for future research **29**

Conclusion **30**

References **31**

Appendices

A(i). Composition criteria for *Romantic Rollercoaster* **35**

A(ii). Musical structure of *Romantic Rollercoaster* **36**

B. Criteria for isolation of segments **37**

C. Scoring system **38**

D. SPSS outputs **39**

List of tables

Table 1. Allocation of expectancy ratings to notes **9**

Table 2. System for grouping the 6 expectancy ratings into 4 categories **12**

Table 3. Timetable for trials **15**

Table 4. Spearman Rho Correlations **20**

List of figures

Figure 1. Model of expectation in music (after Ockelford, 2006) **8**

Figure 2. Score of *Romantic Rollercoaster* **13**

Figure 3. Calculation of proportional accuracy **16**

Figure 4. Calculation of proportional accuracy in the imitation threshold task **17**

Figure 5. Overall proportional accuracy for all notes in each trial **18**

Figure 6. Mean proportional accuracy for notes of each expectancy rating **19**

Figure 7. Mean proportional accuracy for segments of each expectancy category **20**

Abstract

This study investigated the role of musical expectations when an exceptionally talented pianist with savant syndrome (Derek Paravicini) attempted to learn a new piece of music. Derek used a 'play along' method of learning for 10 trials over a 2-month period. Results were analysed using a model of expectation in music adapted from Ockelford (2006), based on zygonic theory (e.g. Ockelford 2005, 2006, 2008, 2011). It was hypothesised that notes with higher expectancy ratings (according to the model) would be predicted more accurately by Derek during his first attempt to play along with the unfamiliar piece, and that an interaction between expectancy and proportional accuracy would occur over subsequent attempts. The first hypothesis was confirmed. Data suggested an interaction effect across trials, although this was not statistically significant. Implications for research into memory and expectation in music are discussed.

Acknowledgements

Many thanks to Derek Paravicini, for his participation in the study. Sincere thanks also to Adam Ockelford, Linda Pring, Evangelos Himonides and Tim Pring.

Note on the nature of collaboration with the University of Roehampton and Institute of Education

As this study was one in a series of four, the experimental paradigm and stimulus piece of music were predetermined. The initial recording of both the stimulus piece and the imitation threshold stimulus were carried out by Adam Ockelford (University of Roehampton, London) and Evangelos Himonides (Institute of Education, London). Production of recorded stimuli and conversion of file formats were also performed by Evangelos Himonides.

Musical memory and expectation in an individual with savant syndrome

Individuals with savant syndrome display an area of highly developed ability, in the context of developmental disabilities, including autism (Pring, 2005; Treffert, 2009). It has been argued that this ability is always related to a highly developed memory for the relevant domain (e.g. Pring, 2008; Treffert, 2009). Savant skills have been demonstrated in several areas, including art (Corrigan, Richards, Treffert & Dager, 2012), calendar calculation (Heavey, Pring & Hermelin, 1999), and music (Hermelin, Oconnor, & Lee, 1987; Ockelford, 2007a; Young & Nettelbeck, 1995).

Derek Paravicini is an extraordinarily talented pianist with savant syndrome, who enjoys performing for others. He is congenitally blind (due to retinopathy of prematurity) and has severe learning difficulties, with a diagnosis of autism (Ockelford, 2007a; Pring, Woolf & Tadic, 2008). He has absolute pitch (Pring, Woolf, & Tadic, 2008) and the ability to disaggregate complex chords consisting of (up to) 10 notes (the physical limitations of playing more than 10 notes at a time may mean that his aural skill extends to even more complex chords; Ockelford, 2008).

This project formed part of the *Fragments of Genius* REMUS (Researching Exceptional Musical Skill) research, which explores musical savant skills, in particular 'learning, memory, musical reproduction and creativity' (Ockelford, 2011, p. 31). Following on from studies such as Sloboda (1985), *Fragments of Genius* is based on the premise that musical analysis of what musicians with savant syndrome play gives insight into their mental processes (Ockelford, 2011, pp. 31-32). This is intended to inform pedagogy both with this group and others, as well as enriching our understanding of musicality (Ockelford & Pring, 2005). The present paper relates to one in a series of four studies involving Derek. The series will provide the first known formal exploration of long-term memory in a musician with savant syndrome (Ockelford, 2012). In each component study, Derek learns a new piece on the piano. Each time, a different learning method is used:

1) 'Listen and play' (e.g. Ockelford, 2008; Ockelford, 2011, and see below)

2) 'A bit at a time' (Derek hears a short section of the piece and is asked to play it before hearing the next section)

3) 'Play along' (discussed in this paper). Derek will learn a piece entitled *Romantic Rollercoaster*.

4) ‘Just listen’ (Derek will hear the piece in regular sessions but will not be asked to play it until session 8)

The ‘listen and play’ study is well documented (Ockelford, 2007b, 2008, 2011; Ockelford & Pring 2005) and revealed much about the interaction between Derek’s musical memory and his creativity. 26 trials took place over a 4 year period. In the first trial, Derek heard a recording of the stimulus piece (entitled *Chromatic Blues*) for the first time and was then asked to play what he could remember of it. In each subsequent trial Derek played what he could remember of the stimulus piece before hearing a recording of it. He was then asked to play the piece again, but was not exposed to the recording between trials, and to the authors’ knowledge did not attempt to play the piece outside of sessions. A method of musical analysis based on Ockelford’s zygonic theory (e.g. Ockelford, 2006) was used to calculate ‘indices of derivation’ of Derek’s attempts from the original stimulus (i.e. the proportion of Derek’s playing that was a replication of *Chromatic Blues*). After 9 trials, around three quarters of Derek’s playing was exactly the same as the stimulus piece (Ockelford, 2011). The results of the first few trials highlighted how, when pushed beyond the limits of his working memory, Derek took the combined influences of the stimulus piece and other known pieces in a similar style to produce something that made ‘musical sense’ (Ockelford, 2011, p. 49).

Adam Ockelford (Derek’s teacher and mentor) has suggested that Derek’s outstanding musical ability means that he is ‘…able to offer direct musical evidence of his thinking in a way that is far beyond the capacity of most people.’ (Ockelford, 2007b, p. 10). Where the ‘listen and play’ study allowed for an insight into Derek’s musical creative processes, the present investigation will focus on Derek’s ability to predict musical events based on preceding structure, and the impact of musical predictability on subsequent memory. It will be argued that conclusions drawn here have relevance for our understanding of music processing more generally, in particular musical expectations, and their importance for memory.

‘Play along’ is an unusual method of learning instrumental music (it is much more common in a vocal setting), yet is often favoured by young musicians with learning disabilities (Ockelford, personal communication, July 2012). To the author’s knowledge, there has been no formal exploration of the processes involved. It is hypothesised that engaging in the ‘play along’ method means that sound must be transferred from the ‘preattentive’ sensory store of echoic memory (Kaernbach, 2004, p. 240), through working memory

(Baddeley, Chincotta, & Adlam, 2001) before it is stored in long term memory. Previous studies have demonstrated the startling accuracy of Derek's echoic memory and his ability to play back complex aural stimuli (Ockelford, 2008; Pring et al., 2008), but the relationship between this skill and his long term memory requires investigation.

'Play along' has some similarities with psychological 'shadowing' studies in which an aural stimulus is to be remembered. However, such studies have generally included a 'distractor' stimulus, in order to examine attention resources (e.g. Banks, Roberts, & Ciranni, 1995; Spence & Read, 2003). The closest parallel task to 'play along' is simultaneous interpretation, which is 'the act of transforming sentences from one language into sentences of a second language in a fluent, continuous manner' (Isham & Lane, 1993, p. 241). When engaging in 'play along', Derek must immediately 'interpret' the auditory input into a motor action. In language, the process of interpretation disrupts subsequent memory for the stimulus (Daro & Fabbro, 1994; Isham & Lane, 1993). It is hypothesised therefore that long term memory for the stimulus piece (*Romantic Rollercoaster*) learnt using the 'play along' method will not be as accurate as that for pieces learnt using the other methods under investigation. Derek's attempts to reproduce the stimulus at the same time as hearing it will interfere with the formation of memory traces for that material.

The use of the 'play along' method of learning brings other implications regarding Derek's response. In the initial trial, Derek will be asked to play simultaneously with a piece of music he has never heard before. For the first trial at least, it will be impossible for him to be correctly and completely in time with the stimulus; even if every note is accurately imitated, there will be a delay as the sound is processed and converted into a motor action. The frequent and complex changes of harmony throughout *Romantic Rollercoaster* mean that this would not result in a sound which is musically coherent. Derek's musicality and the request that he 'play along' *at the same time* as the recording leads to the assumption that, in order comply with task demands and to avoid creating something unmusical, Derek will try to predict what is coming next in the stimulus. The first trial in this study therefore demands an exploration of musical expectation. In later trials, where Derek is asked to 'play along' with the stimulus, it is predicted that he will be able to remember some material from previous trials but will still rely on expectation to some extent. Subsequent trials will therefore allow an assessment of how musical expectedness influences the transfer of music from working- to long term- memory.

Expectation in music

Since Meyer's influential 'Emotion and Meaning in Music' (Meyer, 1956), much empirical work on human cognition of music has focused on the generation and fulfilment or denial of expectations. Expectation has been argued to be important for affective experiences of music (Huron, 2006; Margulis, 2005) and demonstrated to facilitate processing of specific musical elements (Bharucha, 1987). However, little work has been done on the impact of expectedness on memory for music (though see Schmuckler, 1997).

Bharucha (1987) made the distinction between schematic and veridical expectations in music. This was after Tulving (1972), who differentiated semantic and episodic memory. Schematic expectations are those which reflect a knowledge of and memory for the characteristics of a particular musical style, whereas veridical expectations are for specific, known segments of music (Bharucha, 1987). There has been much research on the formation of cognitive schemata for musical styles (e.g. Loui, Wu, Wessel, & Knight, 2009; Saffran, Johnson, Aslin, & Newport, 1999), approaches to music learning (e.g. Byrne, 2005) and the broad characteristics of music which influence memory (e.g. Eschrich, Muente, & Altenmueller, 2008). In contrast, there has been little empirical work on why some specific sections of music are encoded in memory (generating veridical expectations) more quickly than others (although position within the musical structure has been found to be important by Williamon & Egner, 2004). Furthermore, it has been argued that the application of Tulving's distinction to music should not be made too rigidly, since memory for the 'episode' when the music was first encountered is often lost (Huron, 2006, p. 225). Huron also points out that 'Since all semantic memories begin as episodic information, there must be some transition between the two forms' (p. 225). The present study will ask how quickly memories for specific sections of music are formed, and whether they are encoded in the same way as are schemata of musical styles. This will allow a consideration of the potential overlap between the two types of memory. Therefore, an overview of recent work pertaining to the internalisation of musical schemata will be given before Schmuckler's work concerning expectation and memory is discussed. Following this will be an exploration of attempts to model expectation in music, before Ockelford's (2006) model (to be adapted for use in the present experiment) is introduced.

