Individual differences in music performance

John A. Sloboda

Music cognition depends on the existence and deployment of processes for detecting, storing and organizing musical materials according to underlying structural features. Common cultural experiences develop these processes to a certain degree, but specifically designed and supported learning environments are required to achieve the levels of expertise required to perform western art music. Certain motivational and social factors are therefore implicated in the maintenance of activities that promote skill-acquisition, such as practice. Expert musical performance is not just a matter of technical motor skill, it also requires the ability to generate expressively different performances of the same piece of music according to the nature of intended structural and emotional communication. This review examines these abilities and describes how some of them have been shown to have lawful relationships to objective musical and extra-musical parameters. Psychological research is thus engaged in a process of demystifying musical expertise, a process that helps to improve upon culturally prevalent, but ultimately non-explanatory, notions of inborn ‘talent’.

The vast majority of contemporary research on music cognition has focussed on perceptual processes in the listener. This is unsurprising and defensible: the musical experience of the listener is at the heart of all musical activity. Without heard experience composition and performance would have no purpose. In addition, the vast majority of the population in contemporary industrialized nations are listeners rather than performers. It makes sense to focus scientific effort on processes that are shared by the majority of a population. Taking into account performers as well as listeners, the core questions for the study of music cognition are: what representational and control processes underlie people’s ability to recognize, store, recall, transform and generate musical materials?

Research on music perception has established that, as for language, the cognition of music is underpinned by the human ability to extract, store and manipulate a range of abstract structural representations from a complex
multi-dimensional stimulus stream. Although this ability increases throughout life as a result of development and training, one of the most significant outcomes of recent research is the demonstration of how sophisticated these processes are, even in untrained young listeners. Informal interaction with musical materials over the course of normal development appears, in itself, to lead to a relatively high degree of receptive competence for handling the musical materials of that culture. Neural network modelling has successfully simulated central aspects of perceptual tasks in which knowledge of structural factors (often implicit rather than explicit) influences the response of the listener.

For instance, untrained listeners are able to judge musical sequences as well-formed (grammatical) or ill-formed (ungrammatical). They show superior recognition and recall for material that conforms to the structural norms of the prevailing musical idioms, and are likely to accept as similar or identical two different musical sequences that share the same underlying structural description (the musical equivalent of synonyms).

Despite the centrality of perception, the study of music performance has recently established itself as a significant sub-area within music cognition. Music performance has considerable intrinsic interest as an example of a highly complex perceptuo-motor skill, and it has been used as a window onto a better understanding of motor programming and control. In addition, just as the study of speech has shed important light on our understanding of language representation, so the study of performance has also increased our understanding of the organization of musical cognition. Measurement of the microstructure of performance (e.g. timing, prosody, errors) is crucial in both psycholinguistics and music psychology. Only in the last decade has the technical capability for extracting this type of information from musical performances been cheaply and widely available through MIDI (Musical Instrument Digital Interface) and other commercial products – developed mainly for performers and composers rather than researchers, but eagerly exploited by the latter.

This review focuses on two core phenomena in the study of music performance. The first of these is that high levels of music performance skill are rare, at least in industrialized societies. The vast majority of the population have at best rudimentary performance skills. The second phenomenon relates to the widely varying aesthetic and emotional impact that can be created in listeners by different performers, even when they are playing the same piece of music with the same level of overall technical and artistic competence. Two questions arise out of these commonplace phenomena: (1) what mechanisms underlie overall differences in level of performance abilities between individuals? (2) what characteristics of performances that are under the direct control of the performer lead to aesthetically relevant variation in the experience of the listener?

These are important questions, particularly because the folk psychology of the musical world can often seem to be designed to keep the answers shrouded in mystery. The invocation of concepts of ‘talent’ and ‘inspiration’ can often be used to put an end to further attempts to analyse and understand the underlying processes. Such concepts are often no more than redescriptions of the very phenomena that require explanation. They have no real explanatory power, but their power in the lives and motivations of individuals can be disproportionate. The number of people who disengage from musical activity based on the belief that they ‘lack musical talent’ constitutes a continuing cultural and educational tragedy. A proper and scientifically grounded understanding of the factors underlying individual differences has social and educational as well as scientific implications.

