Edukacja. An interdisciplinary approach 2017, 115–128
DOI: 10.24131/3724.170508
ISSN: 0239-6858

Assessment of instrumental music performance: definitions, criteria, measurement

Zofia Mazur, Mariola Łaguna

Institute of Psychology, The John Paul II Catholic University of Lublin*

The concept of music performance achievement or musical achievement is variably understood and operationalised in different ways in research. Assessing the level of instrumental music performance is a complex task, requiring the assessment of many aspects of music performance. The aim of this article is to analyse (a) the ways achievement in music performance is understood, (b) factors affecting the assessment of instrumental music performance, and (c) tools used to measure such achievements. We describe the distinctions made in terms of the level of music expertise attained and the quality of the performance. We also present factors considered when assessing performance achievement: types of musical performance tasks, choice of repertoire, assessment criteria, such as the overall impression of the performance, technical ability, expressive components, and the basic parameters of the quality of the performance. Finally, we discuss available assessment tools, indicating their limitations.

Keywords: music performance achievement; musical attainment; instrumental performance; musical education; measurement.

Music is a significant element of every person's life: everyone listens to music in one form or another, though fewer people perform or create it (McPherson and Renwick, 2011; Sloboda, 2000). Playing a musical instrument, which is one of the most advanced tasks for the human central nervous system, has many benefits, but it requires effort and perseverance (McPherson and Renwick, 2011; Schellenberg, 2004; 2005). According to Polish research, the lack of persistent practice and the difficulty in reconciling education in a music school and general school are the most common reasons for abandoning instrument playing (Twarowska, 2012). The high percentage of resignations (approx. 40% of students according to 2011 data) and the fact that students are discouraged from studying in a music school can also be explained by insufficient musical attainment, which has a negative effect on students' internal motivation for playing an instrument (Kaleńska-Rodzaj, 2014a; McPherson and Renwick, 2011; Twarowska, 2012).

The evaluation of musical performance, the main indicator of musical achievement (Kaleńska-Rodzaj, 2014b), constitutes the basis for being promoted to the next class in a music school, as well as receiving an award in a music competition. It also serves

* Address: Al. Racławickie 14, 20-950 Lublin, Poland.
E-mail: laguna@kul.pl

as feedback on how to improve one's performance, thus directing further development, as well as determining success and motivation for learning (Kaleńska, 2008).

At the same time, a fair evaluation of the performance and a reliable description of the quality of the performed work is a formidable challenge (Kaleńska-Rodzaj, 2014b, Manturzewska, 2014). Playing an instrument is a multidimensional activity, and therefore relevant aspects and methods of evaluating musical achievement must be selected (DeLuca and Bolden, 2014; McPherson and Thompson, 1998; Radocy, 1989; Sloboda, 2000). The concept of "musical achievement" or "music performance achievement" can be interpreted and measured in different ways (Bergee, 2007; Ciorba and Smith, 2009; Simon, 2014). The diversity of how these concepts are described makes it necessary to rationalise the ways of defining and measuring musical achievement in research on music education. Therefore, the purpose of our article is to analyse (a) the ways in which musical achievement is defined in research, (b) the factors determining the evaluation of achievement, and (c) the tools used to measure achievement. The analysis will focus on musical achievement in learning to play an instrument (performance achievement), allowing for a more precise definition of music achievement in this particular area, which differs from the definition of achievement in music theory, e.g. knowledge of musical forms, knowledge of music history or the ability to recognise different instruments (Radocy, 1989).

In order to answer the questions on how to define and measure musical achievement, a review of psychological and educational research on musical performance was conducted. In April 2016, the following databases were searched: Arts & Humanities Citation Index, ERIC, Informa – Taylor & Francis, JSTOR Archival Journals, OneFile, Periodicals Archive Online, ProQuest Education Journals, ProQuest Psychology Journals, ProQuest Research Library, Sage Journals, Sage Publications, Scopus, Social Sciences Citation Index, Taylor&Francis Online Journals. The keywords used to search the databases included the following phrases: "musical attainment" and "music performance achievement". The review also included selected Polish-language publications on the evaluation of playing a solo instrument. This analysis is a qualitative, narrative review of the research, and does not refer to quantitative indicators.

First, we will present two indicators of musical achievement in playing an instrument, which include an evaluation of the musical performance and the number of solo performances, and discuss the difficulties of such an approach. Next, we will discuss two criteria of achievement: the performance level criterion and the more specific performance quality criterion. The next section will focus on specific factors constituting "achievement in playing an instrument". Finally, we will discuss the tools used to measure performance achievement.