Schemata can be rapidly learned, without conscious awareness. Saffran et al. (1999) extended previous work on language acquisition to demonstrate this learning mechanism, using musical tones. Two musical 'vocabularies' were created, each consisting of 6 'words'

(sequences of three tones). Listeners were exposed to a continuous stream of tones, repeated twice, made up of 'words' (in random order) from one of the 'vocabularies' for a total of 21 minutes. They were not informed that these sequences were constructed from 'words'. A forced-choice test showed that 'words' from the 'vocabulary' to which participants had been exposed were judged as more familiar. Similar results were found using a head-turn paradigm with 8-month old infants. It is argued that participants extracted information about the intervals between tones, allowing them to encode the 'words' as distinct entities (because intra-word intervals were more frequently occurring than inter-word intervals). It is suggested that the same, probabilistic, learning mechanism is applied to both linguistic and non-linguistic stimuli (Saffran et al., 1999). Similarly, Loui et al. (2009) found rapid learning of a new musical 'grammar' constructed from chord sequences which used the Bohlen-Pierce scale. Electrophysiological correlates were found which suggested a 'shared neural mechanism for processing well-known as well as novel sound patterns' (Loui et al., 2009, p.457). These studies provide evidence for rapid probabilistic learning, which may use the same cognitive resources as operate in our perception of known music.

Listeners find it easier to remember music which belongs to a schema with which they are familiar (Dowling, 1991). An example of this from the savant syndrome literature can be found in Sloboda (1985). Sloboda compared the accuracy with which a musician with savant syndrome, NP, learnt a tonal piece and an atonal piece, and found that memory for the tonal piece was much more accurate. Sloboda attributed this to the structural framework provided by tonality, within which memory could operate (Sloboda, 1985).

In spite of these findings, there has been little research considering the impact of the expectedness of specific sections of music on subsequent memory for them. In an isolated (to the author's knowledge) example, Schmuckler (1997) examined the importance of conformance to a musical schema for memory for unfamiliar music. Three alternative endings were written for each of 14 folk melodies; pitches were changed but rhythmic information maintained. Participants were played the 'study' melody, and then asked to select this melody from a series of four 'test' melodies. Participants' memory scores were compared against expectancy ratings for each of the melody endings, which were provided by a separate group of participants (Schmuckler, 1997, experiment 2). Memory was found to be better for melodies with high expectancy ratings, but there was no significant difference between those with low and medium expectancy ratings. There was, however, a

general trend towards low expectancy melodies being remembered better than medium expectancy melodies.

It is suggested that the slight increase in memory for melodies with low expectancy ratings might be due to the 'von Restorff effect', which states that events isolated from their background are more salient and so are remembered better (von Restorff, 1933). Schmuckler states that a greater range of expectancy ratings would be necessary to confirm this hypothesis (Schmuckler, 1997). Overall, results are interpreted in the context of tonal schemata, and it is argued that these facilitate both expectation and memory in music processing.

This focus on schematic expectations for unfamiliar music has also been dominant in the several attempts to model or explain expectation as experienced in the perception of music (e.g. Bharucha, 1987; Larson, 2004; Pearce & Wiggins, 2006). Although diverse in nature and scope, there is a common emphasis among these models on probabilistic learning of musical schemata, from which musical expectations are generated. By this mechanism, it is argued that listeners predict what is likely to come next on the basis of what, given the context, has happened most often in their past listening experience. For example, Pearce & Wiggins (2006) created and tested a computational model based on this approach. This model was 'exposed' to examples from a particular corpus of music, and then used to generate predictions about which notes would follow in an 'unfamiliar' musical sequence. The predictions made by the model were found to be consistent with participant responses (Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010; Pearce & Wiggins, 2006). In a similar vein, David Huron's *Sweet Anticipation* (Huron, 2006) summarises work which has shown how the probabilities of musical events account (in part) for 'probe tone' profiles, which chart ratings of how well a tone 'fits' a melodic (tonal) context (therefore representing the expectedness of each scale degree; Krumhansl, 1990). For example, Aarden (2003) found that the expectancy of a particular scale degree, according to the 'probe tone' profiles, reflects the frequency with which that scale degree *closes* a musical phrase in a large selection of Western melodies.

Although useful for understanding our responses to unfamiliar music, the accounts described above do not explore expectancy in known music. Ockelford (2006) sought to address this problem by providing a theoretical model of how specific and general memory traces for music interact over time.

Ockelford's (2006) theory of musical expectation is based on his zygonic theory (e.g. Ockelford, 2006, 2007b, 2008, 2011), which is built on the premise that perceived order in music is a result of the derivation of one musical event from another. It emphasises the importance of relationships between perceived aspects of music ('perspects'), which underlie musical coherence. Therefore, the theory allows for an exploration of musical expectation, which is relevant to both familiar and unfamiliar pieces of music, with the underlying principle that it is order in music which leads to expectations (Ockelford, 2006). Figure 1 (p. 8) summarises this model. Ockelford argues that music is perceived as a series of groups, and expectations can be generated both within (via route **A1**) and between groups (routes **A3**, **B2** and **C3**). Expectations are on a continuum from the general (**1** and **2**) to the specific (**3**). Between-groups expectations are generated from working (**A**), semantic (**B**) and veridical (**C**) memory, and the relative contributions of the different types of expectation change as listeners become more familiar with a piece of music (Ockelford, 2006).

Thorpe, Ockelford, & Aksentijevic (2012) tested aspects of the model empirically using a production paradigm. 40 participants, with varying levels of musical training, were played segments of increasing length of an especially composed melody. The first segment was the first note alone; each subsequent segment included one more sequential note. After each segment, participants were asked to sing the note that would best continue the sequence.

Participant responses were pooled and compared against the predictions of three models based on Ockelford (2006). Of these, the model which best fit participant data incorporated three aspects: the adjacency and recency of the sung tone to the relevant preceding pitches (**A1+B2**), and the between-groups current structures veridically encoded in working memory (**A3**). Of these three aspects, the latter (**A3**) was shown to be by far the most important factor for listener predictions (Thorpe et al., 2012).

Although operating with a lot of constraints and assumptions such as the use of single line, diatonic melodies, Thorpe et al. (2012) is a helpful test of the validity of Ockelford's model. It has helped to identify the importance of between-groups, veridical expectations, even within the first hearing of a piece of music. The present study will extend this research to a fuller musical texture, and investigate how the importance of veridical memory traces changes over repeated exposure to a piece of music, from the perspective of a performer given the task of learning the piece, rather than a listener. The model has been adapted to

accommodate for the findings of Thorpe et al. (2012) and the changes in the experimental paradigm, and is described below (p. 9).

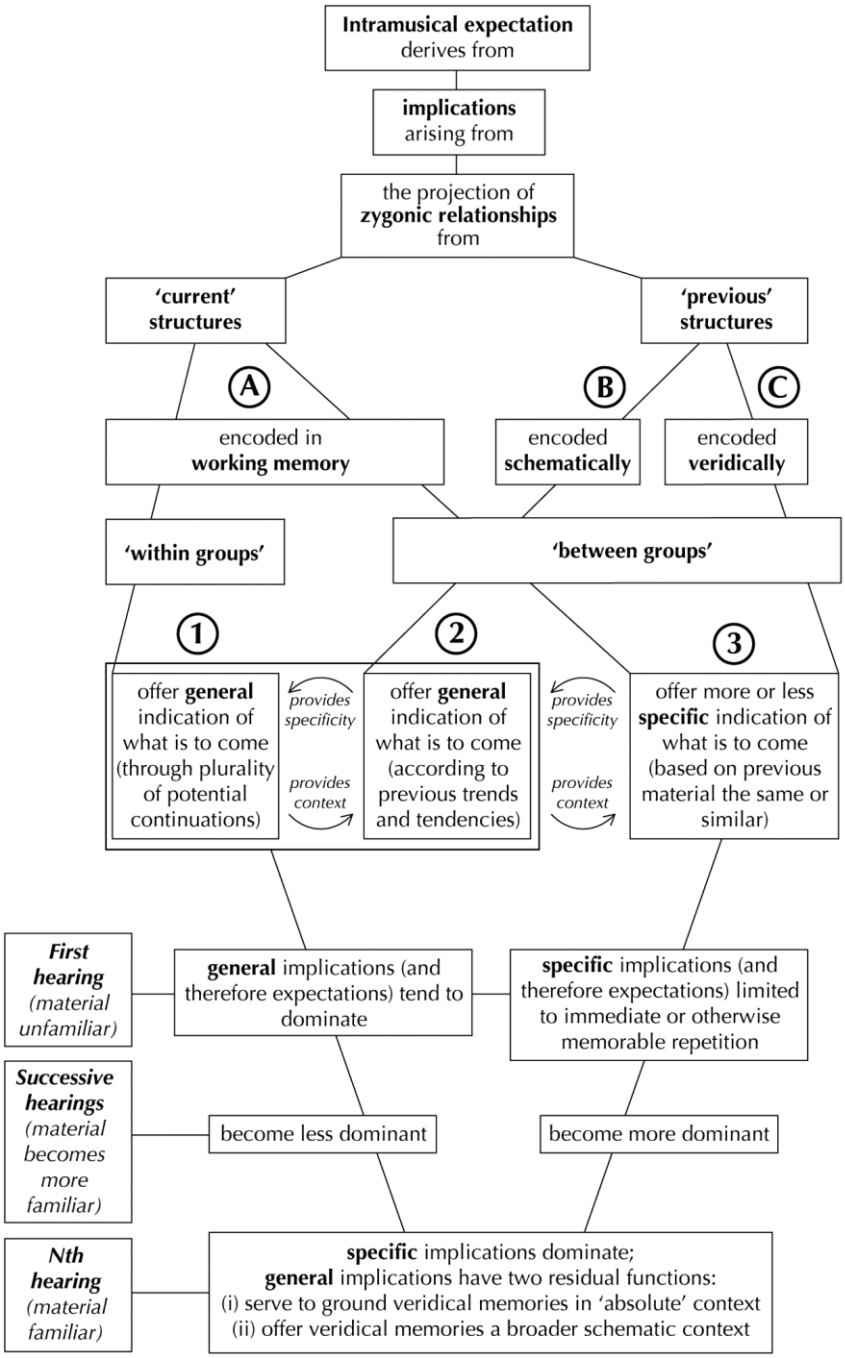


Figure 1: Model of expectation in music (after Ockelford, 2006; reproduced with permission)

It is important to note that the stimulus piece used in this study, *Romantic Rollercoaster,* was purposely written in non-conformance to tonal schemata, therefore minimising the extent to which Derek will be able to rely on schematic information for his musical predictions. Although loosely in the key of A minor, the tonal centre of the piece shifts

constantly, such that there is never tonal stability. These modulations themselves do not conform to any schematic patterns, so the relationships between the keys could not be predicted via schematic expectations (see figure 2, p. 13). As a result, Derek will be unable to base his expectations on relationships between scale steps, nor between keys. Therefore, any expectations demonstrated by Derek must be the result of musical relationships intrinsic to the simulus piece.

The Model

Three factors were taken into account by the model of musical expectation used for the present study; between-groups expectations, adjacency and recency.

Between-groups expectations. A group is defined as having three or more sequential events (notes), within which the 'perspects' (see above, p. 7) are pitch and inter-onset intervals. Notes are classified as predicted by a previous group where they contain a repeat of 'perspects', irrespective of whether or not the group is repeated at the same absolute pitch. Repetitions of groups are recognised as such after the first melodic interval (two events). Where the same group occurs more than twice in a series, the difference in pitch and inter-onset intervals between the first two occurrences is used to predict the pitch and inter-onset intervals between subsequent pairs of groups.