The nature of skilled musical performance

Skilled musical performance has two major components, a technical component and an expressive component. The technical component is related to the mechanics of producing fluent coordinated outputs. For instance, a technically competent piano performance may involve the execution of as many as 20 notes per second, where each note’s duration and loudness is controlled to within very narrow limits of tolerance, and where absolute synchronization between notes played by different hands and fingers is also required. These requirements mean that the constraints operating in music performance are somewhat different from those operating in speech production, or within the types of task (such as tapping or typing) that are habitually studied within the motor behaviour literature.

The expressive component of musical performance is derived from intentional variations in performance parameters chosen by the performer to influence cognitive and aesthetic outcomes for the listener. The main expressive parameters available to performers are those of timing (both in note-onset and note-offset), loudness, pitch and timbre (sound quality). The precise parameters vary from instrument to instrument. For instance, a pianist has limited opportunities to affect timbre and none to affect pitch. Expressivity is also related to knowledge of musical genres – what would be considered appropriate for Chopin would be completely unacceptable for Mozart.

Technical and expressive skill are separate components, even though they interact with, and depend upon, one another. Technical skill is, at least in theory, unrelated to the musical or artistic content of the music. It is possible to perform a piece of music with absolute technical mastery yet with no expressive skill whatsoever. Such performances approach the ‘dead-pan’ rendition that may be produced by programming a computer to deliver the notes of a composition with absolutely equal loudness, timing and timbral characteristics. Dead-pan renditions sound mechanical and lifeless, and have little aesthetic value. Expressive skill demands knowledge of the underlying structural and stylistic constraints of a piece or a genre. However, because effective expressive performance often requires very fine and subtle variations in performance parameters (of the order of 10s of milliseconds, for instance, in the timing domain), expressive intentions
Box 1. The cultural specificity of practice

Table I. Contrasting performance requirements of past centuries and modern-day performances*

<table>
<thead>
<tr>
<th>Performers of past centuries</th>
<th>Modern performances</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Music to be played</strong></td>
<td>Mostly own compositions</td>
</tr>
<tr>
<td><strong>Preparation for a specific performance</strong></td>
<td>Relatively short; often improvisation or sight-reading</td>
</tr>
<tr>
<td><strong>Specialization</strong></td>
<td>Many professional responsibilities: perform on several instruments, lead ensembles, compose, teach, etc.</td>
</tr>
<tr>
<td><strong>Training</strong></td>
<td>Often informal, with late start for instrumental study</td>
</tr>
</tbody>
</table>

* Adapted from Ref. a

Western art music of the last 300 years (sometimes generically called classical music) has been characterized by a steady average increase in technical difficulty over time. Lehmann and Ericsson specified four key contrasts between public music performances of past centuries and the present day that arise from this increase in difficulty (Ref. a) (see Table I). A necessary consequence of this is the flourishing of what Sloboda has described as the classical conservatoire tradition (Ref. b). The characteristics of this tradition are: (a) concern with accurate and faithful reproduction of a printed score, rather than with improvisation or composition; (b) the existence of a central repertoire of extreme technical difficulty; (c) definitions of mastery in terms of ability to perform items from a rather small common core set of compositions within a culture, and (d) explicit or implicit competitive events in which performers are compared with one another by expert judges on their ability to perform identical or closely similar pieces, such judgements forming an important element in decisions about progression and reward within the culture (p. 110). It is the technical demands of this particular repertoire which demand the intensive supervised learning regimes found among high achievers in the recently published studies. Other musical cultures (jazz, pop, folk) where the technical and social constraints are different (and often arguably looser) do not necessarily demand the heroic levels of organized practice demanded of art-music performers, and it would be unwise to generalize the results reported in this paper to such cultures, which have received almost no serious scientific study. One possible reason why the ‘talent’ hypothesis has been so prevalent in Western societies, but so signaly absent in more traditional societies, is that traditional societies have achieved a better match between the technical demands of the music of their cultures and the social resources available for most people to learn to achieve these demands. Such closer match may also be available in Western cultures outside the classical conservatoire tradition (Ref. c). It would be surprising indeed if the ubiquitous perceptual competence displayed even in untrained children were to be unmatched by an equally ubiquitous capacity for performance competence, given the appropriate conditions for its development.