Performance evaluation and the number of performances as indicators of musical achievement in playing an instrument

From the perspective of Polish researchers who analyse the determinants of success in music education, the problem of the lack of a reliable definition of school achievements in playing an instrument is still valid and hinders the conduct of reliable scientific research (Manturzewska, 2014). In spite of the lack of a definition, it is assumed that the main indicator of achievement in school and professional music activity is the evaluation of musical performance by musical experts (Kaleńska-Rodzaj, 2014b; Manturzewska, 2006). In practice, such an evaluation takes place during examinations in music schools, music academies and music competitions.

A musical audition can also be organised at the initiative of a researcher. However, in order to measure achievement or abilities, the researcher should appeal to the opinion of music experts (Manturzewska, 1968).

The idea of considering musical performance evaluation as an indicator of achievement is quite problematic as it is related to the very nature of the phenomenon of musical performance, which is non-quantifiable; hence, it is difficult to determine precise evaluation criteria (Jordan-Szymańska, 2006). Even if the evaluation criteria are defined, music evaluators often use and interpret these criteria quite freely. This is because musical performance criteria are to a large extent aesthetic criteria. These types of criteria are determined "on the one hand by cultural conventions and preferences that are historically conditioned, on the other hand, they are related to the preferences of the evaluator, conditioned by the characteristics of his/her personality and the sum of past experiences" (Jordan-Szymańska, 2006, pp. 48–49).

In the educational practice of a music school or music academy, the imperfection of the performance evaluation criterion is manifested by the fact that the subject of the evaluation and the musical expert capable of making an accurate evaluation of the performance are not clearly defined (Manturzewska, 1990b). For example, the end-of-year school examination may focus on current skills in instrument playing, as well as the pace of progress made in the past year and involvement in practice. The Regulation of the Minister of Culture (Journal of Laws 2004 No. 214, item 2179, as further amended) in Poland states that the conditions and criteria for evaluating the quality of an artistic performance during an end-of--year exam should be determined by the persons managing the institution and included in its statutes. However, in practice, many music schools lack specific and detailed records regulating performance evaluation (Kaleńska, 2008; Kaleńska-Rodzaj, 2014b). The evaluation is often based on intuitive criteria, depending on the teacher's knowledge and experience. Likewise, there is no clarity as to who should evaluate the student, i.e. who is to be the "expert" – the teacher of the main instrument, who is emotionally (positively or negatively) connected with the student, and thus might be unable to pass an objective, impartial judgement, or another teacher of the instrument, who knows the student from previous auditions (Manturzewska, 1990b).

The recognition of experts' opinions as a criterion of achievement in learning to play an instrument, also in the context of a music competition, may give rise to justified doubts as to the reliability and credibility of such an opinion (Manturzewska, 1968). In this case, it is also legitimate to ask about the criteria used by the experts in providing their opinions. Does their evaluation reflect the actual characteristics of the performer or is it influenced by factors not related to the musician, such as the evaluator's preferences or mood? The process of evaluating musical performance is, after all, a derivative of the perception process, and therefore, the characteristics of the evaluator and the situation in which the evaluation is made will always modify the perception and influence the evaluation (Jordan-Szymańska, 2006).

In her research, Maria Manturzewska (1966; 1968; 2006) analysed the question of the credibility and reliability of music performance evaluation during music competitions. By analysing the ratings of the judges of the VI Fryderyk Chopin International Piano Competition, she proved that despite agreement in the ratings, the judges were guided by common criteria only to a small extent. On the one hand, the results show that the indicator on which the judges' based their opinions can be a reliable source of knowledge about the performance, but on

the other hand, they prove that the opinion of a single music expert – even if it is a top-ranking expert in an international jury – is unlikely to be an accurate indicator of musical achievement.

Although performance evaluation is the most commonly used achievement indicator in research, other methods exist to estimate the success of a musician. Not only the quality, but also the number of musical performances of a given musician constitute proof of his/her musical achievement (Bonneville-Roussy, Lavigne and Vallerand, 2011). Such a measure of success may prove appropriate, particularly in the case of solo performers of classical music. This may be explained by the fact that there is great rivalry among such musicians. As a result, we may assume that a musician who has not attained a certain level of achievement would not have been given a chance to play many concerts, because people responsible for the promotion of talents and the organisation of such performances are likely to engage the best musicians.

This method of estimating achievements also has its drawbacks. If we assume that the number of concerts is determined by the decisions of people organising such events, we are once again faced with the question about the experts who decide that certain musicians are better than others: what are their evaluation competences, what evaluation criteria do they use, and is their evaluation of the musician's performance accurate?