Notes can be predicted by any number of previously occurring groups. As such, each note of *Romantic Rollercoaster* (the stimulus piece) was assigned one of six expectancy ratings. The ratings were dependent on the number of between-groups patterns which predict the occurrence of that note, as follows in table 1:

Table 1

Allocation of expectancy ratings to notes, according to the number of patterns which predict that note.

Number of patterns	Expectancy rating
0	0
1	.2
2	.4
3	.6
4	.8
5 or more	1

Adjacency/ recency. Where a note could be predicted differently according to two (or more) previous groups, the model prioritises groups which are:

- At the same absolute pitch as the current group
- Most recently occurring

However, the analyses presented in this paper will consider only the impact of between-groups expectations.

Thus, this paper will investigate the relationship between the expectancy rating of each note of *Romantic Rollercoaster* (according to the model described above) and Derek's ability to predict those notes, in his first encounter with the piece. The paper will also examine the impact of expectancy rating on memory for notes, in Derek's subsequent attempts to play along with *Romantic Rollercoaster.* To facilitate this, there will be an item analysis of selected segments of the piece. Inclusion of data from across the learning period (trials 1, 2, 5 and 10) will enable an investigation of whether there is any interaction between expectancy rating and memory for notes of the stimulus. These analyses require that the data be filtered to include only those notes which Derek accurately predicted, or remembered, in each trial, and not those which he copied upon hearing them in the recoding. It will therefore be necessary to find out how quickly Derek can copy isolated chords from *Romantic Rollercoaster* (his 'imitation threshold'; see p. 16). Memory trials will enable us to ascertain if and when Derek completely learns the piece. Until he is accurate in his memory for *Romantic Rollercoaster*, it can be assumed that notes played correctly and before the imitation threshold represent both remembered and predicted notes (however, a full discussion of the memory trials is beyond the scope of this paper, which will instead focus on the trials in which Derek plays along with the stimulus).

Following Thorpe et al. (2012), it is hypothesised that in trial 1, notes with higher expectancies will be more frequently and more accurately predicted. Derek should score near to chance levels (1/12, as there are 12 semitones in an octave) for notes with an expectancy rating of '0'.

In subsequent trials, it is hypothesised that higher expectancy values will facilitate memory for notes. This follows the finding that schematically expected chords are processed faster than unrelated chords (Bharucha, 1987), and the results of Schmuckler (1997). Given Derek's ability, it is predicted that scores by trial 10 will begin to reach a ceiling effect, demonstrating an interaction between trial (time) and his proportional accuracy for notes of

different expectancy ratings. More expected notes will be committed to memory first; once these are learnt, Derek's cognitive resources will be able to focus on the less expected notes.

Method

Participant

Derek Paravicini is a 33 year-old male with a remarkable ability to play the piano and an extensive memory for music. He is congenitally blind due to retinopathy of prematurity. He has severe learning difficulties and diagnosis of autism, and in 2008 his verbal IQ was 58 (as reported by Pring, Woolf & Tadic, 2008; measured using WAIS-R). He is unable to explain how he learns and plays music.

Design

A single case study design was used in this exploration of musical expectation and memory. Derek was asked to play along with the stimulus piece, *Romantic Rollercoaster,* in 10 trials over a 2-month period. (In these trials, his memory for the piece was also tested, but analysis of these renditions is beyond the scope of this paper). Trials 1, 2, 5 and 10 only were considered in the analysis. (This was due to time constraints. It should be noted that the task undertaken by Derek in trial 1 was different in nature from subsequent trials, as this was the first time he had encountered the piece. Therefore, one set of adjacent trials [1 and 2] have been included in the analyses.)

Each note of the stimulus piece was assigned an expectancy rating, which was correlated with Derek's proportional accuracy for that note in each of the 4 analysed trials. An item analysis was also carried out. The original 6 possible expectancy ratings were grouped into 4 categories. 40 matched, 3-note segments (10 of each of the 4 expectancy categories) were selected (see 'Isolation of segments', p. 12). A 4x4 mixed ANOVA was performed, in which the between-groups variable was expectancy category (1-4), and the within-groups variable was trial number (1, 2, 5 and 10). The dependent variable was Derek's overall proportional accuracy for each segment. Notes were considered accurate only if they were within the imitation threshold (see below, pp. 15-18, for details of how proportional accuracy was calculated, and how the imitation threshold was determined).

Materials

Stimulus piece (*Romantic Rollercoaster)*. It was necessary to compose a new piece of music to ensure that Derek could not have previously heard the stimulus (Derek is unable to say reliably whether or not he recognises a piece; Ockelford, 2008, p. 226). Similar composition criteria to those used for other compositions in the *Fragments of Genius* series were followed, to allow comparison across the different learning methods. These are outlined in Appendix A(i).

The resulting stimulus, *Romantic Rollercoaster* (figure 2; see p. 13), consists of 362 events (notes) over a period of 86.7 seconds, an average rate of 4.2 events/second. Its musical structure is outlined in Appendix A(ii).

Isolation of segments. The nature of the task meant that each note in isolation could score only a restricted range of values (0, 0.25, 0.75 or 1). Therefore, an item analysis was carried out using groups of three notes ('segments') selected from *Romantic Rollercoaster.* Segments contain 3 notes that occur together or sequentially, but all three notes are of the same expectancy. This remit made it impossible to select the same number of segments of each expectancy rating. Therefore, the original 6 possible expectancy ratings (according to the model described above) were grouped into 4 categories, as shown in table 2:

Table 2

System for grouping the 6 expectancy ratings into 4 categories

Original expectancy rating	Expectancy category
0	1
.2	2
.4/.6	3
0.8/1	4

Derek's overall proportional accuracy for each segment was measured (see p. 15 below for details of the scoring system used). Proportional accuracy for these segments could then be compared across trials and across segments of different expectancy categories. In order to ensure that, aside from expectancy category, there were no systematic differences between segments (particularly with regard to complexity) they were matched using the criteria outlined in Appendix B. The resulting segments are marked in figure 2 (p. 13).

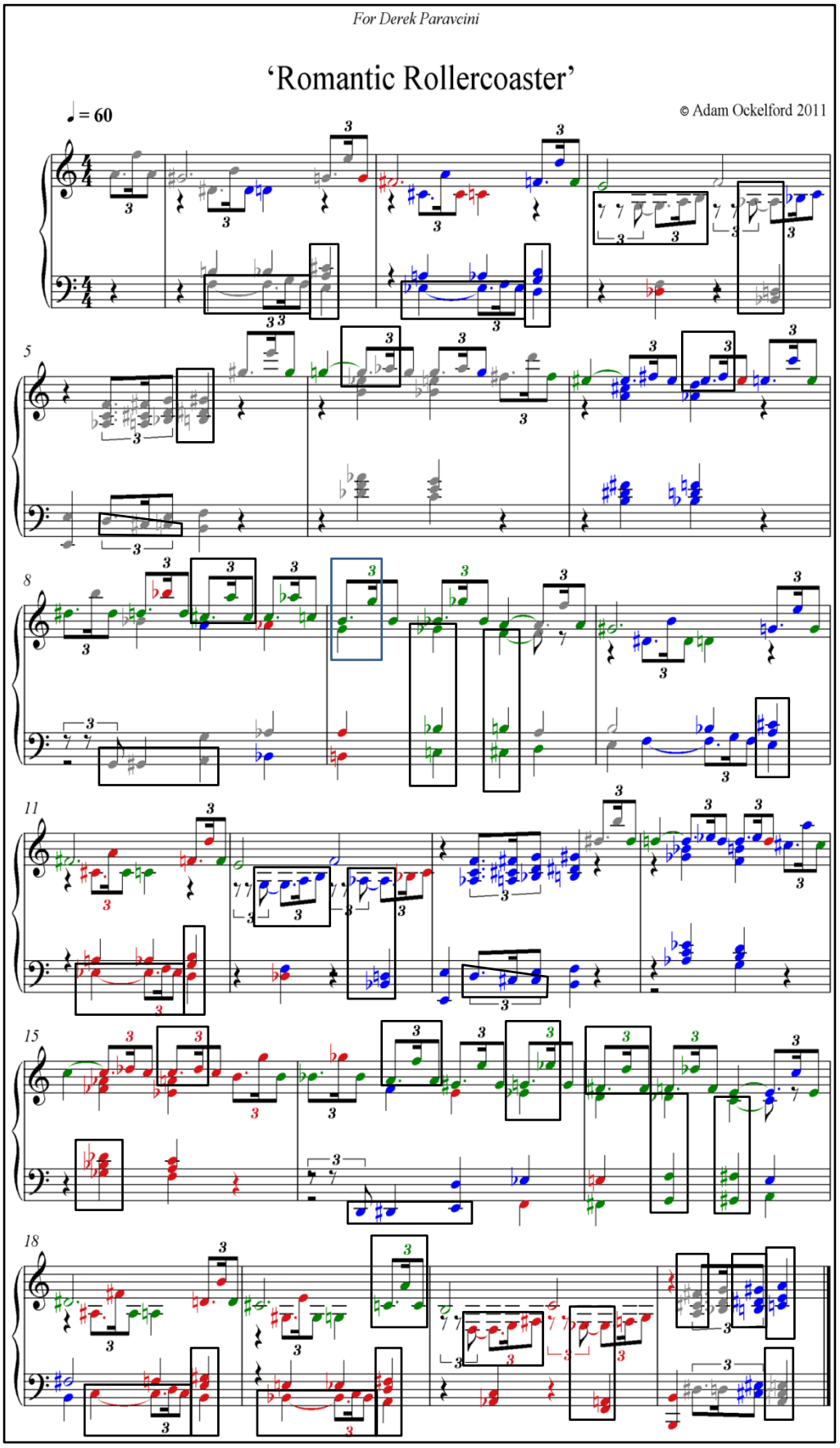


Figure 2: Score of the stimulus piece. Segments used in the item analysis are marked with boxes, note colours show expectancy categories (grey: 0; blue: 1; red: 3; green: 4).

Equipment

The stimulus piece was recorded at the Institute of Education, London, on a Yamaha CP300 Stage Piano using a M-Audio USB UNO MIDI Interface by REAPER Digital Audio Workstation (version 4/x64), onto a Toshiba Satellite R-830 laptop in Waveform Audio File format. This was converted into MIDI format.

The Waveform Audio File version of *Romantic Rollercoaster* was used in the experiment so that the stimulus sounded realistically like a piano performance of the piece. (The MIDI version was played precisely simultaneously but muted, as this was necessary for the analysis stage). The stimulus was played using REAPER on the laptop, via Edirol MA-10A speakers.

Derek played a Korg SP-200 88 note touch-sensitive hammer action keyboard, set to 'Grand piano 1' sound and amplified using a Roland 4-ch Mixing Keyboard Amplifier. Derek's playing was recorded in MIDI format by REAPER Digital Audio Workstation using the M-Audio MIDI Interface to the Toshiba laptop.

A video recording of Derek's hands on the keyboard was made using a Vivitar DVR 985HD camcorder.

The data gathered in MIDI format by REAPER was converted into csv (Comma Separated Values) format and analysed using Microsoft Excel (2010) and IBM SPSS Statistics, version 19.