These conditions are summarized by Ericsson as follows: ‘The most effective learning requires a well-defined task with an appropriate difficulty level for the particular individual, informative feedback, and opportunities for repetition and corrections of errors. When all these elements are present the term deliberate practice [can be used] to characterize training activities’ (Ref. d, pp. 20–21).

References


determinants of performance achievement

There now exists strong evidence that individual differences in both technical and expressive skill are strongly related to differences in the amount of relevant learning activities that have taken place. A number of studies have provided reliable indicators of relevant learning through estimates of total number of hours of formal deliberate practice over the lifespan (see Box 1). For instance, a comparison of different violin students in a music academy showed that those students rated by their teachers as most musically and technically advanced had accumulated vastly more hours of practice by the same age than those rated only average (e.g. 7000 hours as opposed to 3000 hours by the age of 18) (Ref. 18). A second study showed that the amount of practice correlated strongly with a number of tests of pure technical ability (e.g. finger tapping, bimanual co-ordination), even within a high ability group.

Sloboda et al. extended these findings to performers drawn from a wider range of ability levels, including a substantial cohort of young people who had learned instruments for less than three years before abandoning study. Such a cohort would normally be thought of as including mainly the ‘untalented’ or ‘less talented’. Sloboda et al. also used a more objective measure of musical performance achievement than the individual teacher ratings.
used in earlier studies. This measure was the level of achievement on national graded performance examinations (from Grade 1 (lowest) to Grade 8 (highest)). The data showed that there was a direct positive correlation between hours of practice and grade achievement, and the slope of this line was identical for all cohorts, whatever the level of final achievement. Figure 1 shows the average hours of practice required for the attainment of each grade level across all cohorts.

A correlation between amount of practice and level of achievement does not by itself prove that practice causes the achievement. Some proponents of the ‘talent’ account argue that it is the existence of talent that causes high levels of practice through, for instance, the positive feedback obtained by success\textsuperscript{22,23}. However, this explanation sits uncomfortably with a number of pieces of evidence. For instance, O’Neill\textsuperscript{22} measured musical and intellectual aptitude of a cohort of 7–9-year-old children who were just about to begin instrumental tuition and practice for the first time. She also measured a general motivational variable (persistence in the face of failure on the Wisconsin Card Sorting Test) and amount of practice undertaken during a music performance assignment that was given to all children some months into the learning process. Performance outcomes on this assignment were predicted solely by the general motivational variable and the amount of practice undertaken. Intelligence and musical aptitude had no influence on outcomes.

There is evidence that effective practice is difficult. It is neither inherently motivating (especially for young people\textsuperscript{23}), nor easy to do well. Indeed, without guidance and coaching, apprentice performers either engage in practice strategies that are clearly ineffective (e.g. repeatedly playing through entire pieces without working on problematic sections\textsuperscript{24}), or fail to maintain a regular practice schedule.

Historical research suggests that many piano prodigies had live-in tutors and coaches that monitored and supervised their practice activities on a daily basis\textsuperscript{25}. Contemporary evidence has been provided of a direct link between the intensity of parental intervention in learning activities (e.g. direct daily supervision of practice, direct contact with teachers) and achievement outcomes in a study of 240 young British musicians\textsuperscript{26}. Results like these suggest that high achievement in Western art music (see Box 1) is statistically rare because the necessary cognitive and motivational scaffolding that must be provided within the home is also equally rare.

Although none of the evidence reviewed above rules out the possibility that differential individual outcomes are directly contributed to by factors other than quantity and quality of practice (and those factors, such as motivation, that directly affect practising behaviour), it has proved remarkably difficult to find agreement on what would constitute acceptable evidence for the contribution of such factors. Many commentators are wedded to the belief that there must be some differences in basic cognitive capacity between high and low musical achievers (possibly inherited and not a result of practice). But until new forms of evidence become available, it may be impossible to characterize or quantify such supposed differences.