Procedure

Ten sessions took place over a two-month period, with a break of a month between sessions 5 and 6, as shown in table 3 (p. 15):

Table 3

Timetable for trials

Session	'Play along' trial	'Memory' trial	Date of session
1	1	-	05/03
2	2	1	07/03
3	3	2	12/03
4	4	3	14/03
5	5	4	22/03
6	6	5	16/04
7	7	6	18/04
8	8	7	23/04
9	9	8	25/04
10	10	9	30/04

Procedure for each session. Each session took place in a quiet room at the University of Roehampton, London. In each of sessions 2-10 the protocol was as follows:

1) 'Memory' trial: Derek played what he could remember of *Romantic Rollercoaster*
2) 'Play along' trial: Derek played along with the recorded stimulus
3) Derek listened to the recorded stimulus

Session 1 consisted of steps 2) and 3) only. Derek's playing was recorded to the nearest 2000th of a second. He was thanked on each occasion but no comment was made as to how accurately he had performed the task. Derek had no exposure to the stimulus piece between trials. Trials 1, 2, 5 and 10 were selected for inclusion in the current analyses to represent data from across the experimental time period.

Scoring system. Proportional accuracy was measured using a method based on zygonic theory. The basic premise is shown in figure 3 (p. 16), where S = stimulus piece and D = Derek's attempt to 'play along'. The 'relationships' can be in any aspect of music; here, pitch was the primary aspect under investigation, but rhythmic data was also taken into account for those notes which scored correctly for pitch.

$$\text{Proportional accuracy} = \frac{\text{Total relationships between S and D}}{\text{Potential relationships between S and D}}$$

Figure 3: Basic formula for the calculation of the proportional accuracy of Derek's attempts to play along (D), as compared with the stimulus piece (S) (adapted from Ockelford, 2012).

Data was rounded to the nearest 10ms. Each note of the stimulus was considered separately. If Derek matched that note, it was scored as detailed in Appendix C.

In the item analysis, the proportional accuracies of each of the 3 notes in each segment were averaged to produce the proportional accuracy score for that segment.

Imitation threshold task

It was necessary to find out the fastest possible time (the 'imitation threshold') in which Derek could copy individual chords from *Romantic Rollercoaster,* isolated so that he was unable to predict from a musical context what he would hear next. In trial one, any notes which were played before this threshold must have been predicted. After trial one, any notes played within this threshold may have been predicted, or may have been remembered from previous hearings of the stimulus. (It is possible that notes played after the threshold may also have been at least partially anticipated or remembered, but this is impossible to test empirically). Therefore, only notes played within the threshold were used to assess the accuracy of the model. A conservative threshold was chosen to ensure that notes included in the analyses were highly unlikely to have been copied, rather than anticipated.

Materials for the imitation threshold task. 30 chords were selected from the stimulus piece and recorded live using a Yamaha CP300 Stage Piano onto REAPER. The chords were of different densities, ranging from 1 to 6 notes; 5 chords of each density were used. All notes within a given chord were quantised to ensure synchronised onset times, but minor variations in duration and velocity of notes were maintained to preserve ecological validity. The 30 chords were each used 5 times in the stimulus sequence. The order of the resulting 150 chords was randomised using an algorithm generated from *random.org* (Haahr, 1998), which offers a 'true random number service that generates randomness via atmospheric noise' (Haahr, 1998, "Introduction to Randomness and Random Numbers", para. 1). Post-hoc validation was carried out on the resulting sequence to check for immediate repetitions of chords. Two such instances were identified and corrected. Pauses

between chords were also varied randomly (minimum time between chord onsets: 2 seconds; maximum: 4 seconds).

During the experiment, the same equipment and computer programs were used as for sessions 1 – 10.

Procedure for imitation threshold task. This took place in a quiet room at the University of Roehampton. A 'listen and play' paradigm was used. Derek was told that he would hear a sequence of chords, and asked to play each of them on the piano as quickly as possible after hearing them. During the task, Derek was reminded gently to play the chords 'straight' (not arpeggiated). Derek was congratulated and thanked after the task.

Scoring in imitation threshold task. Data was rounded to the nearest 10ms. Proportional accuracy for the imitation threshold experiment was measured using a method based on zygonic theory, as above. Notes played correctly scored 1; notes of the correct pitch class but wrong octave scored 0.5. An overall proportional accuracy score for each chord was calculated as shown in figure 4:

$$\text{Proportional accuracy} = \frac{\text{Score for chord x}}{\text{Total number of pitches in chord x}}$$

Figure 4: Formula for the calculation of Derek's proportional accuracy for each chord in the imitation threshold task.

Derek's reaction time for each chord was calculated by averaging the onset times of each component note (of the chord played by Derek), and subtracting from this the onset time of the stimulus chord.

Results of the imitation threshold experiment. The first 5 chords of the sequence were excluded from the analysis to allow Derek to become familiar with the task (i.e. the first five chords were treated as a practice task). For the remaining 145 chords, Derek's mean proportional accuracy was 0.972 (SD: 0.062). His average latency was 0.76s (SD: 0.17s, range: 0.49-1.47s).

These results demonstrate Derek's ability to reproduce complex aural stimuli quickly and accurately. In light of this, it was necessary to apply a conservative imitation threshold to the trial data. This was to ensure that none of the notes included in the analysis could have been copied from the recording of *Romantic Rollercoaster,* rather than anticipated or, after

trial 1, remembered. Therefore, the imitation threshold was set at 0.4s (as this is below Derek's fastest time in the above described experiment).

Results

Analysis of whole dataset

Figure 5 shows Derek's overall proportional accuracy for all notes in each of trials 1, 2, 5 and 10, before and after application of the imitation threshold:

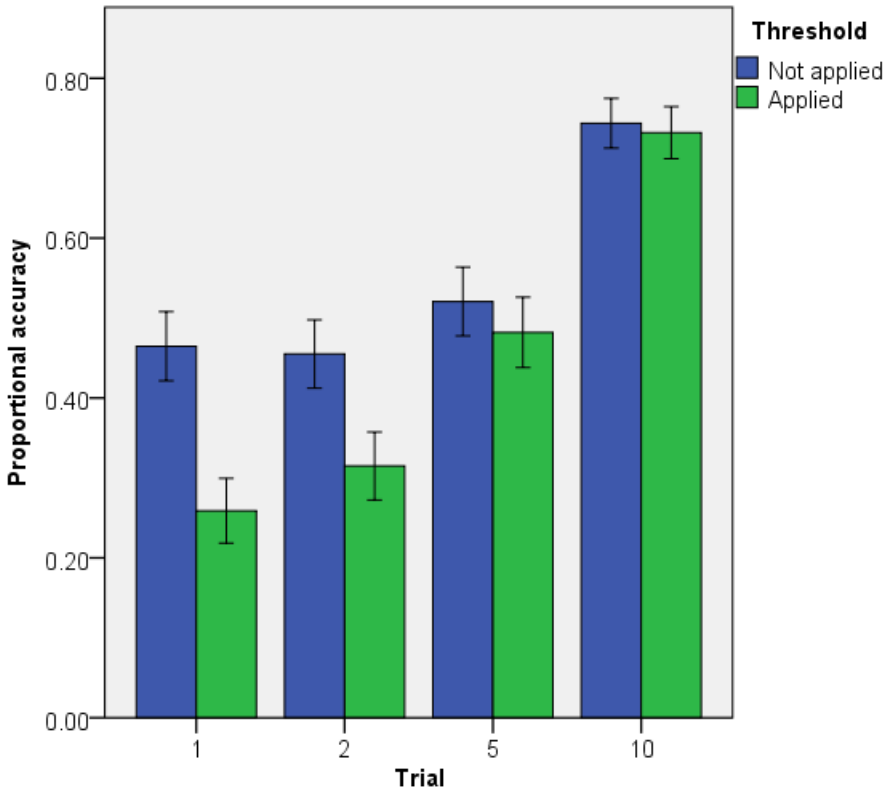


Figure 5: Derek's overall proportional accuracy for all notes in each trial, without and with the application of the imitation threshold. Error bars represent 95% CIs.

There were significant differences between proportional accuracy score before and after the application of the imitation threshold in trial 1 ($t(719) = 6.84$, $p < .001$) and 2 ($t(722) = 4.57$, $p < .001$), but not in trial 5 ($t(722) = 1.24$, $p = .109$) or 10 ($t(720) = .51$, $p = .304$) (all significance levels one-tailed).

Following the application of the imitation threshold, notes were analysed separately according to expectancy rating. When the whole dataset was included, Derek's average

proportional accuracies for notes of each expectancy rating across trials 1, 2, 5 and 10 were as shown in figure 6:

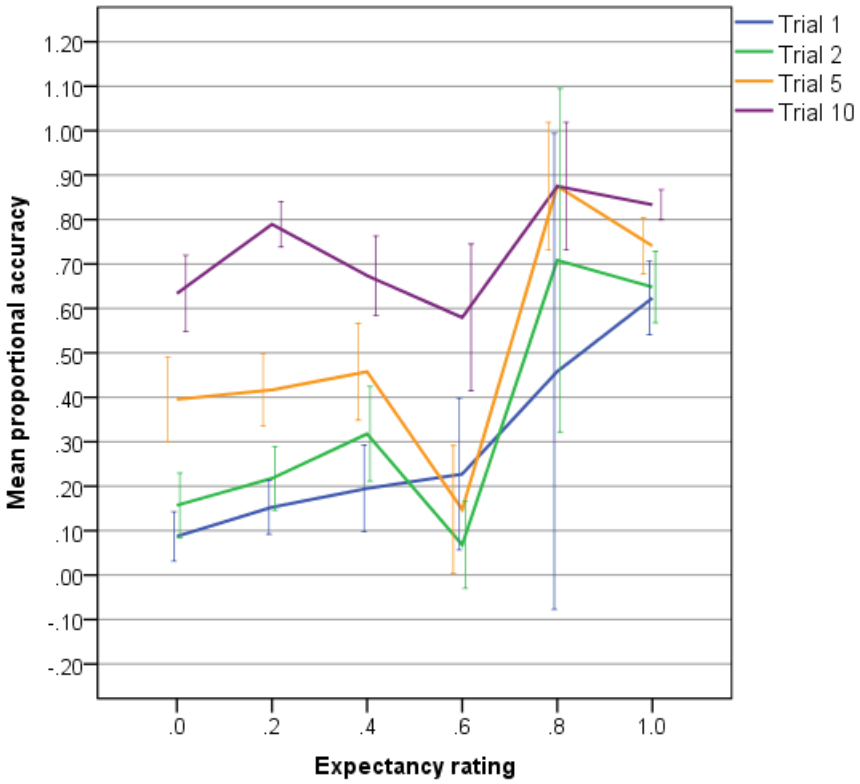


Figure 6: Derek's mean proportional accuracy for notes of each expectancy rating across trials 1, 2, 5 and 10. Error bars represent 95% CIs. Only notes played within the imitation threshold are included.

Spearman's correlations showed that the relationship between expectancy rating and proportional accuracy decreased as trial number increased, as shown in table 4 (p. 20).[2] (It should be noted that these correlations may be unreliable as the scale used for proportional accuracy contained few gradations. This is why an item analysis was also carried out [see below, p. 20]).

[2] A non-parametric measure of correlation was used because the data was not normally distributed, and equivalent parametric tests yielded different overall results from those shown here.

Table 4

Spearman Rho Correlations between proportional accuracy and expectancy rating of notes, across trials 1, 2, 5 and 10

	r_s
Trial	
1	.455*
2	.383*
5	.217*
10	.064**

*p < .001 **n.s. (p > .2)

Item analysis

Derek's average proportional accuracy scores for the selected segments (10 of each expectancy category), across trials 1, 2, 5 and 10, are shown in figure 7:

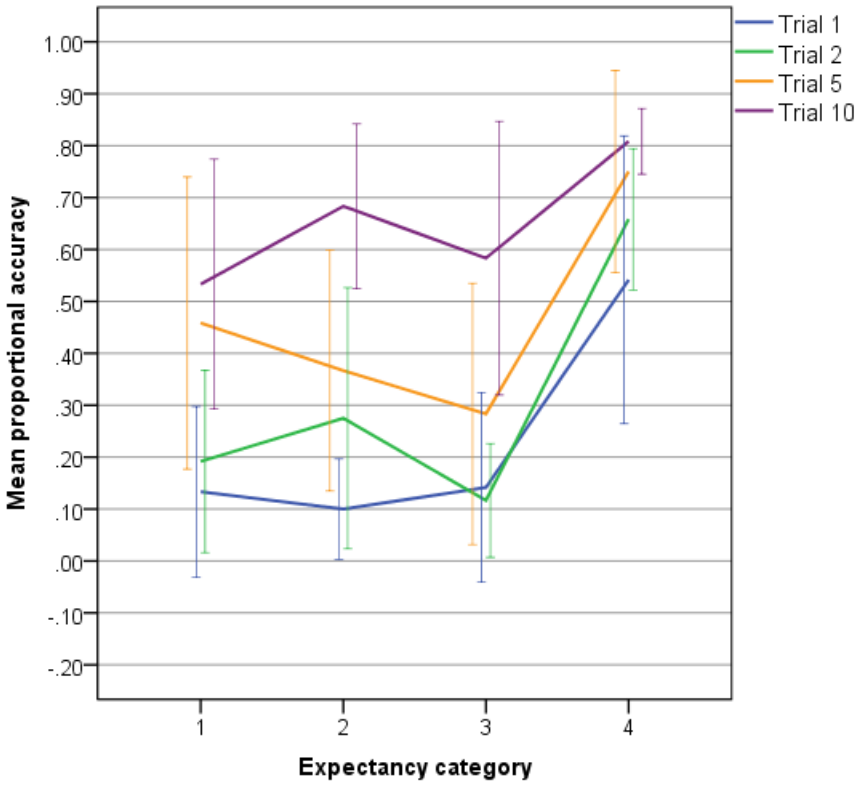


Figure 7: Derek's mean proportional accuracy score for the segments of each expectancy category, across trials 1, 2, 5 and 10. Error bars represent 95% CIs.

Mauchley's test indicated that assumptions of sphericity had not been violated ($\chi^2(5) = 1.245$, $p = .941$).[3] A 4x4 mixed ANOVA showed a significant effect of expectancy category ($F(3, 36) = 9.063$, $p < .001$). Post-hoc tests (with Bonferroni correction applied) showed that Derek was significantly more accurate for segments in expectancy category 4 as compared to category 1 ($p = .001$), category 2 ($p = .003$) and category 3 ($p < .001$), but there were no significant differences between categories 1, 2 and 3. There was a significant effect of trial ($F(3, 108) = 24.496$, $p < .001$). Post-hoc pairwise comparisons (adjusted using Bonferroni) showed significant differences between all pairs of trials apart from between trials 1 and 2 (values shown in Appendix D). The interaction which appears in figure 7 did not come close to significance ($F(9, 108) = 1.115$, $p = .359$), due to the high variability of results. (All significance levels two-tailed)

However, separate univariate ANOVAs for each trial (with alpha level, adjusted for multiple comparisons using Bonferroni, set at .0125) showed that the effect of expectancy category remained significant in trials 1 ($F(3, 36) = 6.117$, $p = .002$) and 2 ($F(3, 36) = 9.527$, $p = <.001$), but not in trial 5 ($F(3, 36) = 3.610$, $p = .022$) or 10 ($F(3, 36) = 1.929$, $p = .142$). (All significance levels two-tailed)

Discussion

This study investigated the use of the 'play along' method of learning music by an individual with savant syndrome (Derek Paravicini). Derek attempted to learn an especially composed piece of music across 10 trials over a 2-month period. The study addressed the importance of musical expectancy for this kind of learning, and the accuracy of a model of expectation in music, adapted from Ockelford (2006). The paradigm allowed for a re-consideration of the distinction between schematic and veridical expectations (Bharucha, 1987). In the present paper, only trials in which Derek played along with the stimulus piece were considered. Two hypotheses were put forward:

1. Derek's proportional accuracy for notes in the 'play along' task would correlate with the expectancy values of those notes (as predicted by the model) in trial 1.

[3] Although assumptions of normality and homogeneity of variance were violated in some cases (and this could not be rectified using data transformations), parametric tests were conducted as they yielded the same overall results the non-parametric alternatives, except in the case of a Jonckheere-Terpstra test of expectancy category in trial one, which was close to significance ($p = .0127$, $\alpha = .0125$). (See Appendix D)

2. Higher expectancy values would facilitate memory for notes in subsequent trials, and an interaction between proportional accuracy and expectancy would be found as accuracy reached a ceiling.

Significant differences between Derek's proportional accuracy scores before and after the application of the imitation threshold (set at 0.4 seconds) were present in trials 1 and 2, but not in trials 5 and 10. This showed that a significant proportion of accurate notes in the earlier trials were echoed (but this was not the case in later trials). After the application of the imitation threshold to the data, there was a strong correlation between expectancy rating and proportional accuracy in trial 1, which declined as trial number increased. The potential unreliability of these correlations is acknowledged, which is why an item analysis was carried out using segments of *Romantic Rollercoaster.* This analysis showed that Derek was significantly more accurate both for later trials and for notes in a higher expectancy category. Further analysis found this difference to be between expectancy category 4 and all others. Separate analyses for each trial showed that the significant effect of expectancy was no longer present in trial 10. Thus, expectancy category was not an important factor in Derek's ability to remember (or predict) notes accurately in trial 10. This, taken with the correlations described above, suggests that an interaction effect might be present with a larger sample size which contains less variability. However, this conclusion can be made only speculatively as the item analysis did not show a significant interaction effect.

Hypothesis 1 is supported by the strong correlation between Derek's proportional accuracy and each note's expectancy rating in trial 1, suggesting Ockelford's model to be a good predictor of accuracy for the first encounter with a piece. In this trial, Derek scored near to chance levels for notes with expectancy rating of 0, which further evidences the importance of between-groups relationships for correct expectations. Hypothesis 1 is also supported to some extent by the item analysis for trial 1, although significant differences occurred only between expectancy category 4, and all other categories. This could suggest that the model contains too many distinctions, and that there are real differences only when a group has occurred 4 or more times already in the piece. However, the apparent lack of a difference in proportional accuracy between notes of expectancy categories 1, 2 and 3 could also have been due to a relatively small sample size (10 segments in each group) or the necessity to collapse the original 6 expectancy ratings into 4 categories (so any distinction between expectancy ratings 0.4 and 0.6, and 0.8 and 1, was lost). In light of this, and the strong correlation found in the analysis of the whole dataset, it is likely that the expectancy ratings

generated by the model provide an accurate measure of Derek's ability to predict those notes, and therefore can be used to assess the impact of expectancy on accuracy and memory in subsequent trials.

The overall increase in proportional accuracy across trials (after the application of the imitation threshold; see figure 5, p. 18) confirms that at least some of the accurate notes in trials 2, 5 and 10 were a result of long term memory for the piece. On the other hand, the memory trials (to be described elsewhere) showed that, even by trial 10, Derek did not learn the piece in its entirety. Therefore, it is likely that overall proportional accuracy scores for trials 2, 5 and 10 were the result of a combination of long term memory and between-groups expectations, and so can be used to explore the relationship between the two.

Hypothesis 2 related to trials 2, 5 and 10. The data suggests that expectancy does facilitate memory to some extent, with the most notable difference occurring for the highest expectancy category. Support is also found for the second part of hypothesis 2; a ceiling effect in later trials. Although there was no significant interaction between expectancy category and proportional accuracy when the trials were analysed together, the results of separate analyses suggest that the ascent of memory renders expectancy values less important.

The accuracy and speed with which Derek echoed the chords in the imitation threshold experiment are in line with previous findings from similar studies involving Derek (Ockelford, 2008; Pring et al., 2008) and highlight the striking acuity of his echoic memory and his ability to reproduce what he hears using the keyboard.[4] This aural aptitude was also demonstrated in trial 1, where Derek correctly played (by predicting or copying) 46% of all notes (see figure 5, p. 18). This skill is a necessary prerequisite to accurate prediction in using the 'play along' method for a piece as complex as *Romantic Rollercoaster*, so also explains why Derek was able to play some notes (26%) before the imitation threshold in trial 1, his first encounter with the piece. It should also be noted that the lack of significant differences between proportional accuracy before and after the application of the imitation threshold in trials 5 and 10 show that Derek was minimally, if at all, using his ability to copy aural stimuli in these trials (instead, notes were correct because of accurate prediction or memory).

[4] These previous studies included comparison subjects (professional musicians) and are evidence that Derek's ability is exceptional and beyond the capabilities of highly trained musicians.

Given Derek's demonstrated and unusual ability, it is reasonable to query the legitimacy of generalizing the findings of the present study; in particular whether the paradigm used is an accurate test of Ockelford's model (Ockelford, 2006). That Derek's musical abilities are far beyond those of most other people is undisputed; a more pressing question is whether or not he processes music in a fundamentally different way. Although impossible to separate out the influences of different aspects of Derek's personality, ability and diagnosis on his responses, some tentative suggestions are required relating to visual impairment, and savant syndrome (to which autism is inextricably linked), when addressing this question.

Musicians with visual impairment rely more (than most sighted musicians) or entirely (as in Derek's case) on learning music by ear. Visual impairment has been shown to facilitate certain aspects of sound processing, such as detecting differences in durations of sounds (Van der Lubbe, Van Mierlo, & Postma, 2010), detecting differences in vowel sounds and emotions in sound (Klinge, Roeder, & Buechel, 2010) and perception of echo cues (Dufour, Despres, & Candas, 2005). Differences in brain structure (e.g. Elbert et al., 2002) and processing networks (e.g. Collignon, Davare, Olivier, & De Volder, 2009; Lewis et al., 2011) have also been demonstrated. Therefore, it could be that people with visual impairment utilise different brain networks (from sighted individuals) in generating musical expectations. However, this is unlikely given findings that it is the early stages of sound processing which are different (and enhanced) in blind individuals (e.g. Rokem & Ahissar, 2009; Stevens & Weaver, 2005, 2009). The method of generating musical expectations suggested in the present paper is rather the result of higher-order mental processes, implying that it could be applicable to the wider population.