**Expressive variation between performances**

Mastery of musical performance does not mean conformity. Performances by internationally renowned musicians can be vastly different. Repp, in an extensive series of studies has measured physical and perceptual characteristics of a large number of performances of the same piece of music (partly from commercial recordings and partly from laboratory performances)\textsuperscript{27–29}. In many cases it is the most eminent performers whose performances are most exaggeratedly different, both from each other and from the statistical mean. This is not, however, arbitrary whim, as an abiding feature of expert musical performance, noted in the work by Seashore\textsuperscript{30} onwards, is the astonishing level of consistency between performances achieved by many expert performers on successive performances of the same piece. In many cases, expressive parameters are exactly reproduced (see Fig. 2). The implication has to be drawn that these expressive variations are intentional and planned.

Much of the research effort in this area of the last two decades has been devoted to probing how expressive performance is achieved and perceived. Are there rules and regularities underlying expressive deviations, and are there adequate theories to account for the effectiveness and impact of these deviations? Two lines of explanation have yielded the most fruit. One of these considers expressive variation as a means of signalling or emphasizing structural features of the music. The second considers expressive variation as a means of signalling information about the ‘character’ of the music, in particular its emotional significance. These explanations actually overlap, in that some forms of structural communication turn out, in and of themselves, to have emotional impact.
Expression as structural communication

A clear example of the use of expressive performance to communicate structure is provided by studies of pianists’ use of ‘rubato’. Rubato is the technical term for planned deviations from a regular pulse achieved by increasing and then decreasing the time interval between the onsets of successive notes, thus slowing and quickening the pulse. A range of studies have now established that solo pianists use rubato to signal structural boundaries in the music (e.g. ends of phrases)\textsuperscript{35,36}. This is normally achieved by maximal slowing of the tempo at the most significant boundaries, and speeding in between boundaries. These conventions are reflected in listener response, as an intriguing series of experiments by Repp shows\textsuperscript{32}. In this study he introduced a mechanical timing perturbation (a slight slowing of tempo) into an otherwise metronomic (dead-pan) synthesized musical performance, and asked listeners to detect its occurrence. Listeners were significantly poorer at this task when the perturbation occurred at a phrase boundary. The same perturbation in a mid-phrase location is instantly detectable. This finding may reflect the cultural assimilation of performance conventions by the listener (although later studies have provided evidence to suggest that elementary perceptual grouping processes are implicated in this phenomenon, which would provide a causal account of the development of rubato: players slow down at phrase boundaries because of the perceptual biases they possess as listeners)\textsuperscript{33,34}.

Expressively powerful performances may be created by the use of unexpected or unconventional devices by the performer. This phenomenon has been investigated by asking professional pianists to record their chosen interpretations of a single Chopin Prelude (no. 4 in E minor) onto a MIDI piano, which allowed extraction of note-by-note timing and intensity information\textsuperscript{35}. These performances were then played to a panel of musician listeners who were asked to adjust a movable pointer according to the degree of expressiveness present in the performance from moment to moment. This allowed an ‘expressiveness graph’ of each performance to be derived. Figure 3 shows the graphs of average rated expressiveness values across 27 listeners produced by two different performances. Although the graphs share the same global shape, rising to an emotional height in the final quarter of the piece before subsiding, there are significant local divergences. The most noticeable of these is near the beginning, where one performance suddenly rises in expressiveness.

Examination of the MIDI data showed objective performance features that differentiate performances at such points of emotional divergence. In this particular example, a slowing of tempo and a decrease in intensity is to be observed in the more emotional performance, one whole bar before the end of the first phrase, where such a slowing would more conventionally occur. It appears to be the stylistic unexpectedness of such expressive devices that gives them their aesthetic and emotional power. Corroboration that violations of structurally-based expectancies are implicated in emotional responses to music is also provided by analyses of listener responses to different compositions\textsuperscript{36,37}. Emotional ‘high-spots’ generating tears and other physiological manifestations tend to coincide with structural features that manipulate listener expectancies.