A similar argument can be made in the case of savant syndrome (and autism). The development of savant abilities is believed to be the result of autistic traits of 'weak central coherence' (Frith & Happe, 1994) and 'enhanced perceptual functioning' (Mottron, Peretz, & Menard, 2000) (see Pring, 2008). Also in line with these two theories, autism has been shown to provide an advantage in processing 'local' aspects of music, such as pitch identity (Heaton, 2003) and subtle changes in pitch (Bonnel et al., 2003; Heaton, 2005). In themselves, these abilities should not affect the higher-order processes of musical expectation in people with savant syndrome and autism, unless they preclude an awareness of musical context. Although it has been argued that autism does result in reduced global processing of music (Foxton et al., 2003), there is compelling evidence to the contrary (Heaton, Williams, Cummins, & Happe, 2007; Quintin, Bhatara, Poissant, Fombonne, &

Levitin, 2012). Therefore it can be suggested that the way in which Derek generates his musical expectations is not specific to those with autism or savant syndrome.[5]

Previous work has highlighted the similarities in music processing between Derek and a professional musician, and a young child singing a simple song (Ockelford, 2008). It is on this basis of shared music processing networks that the *Fragments of Genius* series is argued to be relevant for music psychology in general, and not just those with savant syndrome. The present study has contributed further justification of this approach. A model of expectancy was used which was based on previous empirical work with participants whose formal musical training ranged from none, to over 8 years (Thorpe et al., 2012). Ratings generated by this model were found to correlate well with Derek's first encounter with *Romantic Rollercoaster*. Not only does this help to justify use of the model in an investigation of expectancy and memory in the later trials, but it demonstrates the similarities and comparability between Derek's cognition of music and that of Thorpe et al.'s participants, suggesting that fundamental methods of music processing are similar across a variety of stimulus complexities and participant abilities. The implications of the findings of the present study will therefore be discussed with reference to several areas: expectancy and memory, Ockelford's model and general research into musical expectancy, as well as the 'listen and play' method used. Suggestions for further work will also be given.

Expectancy and memory

The present study supports the finding that more expected musical events are remembered better (Schmuckler, 1997) and extends this beyond the standard tonal framework to which Schmuckler attributes this effect. This evidences the expectancy of the notes themselves as important for memory, rather than the schema to which they belong. The present experiment has also extended this work by introducing more subtle gradations of expectancy (6, as opposed to 3 in Schmuckler, 1997). These were determined on the basis of empirical testing of a model (which has been shown in trial 1 to be an accurate measure of Derek's ability to predict), rather than participant ratings (Schmuckler, 1997).

[5] Future work is planned in which a sighted, neurotypical concert pianist will attempt the same task asked of Derek in trial 1 (using the same stimulus piece). As well as allowing comparison of proportional accuracy scores between the two individuals, analysis of the concert pianist's results will show whether or not they use similar processes of generating musical expectations as does Derek. This will provide further evidence as to whether or not the findings discussed here are generalizable (albeit to an individual with a high level of musical training).

Figures 6 (p. 19) and 7 (p. 20) show a dip in proportional accuracy for notes of expectancy rating 0.6 and expectancy category 3, respectively. Although differences were not significant, the consistency of this decline warrants investigation, particularly as it did not appear in the first trial and so cannot be explained as low expectancy for those notes. These results could be taken as evidence in support of Schmuckler's finding that low-expectancy notes were remembered slightly better than medium-expectancy notes, which were remembered least well (Schmuckler, 1997). This is in line with the 'von Restorff effect' that highly unexpected notes have greater salience and so are fixed in memory more strongly (Schmuckler, 1997). However, in all the trials discussed in the present paper, notes of the second expectancy category scored higher than those in the first. The decline for rating 0.6 (/category 3) could instead be the result of a relatively small 'sample' size; only 22 notes in *Romantic Rollercoaster* have the rating 0.6. It could be that these 22 notes were difficult to remember for other, musical, reasons. Clarification should be sought using pieces composed to include a more equal number of notes from each expectancy category.

These results could also have some implications for our understanding of working memory for music. Several authors have suggested that Baddeley and Hitch's 'Working Memory Model' (Baddeley, 1990) may require some alteration to incorporate a separate store or rehearsal process for musical stimuli (e.g. Ockelford, 2007b; Pechmann & Mohr, 1992; Williamson, Baddeley, & Hitch, 2010). The present study has relevance for this debate because any kind of musical expectations must involve working memory. Expectations require decisions about what is likely to come next, made on the basis of memory for what has just occurred in the music (these decisions are made with relation to previous listening experiences, whether intrinsic or extrinsic to the piece in question). Derek's ability to accurately predict at least 26% of notes in his first encounter with *Romantic Rollercoaster* (a harmonically complex piece) evidences his impressive working memory skills within this domain (see also reports of the 'listen and play' study, e.g. Ockelford, 2011).[6] Given the context of Derek's general disabilities, the findings presented above provide evidence in favour of some kind of separation for musical (as opposed to linguistic) stimuli in working memory (although it would be necessary to measure Derek's verbal working memory ability to confirm this suggestion).

[6] However, a study involving a comparison subject(s) is required to confirm the uniqueness of Derek's ability.

Implications for the model (of expectation in music)

The results presented here extend the findings of Thorpe et al. (2012) to show the importance of between-groups predictions for accurate expectation in music, within a full musical texture and a learning paradigm. Ockelford's model is intended to be relevant for both first and subsequent encounters with a piece of music. The strong correlation seen in trial 1 suggests that it is accurate for the first encounter. In later trials, its accuracy is less clear. Weaker correlations for trials 2 and 5 and significant differences (in the item analysis) for expectancy category 4 only, suggest that between-groups predictions are of importance for memory only where the musical group in question has appeared four or more times in the past. This would imply that the emphasis placed by the model on between-groups expectations should change according to familiarity with a piece, such that the finer-grained distinctions between 0, 1, 2 and 3 previous occurrences of a group are no longer taken into account past the first encounter with a piece of music. This suggestion could be confirmed or refuted by further empirical work involving a bigger dataset.

There should also be an assessment of inaccurate notes played by Derek. In some instances, the notes predicted by the model are not those which actually occur in *Romantic Rollercoaster* (when the adjacency and recency of the predictor groups are taken into account). Derek's 'wrong' notes in trial 1 (which demonstrate his musical expectations) should be compared with the model's output. Alongside this, understanding of the importance of between-groups expectations could be further refined by assessing any differences between groups repeated at the same absolute pitch and those which are not.

Implications for music and expectancy research

The compatibility between the results seen here and the predictions of a model based on the frequency with which a group of notes has occurred in the past suggests that veridical expectations are generated via the same probabilistic learning mechanism as are musical schemata. Events are predicted on the basis of how often they have occurred in the past, given a particular context (Huron, 2006; Pearce & Wiggins, 2006; Saffran et al., 1999). This affirms Huron's suggestion that the distinction between the two types of memory may have been made too rigidly (Huron, 2006, p. 225).

Ockelford (2006) suggests that the between-groups expectations found here to be so important for prediction and for memory are generated mainly by schematic expectations at first encounter with a piece. Veridically-encoded expectations are argued to become

more dominant as familiarity increases. However, the use of *Romantic Rollercoaster* meant that possible reliance on schematic expectations in the first instance was minimal, as the piece does not conform to standard Western patterns of harmony (though it should be acknowledged that its relatively simple overall structure may have led to schematic, form-based expectations). Therefore, Derek may have approached it as a new harmonic schema to be learned (this process may or may not have been conscious). It has been found that listeners are capable of putting off expectations pertaining to a familiar schema in order to approach a new one (Castellano, Krumhansl, & Bharucha, 1984). Listeners are perceptive to the probabilistic information of a new schema (Loui et al., 2009; Saffran et al., 1999), so Derek may have encoded and reproduced likely continuations from any given moment of the piece. However, the nature of the task (requiring Derek to play *Romantic Rollercoaster* as accurately as possible, rather than an improvisation in a similar style) means that it was his veridical memory for the piece which was assessed. Furthermore, his accuracy across the trials suggests that *Romantic Rollercoaster* was unlikely to have been approached as a schema, justifying the above conclusion that parallels between the encoding of schematic and veridical expectations can be drawn.[7]

Comparison with previous *Fragments of Genius* study

Although a full analysis is beyond the scope of this paper, an initial examination of the memory trials showed that Derek did not fully learn the piece by the end of the study (sections of the music were remembered correctly, but the overall structure of *Romantic Rollercoaster* was not). The nature of the 'play along' task meant that it was necessary for a slightly different system of scoring to be used from the *Chromatic Blues* study (in which a 'listen and play' paradigm was used).[8] Therefore, data from the two experiments cannot be treated as precisely equivalent, but can be compared approximately as both give scores of the proportion of Derek's playing which was derived from the original composition (calculated using zygonic theory). Trials in the first stage of the 'listen and play' study were carried out over the same time frame as the present study (although several years earlier). At the equivalent trial in the *Chromatic Blues* study, Derek scored a proportional accuracy of .76 (Ockelford, 2011), compared to .73 within the imitation threshold in trial 10 of the present study. Unlike 'listen and play', the task of playing along with the stimulus meant

[7] It would also be of interest to see what Derek would play if he was asked to improvise in the style of *Romantic Rollercoaster*. This could help us to understand the extent to which its harmonic patterns had become encoded schematically.

[8] The same system as used for *Chromatic Blues* can be applied to the *Romantic Rollercoaster* memory trials, however.

that some of these notes are likely to have been anticipated rather than remembered. If 'play along' and 'listen and play' were equally effective learning methods, the *Romantic Rollercoaster* trials (in which Derek played along with the stimulus) should have yielded higher overall scores. That Derek scored lower in the final *Romantic Rollercoaster* 'play along' trial than the equivalent *Chromatic Blues* trial suggests that 'play along' is a less efficient learning method than 'listen and play'. A more thorough analysis of the 'memory' *Romantic Rollercoaster* trials will allow for confirmation of this finding, from which implications for pedagogy can emerge. This initial finding also supports the drawing of parallels between 'play along' and simultaneous interpretation, as it has been found that the latter process also disrupts subsequent memory for the stimulus material (Daro & Fabbro, 1994; Isham & Lane, 1993).

Further implications for future research

This work has provided insights into the mental processes of someone who chooses to learn by playing along with a piece of music, and has highlighted the importance of musical expectations generated from the relationships between musical groups. The study as reported here could not be replicated using many other participants for pieces of such a complex nature, as Derek is highly unusual in his ability to perform this task. However, similar analytical techniques could now be applied to the other pieces in the *Fragments of Genius* series, to assess whether note predictability has the same impact on memory when Derek uses different learning methods.

The paradigms used across the series provide a more suitable test of the accuracy of expectations, or memory, than studies where production of the music is not required (or the production is less complex). This is because notes are not scored unless they are an accurate reproduction of the stimulus, and so do not reflect a reliance on 'generalized memory' (see McAuley, Stevens, & Humphreys, 2004), which could be what is tested in simpler paradigms. Thorpe et al. (2012) provides an example of one such simpler method, which was more artificial and was not a learning method. However, their results are supported strongly by those reported in this paper, which justifies their method as an accurate assessment of Ockelford's model. Therefore, future work could use variations on their method, with a wider range of participants. The accuracy of the model should also be assessed across different modes of engagement with music, including listening, and learning using different methods.

This study has highlighted similarities between schematically and veridically generated expectations. Future research could therefore examine further how 'episodic' memories (leading to veridical expectations) are obtained. In particular, there should be investigation of whether the use of probabilistic learning remains an important factor when the piece in question conforms to a familiar harmonic schema. Further work investigating the impact of expectedness on memory is also required, using greater sample sizes (number of notes), in order to allow for further refinement of Ockelford's model. In addition, a more complex scoring system could give a greater range of potential scores. A wider range of musical aspects could be considered, such as dynamics and articulation. This would allow for the retention of more information about participant responses.

Conclusion

This study has demonstrated something of the exceptional talent of Derek Paravicini, in particular his highly accurate echoic memory, and his ability to predict musical events based on factors intrinsic to the piece in question. It aimed to extend our understanding of the latter process and to test the accuracy of a model of musical expectation based on Ockelford's zygonic theory (Ockelford, 2006). Between-groups expectations (a central element of the model) were found to be important for Derek's first encounter with a piece, as well as for subsequent memory for it (although an item analysis showed a distinction only between the most expected notes and all others). Overall, this research highlights parallels between schematic and veridically-encoded expectations, and the importance of expectation for learning using the 'play along' method.

References

Aarden, B. (2003). *Dynamic melodic expectancy*. (Unpublished doctoral dissertation). Ohio State University, Columbus.

Baddeley, A. (1990). *Human memory: Theory and practice*. Boston: Allyn and Bacon.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: Evidence from task switching. *Journal of Experimental Psychology-General, 130*(4), 641-657.

Banks, W. P., Roberts, D., & Ciranni, M. (1995). Negative priming in auditory attention. *Journal of Experimental Psychology-Human Perception and Performance, 21*(6), 1354-1361.

Bharucha, J. J. (1987). Music cognition and perceptual facilitation – a connectionist framework. *Music Perception, 5*(1), 1-30.

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience, 15*(2), 226-235.

Byrne, C. (2005). Pedagogical communication in the music classroom. In D. Miell, R. MacDonald & D. Hargreaves (Eds.), *Musical Communication* (pp. 301-319). Oxford: Oxford University Press.

Castellano, M. A., Krumhansl, C. L., & Bharucha, J. J. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology-General, 113*(3), 394-412.

Collignon, O., Davare, M., Olivier, E., & De Volder, A. G. (2009). Reorganisation of the Right Occipito-Parietal Stream for Auditory Spatial Processing in Early Blind Humans. A Transcranial Magnetic Stimulation Study. *Brain Topography, 21*(3-4), 232-40.

Daro, V., & Fabbro, F. (1994). Verbal memory during simultaneous interpretation – effects of phonological interference. *Applied Linguistics, 15*(4), 365-381.

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics, 50*(4), 305-313.

Dufour, A., Despres, O., & Candas, V. (2005). Enhanced sensitivity to echo cues in blind subjects. *Experimental Brain Research, 165*(4), 515–519.

Elbert, T., Sterr, A., Rockstroh, B., Pantev, C., Muller, M. M., & Taub, E. (2002). Expansion of the tonotopic area in the auditory cortex of the blind. *Journal of Neuroscience, 22*(22), 9941-9944.

Eschrich, S., Muente, T. F., & Altenmueller, E. O. (2008). Unforgettable film music: The role of emotion in episodic long-term memory for music. *Bmc Neuroscience, 9*, 48.

Foxton, J. M., Stewart, M. E., Barnard, L., Rodgers, J., Young, A. H., O'Brien, G., & Griffiths, T. D. (2003). Absence of auditory 'global interference' in autism. *Brain, 126*, 2703-2709.

Frith, U., & Happe, F. (1994). Autism – beyond theory of mind. *Cognition, 50*(1-3), 115-132.

Haahr, M. (© 1998-2012). *Random.org.* Retrieved 26/06/2012, from www.random.org

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 44*(4), 543-551.

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders, 35*(6), 787-793.

Heaton, P., Williams, K., Cummins, O., & Happe, F. G. E. (2007). Beyond perception: Musical representation and on-line processing in autism. *Journal of Autism and Developmental Disorders, 37*(7), 1355-1360.

Hermelin, B., Oconnor, N., & Lee, S. (1987). Musical inventiveness of 5 idiots-savants. *Psychological Medicine, 17*(3), 685-694.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Isham, W. P., & Lane, H. (1993). Simultaneous interpretation and the recall of source-language sentences. *Language and Cognitive Processes, 8*(3), 241-264.

Kaernbach, C. (2004). The memory of noise. *Experimental Psychology, 51*(4), 240-248.

Klinge, C., Roeder, B., & Buechel, C. (2010). Increased amygdala activation to emotional auditory stimuli in the blind. *Brain, 133,* 1729-1736.

Krumhansl, C. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Larson, S. (2004). Musical forces and melodic expectations: Comparing computer models and experimental results. *Music Perception, 21*(4), 457-498.

Lewis, J. W., Frum, C., Brefczynski-Lewis, J. A., Talkington, W. J., Walker, N. A., Rapuano, K. M., & Kovach, A. L. (2011). Cortical Network Differences in the Sighted Versus Early Blind for Recognition of Human-Produced Action Sounds. *Human Brain Mapping, 32*(12), 2241-2255.

Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009). A Generalized Mechanism for Perception of Pitch Patterns. *Journal of Neuroscience, 29*(2), 454-459.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception, 22*(4), 663-713.

McAuley, J. D., Stevens, C., & Humphreys, M. S. (2004). Play it again: did this melody occur more frequently or was it heard more recently? The role of stimulus familiarity in episodic recognition of music. *Acta Psychologica, 116*(1), 93-108.

Meyer, L. B. (1956). *Emotion and Meaning in Music*. Chicago: University of Chicago Press.

Mottron, L., Peretz, I., & Menard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41*(8), 1057-1065.

Ockelford, A. (2006). Implication and expectation in music: a zygonic model. *Psychology of Music, 34*(81), 81-142.

Ockelford, A. (2007a). *In the Key of Genius: The Extraordinary Life of Derek Paravicini*. London: Hutchinson.

Ockelford, A. (2007b). A music module in working memory? Evidence from the performance of a prodigious musical savant. *Musicae Scientiae*, 5-36.

Ockelford, A. (2008). *Music for children and young people with complex needs*. Oxford: OUP.

Ockelford, A. (2011). Imagination feeds memory: Exploring evidence from a musical savant using zygonic theory. In D. Hargreaves, D. Miell & R. MacDonald (Eds.), *Musical Imaginations: Multidisciplinary Perspectives on Creativity, Performance and Perception* (pp. 31-61). Oxford: OUP.

Ockelford, A. (2012). *Applied Musicology: Using Zygonic Theory to Inform Music Education, Therapy and Psychology Research*: Oxford University Press.

Ockelford, A., & Pring, L. (2005). Learning and creativity in a prodigious musical savant. *International Congress Series, 1282*, 903-907.

Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *Neuroimage, 50*(1), 302-313.

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception, 23*(5), 377-405.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch – implications for a working-memory model. *Memory & Cognition, 20*(3), 314-320.

Pring, L. (2008). Memory characteristics in individuals with savant skills. In J. Boucher & D. Bowler (Eds.), *Memory in Autism* (pp. 210-230). New York: Cambridge University Press.

Pring, L., Woolf, K., & Tadic, V. (2008). Melody and pitch processing in five musical savants with congenital blindness. *Perception, 37*, 290-307.

Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. (2012). Processing of musical structure by high-functioningadolescents with autism spectrum disorders. *Child Neuropsychology, iFirst*, 1-26.

Rokem, A., & Ahissar, M. (2009). Interactions of cognitive and auditory abilities in congenitally blind individuals. *Neuropsychologia, 47*(3), 843-848.

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.

Schmuckler, M. A. (1997). Expectancy effects in memory for melodies. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 51*(4), 292-306.

Sloboda, J. (1985). An exceptional musical memory. *Music Perception, 3*(2), 155-170.

Spence, C., & Read, L. (2003). Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *Psychological Science, 14*(3), 251-256.

Stevens, A. A., & Weaver, K. (2005). Auditory perceptual consolidation in early-onset blindness. *Neuropsychologia, 43*(13), 1901–1910.

Stevens, A. A., & Weaver, K. (2009). Functional characteristics of auditory cortex in the blind. *Behavioural Brain Research, 196*(1), 134–138.

Thorpe, M., Ockelford, A., & Aksentijevic, A. (2012). An empirical exploration of the zygonic model of expectation in music. *Psychology of Music, 40*(4), 429-470.

Van der Lubbe, R. H. J., Van Mierlo, C. M., & Postma, A. (2010). The Involvement of Occipital Cortex in the Early Blind in Auditory and Tactile Duration Discrimination Tasks. *Journal of Cognitive Neuroscience, 22*(7), 1541-1556.

Von Restorff, H. (1933). About the effect of domain formation in the track field. *Psychologische Forschung, 18*(1), 299–342 .

Williamon, A., & Egner, T. (2004). Memory structures for encoding and retrieving a piece of music: an ERP investigation. *Brain research. Cognitive brain research, 22*(1), 36-44.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38*(2), 163-175.

Young, R. L., & Nettelbeck, T. (1995). The abilities of a musical savant and his family. *Journal of Autism and Developmental Disorders, 25*(3), 231-248.

Appendix A

Appendix A(i)

Criteria used in the composition of Romantic Rollercoaster (adapted from Ockelford, 2011, p. 32; *additions in italics*):

a) The style should be broadly familiar to Derek

b) *The stimulus should be sufficiently harmonically complex and unstable that Derek is unable to rely on expectations based on a tonal schema to predict what is coming next.*

c) The structure should be relatively easy to discern, including both the repetition and transformation of key materials, thereby enabling the researchers to gauge the impact of these relationships on the learning process. *The structure should be broadly similar to that of the other pieces in the* Fragments of Genius *series. This will allow for comparison of equivalent musical sections across the pieces in future work. (Comparison of structurally equivalent points, across different pieces, can be justified by the finding that musicians typically remember structurally important points more accurately than structurally unimportant points; Williamon & Egner, 2004).*

d) The piece should be of sufficient difficulty for Derek to find it challenging though possible to learn after a number of hearings, given its complexity, tempo and length.

e) It should be well within his capacity to play so that technical considerations would not interfere with issues of music-processing. *Fulfilment criteria d) and e) is made possible by Adam Ockelford's intimate knowledge of the extent of Derek's technical abilities (Ockelford, 2007a).*

f) The stimulus should have a memorable name to enable Derek to form an association between the music and its title (Ockelford, 2012).

g) There should be no immediate repetitions of notes, so that accurate pairing of notes played by Derek and stimulus notes can take place at the analysis stage.

Appendix A(ii)

Table A1

Musical structure of *Romantic Rollercoaster*

Section	Bars (bar.beat)
A	0.4-5.3
B	5.4-9.3
A (reprise)	9.4-13.3
B (transposed)	13.4-17.3
A (transposed)	17.4-21.3

Appendix B

Criteria used to ensure matched segments:

- There must be 10 segments for each of the four categories of expectancy.
- 4 of these 10 segments should consist of 3 sequential notes, 4 should consist of 3 simultaneously occurring notes (a chord), and 2 should consist of a combination of simultaneously and sequentially occurring notes.
- Each group of 10 segments should be equally distributed about the musical texture (as Derek's memory in the 'listen and play' study was more accurate for the outer parts of the texture, suggesting that these notes may be easier for him to memorise; Ockelford, 2012)
- Each group of 10 segments should be equally distributed within the musical structure (so segments should be equally distributed between structurally important and less important points).
- It would be impossible to ensure that segments from each of the four categories were distributed equally throughout the piece. This is because the method of rating expectancy was based on how many previously occurring groups within the piece predicted each note. This meant that notes with higher expectancy ratings could not occur at the beginning of the stimulus.