Expression as communication of character

It is a long-established finding that, within a musical culture, listeners show considerable agreement concerning the adjectives that best describe a particular passage of music\textsuperscript{38}. 
These agreements seem to be based on associative links between auditory features of music and non-musical phenomena. A compilation of findings from a range of studies shows that there are systematic relationships between aspects of musical materials such as speed, timbre, pitch range and perceived emotional qualities (such as happiness, sadness, solemnity). This work has recently been extended by studies that have shown that similar relationships might be extended to semantic dimensions (such as hot–cold, male–female, etc.). These results stimulate the conjecture that musical stimuli very often project a ‘virtual person’ in the mind of the listener.

Recently, a similar systematic approach has been applied to music performance. In these studies, performers were asked to play the same short piece of music to depict different emotional characters. These were the five emotions of happiness, sadness, fearfulness, anger and tenderness, which are generally agreed by emotion researchers to be among the most ‘basic’. Not only was there a strong relationship between intended emotion and objective performance features, but naïve listeners showed significant ability to identify the intended emotion. The relative contribution of different features (speed, intensity, articulation etc.) was investigated through the construction of computer-synthesized performances that preserved one or more of the dimensions of variation present in the original performances. It has been argued convincingly that the emotional ‘code’ used by performers and listeners is closely related to the prosodic code, which communicates emotion in human vocal expression, and thus may have at least some innate components (P.N. Juslin, unpublished). However, performers and listeners do vary in the efficiency with which they can use the emotional code, and Juslin has provided some data that suggests that the application of ‘universal’ emotional codes to the specific structures of music requires specific, targeted learning experiences. Thus, at least part of the differences in levels of expressive skill between performers may reflect the different learning environments in which they acquired their skill.

Conclusions
Individual differences between music performers and performances are complex and multidimensional. The psychological mechanisms underlying these differences are, therefore, equally complex and multiple. The examples given in this paper show that significant aspects of these differences can be formally characterized, are amenable to empirical psychological investigation, and can be accounted for in terms of theories and mechanisms that underlie a range of cognitive achievements. Music performance, whilst possibly mobilizing a unique combination of psychological resources, is just as amenable to a componential analysis as other more well-studied cognitive skills. It can be argued, therefore, that music performance is no more mysterious and inscrutable than any other human activity, and skilled musical activity is dependent on the mobilization and co-ordination of structural knowledge. Psychological research has contributed to a demystification of music perception and musical skill. However, we have a long way to go before a full predictive account of music performance is available. No machine programmed with the insights obtained by music psychologists has yet come anywhere close to capturing the richness and coherence of an expert live musical performance.

Outstanding questions
- How does content of practice of a piece of music moderate the strong main effects of amount of practice per se, in respect of technical and expressive musical skill?
- What factors determine whether an expressive deviation is experienced as appropriate (and thus aesthetically powerful) or simply arbitrary (and thus aesthetically disrupting)?
- Are there forms of emotional communication in music that transcend cultures, as the hypothesis of a ‘universal’ emotional code would demand?
- Do structural and ‘character-based’ expressive deviations account for the most significant expressive variation to be found in music performance, or do other features yet-to-be characterized contribute significantly to the variance?
- To what extent do existing methods of measuring listener response provide valid and reliable indicators of aesthetic and emotional involvement? Do brain imaging techniques offer improved means of tracking such involvement during the course of a piece of music?

References

BOOK REVIEWS

Publishers – send copies of books you would like to be reviewed in TICS to the Editor.
Readers – would you like to review a particular book for the journal? Contact the Editor to discuss the project.
Other types of publications (software, CD-ROMs, films, art exhibitions, etc.) will be considered as well. Send any ideas or proposals for reviews to the Editor.

Erratum

In the article by Keenan et al. in the September issue [Keenan, J.P. et al. (2000) Self-recognition and the right prefrontal cortex. TICS 4, 338–344], there was an error with respect to the citation of reference 35. The following sentence on p. 342 cited reference 33 but should instead have read: ‘Although the data are controversial, some have indicated that children with autism and adults with Alzheimer’s disease have self-recognition deficits.’

Also, the figure in Box 1 on p. 339 of the same article should have been credited to reference 39. Permission to reproduce this figure was obtained from the original publishers.

We apologise to the author and readers for this oversight.

Pil: 51364-6613(00)01544-8