Appendix C

Scoring system for notes matched by Derek:

- If the note was at the correct octave, and the onset time of Derek's note was within -0.5 seconds and the imitation threshold (as determined by the imitation threshold experiment; see p. 16) of the stimulus note, score 1.
- If the note onset was within the correct time and was the correct pitch class but was at the wrong octave, score 0.5.
- All other notes score 0.
- Each of Derek's notes could only be scored once (in relation to one note of the stimulus).
- If Derek played the same note twice in succession, and both notes matched the above criteria, the first note was used in the analyses. Given that all notes which met the above criteria must have been predicted (or remembered, after the first trial), the duration of the first note of any such pair can be assumed to be of the length that Derek predicted (or remembered) it to be.

The timing thresholds therefore ensured that all notes that were included in the calculation of proportional accuracy must have been predicted (or remembered) accurately by Derek, rather than copied from the stimulus.

In addition, duration data was taken into account, to allow for consideration of another aspect of music and so give a more precise picture of Derek's overall accuracy. By virtue of the nature of the experiment, proportional onset accuracy was automatically included in the calculation of proportional pitch accuracy, and only notes scoring for proportional pitch accuracy were assessed for proportional duration accuracy. The latter was scored as follows:

- Notes which score 0.5 or 1 for pitch and are of duration within 0.1 seconds of the stimulus note duration, score 1.
- Notes which score 0.5 or 1 for pitch but are not of duration within 0.1 seconds of the stimulus duration, score 0.5.

The overall proportional accuracy for each note was calculated by finding the average of the pitch and rhythm scores for that note.

Appendix D

SPSS outputs

1. t-tests comparing proportional accuracy scores before and after the application of the imitation threshold.

Equal variances should not be assumed. P values should be halved (1-tailed).

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Trial 1	Equal variances assumed	18.026	.000	6.841	722	.000	.20580	.03008	.14674	.26486
	Equal variances not assumed			6.841	718.949	.000	.20580	.03008	.14674	.26486
Trial 2	Equal variances assumed	.432	.511	4.573	722	.000	.14019	.03066	.08001	.20038
	Equal variances not assumed			4.573	721.994	.000	.14019	.03066	.08001	.20038
Trial 5	Equal variances assumed	3.880	.049	1.236	722	.217	.03867	.03129	-.02276	.10011
	Equal variances not assumed			1.236	721.657	.217	.03867	.03129	-.02276	.10011
Trial 10	Equal variances assumed	1.264	.261	.513	722	.608	.01174	.02290	-.03323	.05671
	Equal variances not assumed			.513	720.421	.608	.01174	.02290	-.03323	.05671

2. The equivalent non-parametric tests (Kruskal Wallis; demonstrating the same overall result as when parametric statistics were used

Test Statistics[a,b]

	Trial1	Trial2	Trial5	Trial10
Chi-Square	42.260	18.345	1.197	.060
df	1	1	1	1
Asymp. Sig.	.000	.000	.274	.807

a. Kruskal Wallis Test

b. Grouping Variable: Threshold

3. Correlations of proportional accuracy with expectancy rating (considering notes played before the imitation threshold, across the whole dataset)

Correlations

			Expectancy
Spearman's rho	Trial1	Correlation Coefficient	.455**
		Sig. (2-tailed)	.000
		N	362
	Trial2	Correlation Coefficient	.383**
		Sig. (2-tailed)	.000
		N	362
	Trial5	Correlation Coefficient	.217**
		Sig. (2-tailed)	.000
		N	362
	Trial10	Correlation Coefficient	.064
		Sig. (2-tailed)	.224
		N	362

**. Correlation is significant at the 0.01 level (2-tailed).

4. Parametric statistical tests used in the item analysis

4x4 mixed ANOVA of expectancy category [1-4] and trial [1, 2, 5, 10] with post-hoc tests for expectancy category, and pairwise comparisons for trial:

Mauchly's Test of Sphericity[b]

Measure:MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Trial	.965	1.245	5	.941	.976	1.000	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.

a. May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

b. Design: Intercept + Expectancy

Within Subjects Design: Trial

Tests of Within-Subjects Effects

Measure:MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Trial	Sphericity Assumed	4.165	3	1.388	24.496	.000
	Greenhouse-Geisser	4.165	2.927	1.423	24.496	.000
	Huynh-Feldt	4.165	3.000	1.388	24.496	.000
	Lower-bound	4.165	1.000	4.165	24.496	.000
Trial * Expectancy	Sphericity Assumed	.569	9	.063	1.115	.359
	Greenhouse-Geisser	.569	8.780	.065	1.115	.359
	Huynh-Feldt	.569	9.000	.063	1.115	.359
	Lower-bound	.569	3.000	.190	1.115	.356
Error(Trial)	Sphericity Assumed	6.122	108	.057		
	Greenhouse-Geisser	6.122	105.363	.058		
	Huynh-Feldt	6.122	108.000	.057		
	Lower-bound	6.122	36.000	.170		

Tests of Between-Subjects Effects

Measure:MEASURE_1

Transformed Variable:Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	27.432	1	27.432	179.106	.000
Expectancy	4.164	3	1.388	9.063	.000
Error	5.514	36	.153		

Bonferroni and Games-Howell post-hoc tests for expectancy category:

Multiple Comparisons

Measure:MEASURE_1

	(I) Expectancy	(J) Expectancy	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Bonferroni	1	2	-.0271	.08751	1.000	-.2714	.2172
		3	.0479	.08751	1.000	-.1964	.2922
		4	-.3604*	.08751	.001	-.6047	-.1161
	2	1	.0271	.08751	1.000	-.2172	.2714
		3	.0750	.08751	1.000	-.1693	.3193
		4	-.3333*	.08751	.003	-.5777	-.0890
	3	1	-.0479	.08751	1.000	-.2922	.1964
		2	-.0750	.08751	1.000	-.3193	.1693
		4	-.4083*	.08751	.000	-.6527	-.1640
	4	1	.3604*	.08751	.001	.1161	.6047
		2	.3333*	.08751	.003	.0890	.5777
		3	.4083*	.08751	.000	.1640	.6527
Games-Howell	1	2	-.0271	.09251	.991	-.2891	.2349
		3	.0479	.09348	.955	-.2167	.3125
		4	-.3604*	.08840	.004	-.6118	-.1090
	2	1	.0271	.09251	.991	-.2349	.2891
		3	.0750	.08661	.822	-.1698	.3198
		4	-.3333*	.08110	.003	-.5628	-.1038
	3	1	-.0479	.09348	.955	-.3125	.2167
		2	-.0750	.08661	.822	-.3198	.1698
		4	-.4083*	.08221	.001	-.6411	-.1756
	4	1	.3604*	.08840	.004	.1090	.6118
		2	.3333*	.08110	.003	.1038	.5628
		3	.4083*	.08221	.001	.1756	.6411

Based on observed means.

The error term is Mean Square(Error) = .038.

*. The mean difference is significant at the .05 level.

Pairwise comparisons for trial:

Pairwise Comparisons

Measure:MEASURE_1

(I) Trial	(J) Trial	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	-.081	.051	.700	-.222	.060
	3	-.235*	.052	.000	-.380	-.090
	4	-.423*	.051	.000	-.567	-.279
2	1	.081	.051	.700	-.060	.222
	3	-.154*	.052	.033	-.300	-.008
	4	-.342*	.054	.000	-.493	-.190
3	1	.235*	.052	.000	.090	.380
	2	.154*	.052	.033	.008	.300
	4	-.188*	.059	.017	-.351	-.024
4	1	.423*	.051	.000	.279	.567
	2	.342*	.054	.000	.190	.493
	3	.188*	.059	.017	.024	.351

Based on estimated marginal means

a. Adjustment for multiple comparisons: Bonferroni.

*. The mean difference is significant at the .05 level.

Separate univariate ANOVAs for expectancy category, for each trial:

Trial 1:

Tests of Between-Subjects Effects

Dependent Variable:Trial1

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.312[a]	3	.437	6.117	.002
Intercept	2.101	1	2.101	29.385	.000
Expectancy	1.312	3	.437	6.117	.002
Error	2.574	36	.071		
Total	5.986	40			
Corrected Total	3.885	39			

a. R Squared = .338 (Adjusted R Squared = .282)

Trial 2:

Tests of Between-Subjects Effects

Dependent Variable:Trial2

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.739[a]	3	.580	9.527	.000
Intercept	3.854	1	3.854	63.331	.000
Expectancy	1.739	3	.580	9.527	.000
Error	2.191	36	.061		
Total	7.785	40			
Corrected Total	3.930	39			

a. R Squared = .443 (Adjusted R Squared = .396)

Trial 5:

Tests of Between-Subjects Effects

Dependent Variable:Trial5

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1.239[a]	3	.413	3.610	.022
Intercept	8.634	1	8.634	75.436	.000
Expectancy	1.239	3	.413	3.610	.022
Error	4.120	36	.114		
Total	13.993	40			
Corrected Total	5.360	39			

a. R Squared = .231 (Adjusted R Squared = .167)

Trial 10:

Tests of Between-Subjects Effects

Dependent Variable:Trial10

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	.442[a]	3	.147	1.929	.142
Intercept	17.009	1	17.009	222.601	.000
Expectancy	.442	3	.147	1.929	.142
Error	2.751	36	.076		
Total	20.201	40			
Corrected Total	3.193	39			

a. R Squared = .138 (Adjusted R Squared = .067)

5. Tests of normality and homogeneity of variance for the data used in the item analysis

Tests of Normality

		Shapiro-Wilk		
	Expectancy	Statistic	df	Sig.
Trial1	1	.643	10	.000
	2	.737	10	.002
	3	.640	10	.000
	4	.885	10	.148
Trial2	1	.784	10	.009
	2	.792	10	.012
	3	.703	10	.001
	4	.928	10	.430
Trial5	1	.845	10	.050
	2	.870	10	.101
	3	.808	10	.018
	4	.589	10	.000
Trial10	1	.895	10	.195
	2	.866	10	.089
	3	.853	10	.062
	4	.874	10	.111

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
Trial1	4.992	3	36	.005
Trial2	2.651	3	36	.063
Trial5	1.978	3	36	.135
Trial10	4.750	3	36	.007

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Expectancy

Within Subjects Design: Trial

6. Equivalent non-parametric statistic tests for the item analysis

These give the same overall results as the parametric tests used (except in the case of a Jonckheere-Terpstra test of expectancy category in trial one, which was close to significance).

Jonckheere-Terpstra test, to demonstrate effect of expectancy category in each of the four trials

Jonckheere-Terpstra Test[a]

	Trial1	Trial2	Trial5	Trial10
Number of Levels in Expectancy	4	4	4	4
N	40	40	40	40
Observed J-T Statistic	394.000	409.500	350.500	373.500
Mean J-T Statistic	300.000	300.000	300.000	300.000
Std. Deviation of J-T Statistic	37.728	39.904	40.700	40.535
Std. J-T Statistic	2.492	2.744	1.241	1.813
Asymp. Sig. (2-tailed)	.013	.006	.215	.070

a. Grouping Variable: Expectancy

(Significance level adjusted for multiple comparisons, and set at .0125)

Friedman test, showing significant differences between trials

Ranks

	Mean Rank
Expectancy	5.00
Trial1	1.90
Trial2	2.01
Trial5	2.74
Trial10	3.35

Test Statistics[a]

N	40
Chi-Square	113.994
df	4
Asymp. Sig.	.000

a. Friedman Test