In summary, both performance indicators – the quality criterion and the number of performances – are based on the outcome of the musical performance evaluation. Regardless of the circumstances under which this evaluation is made (whether during concerts, examinations, competitions or research), it is an imperfect indicator of achievement because it is based on subjective, culturally and personally determined aesthetic judgements.

The level and quality of performance

Despite the lack of generally accepted, well-defined performance evaluation criteria (Kaleńska-Rodzaj, 2014b), the literature indicates that expert assessment covers both the quality of the performance and the level of expertise attained (Manturzewska, 1968). A review of English-language publications showed that both criteria are taken into account in the research of achievement in instrumental performance (Hallam, 2013), therefore, we will explain their significance.

The level of performance (Hallam, 2013; Lehmann, Sloboda and Woody, 2007) can be defined as the overall level of advancement in musical performance. In order to assess the performance level of a given person, one needs to relate his or her current musical skills to the expertise of others in the selected reference group. Researchers may, for example, be interested in the person's performance level compared to the rest of the population or in comparison with persons in a smaller social group, e.g. music school students.

Different performance levels and different criteria for classifying individuals to different levels are assumed. Figure 1 shows the proposal of Andreas Lehmann and colleagues (Lehmann et al., 2007), which distinguishes four levels of musical skills based on the musical training criteria of a given person (this principle does not refer to the distinction between the level of music experts and the elite of musicians). The lower part of the pyramid represents the level of performance achieved by persons who have not received specific training in music and have average musical skills, for example, they can sing a simple song. The middle level represents the abilities of students of music schools and music academies – this group is characterised by the most varied level of performance. Music experts are a less numerous group that have received advanced musical education.

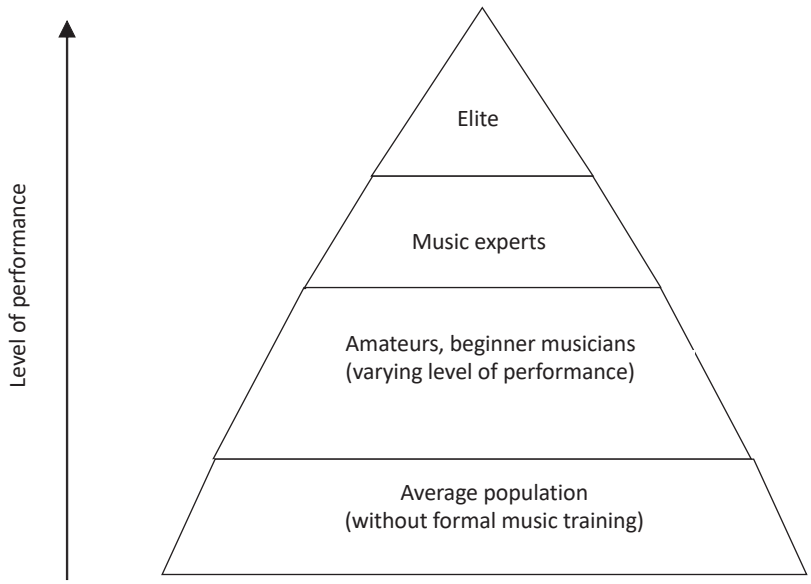

Figure 1. The model of the distribution of musical skills in society.
Based on: Lehmann et al. (2007, p. 16).

The elite includes successful individuals, recognised by other music experts as models to follow. They are, for example, world-famous musicians making studio recordings.

The concept of musical skill levels was utilised in Manturzewska's research (1990a). Based on the analysis of biographies of musicians, she distinguished six stages in the development of a professional musician. Some of the criteria used to distinguish the levels or phases of development include: the criterion of criterion of changes in musical behaviour and form of musical expression, the criterion of changes in motivation and interest in music, the criterion of changes in musical achievement and activity.

Another classification of performance levels, popular in the literature on music performance, differentiates and clarifies the achievement of students of music schools (Figure 2). This was the division used, among others, by Susan Hallam in her research (2013; Hallam et al., 2012). Hallam used a rating scale according to the class divisions of a music school. She assumed that the level of students' expertise attained in a music school was compatible with level of the most recently graded instrumental examination that they had taken. The criterion for assigning a specific performance level according to this classification includes the fulfilment of the programme requirements for the specific year of study. At the same time, persons who have similar skills at the same level of performance – namely students from the same class – can differ in terms of quality (Figure 2). This is manifested, for example, in the different marks received by students during the school exam. Noting this, Hallam introduced an additional criterion for determining musical accomplishments in instrument playing – the quality of performance (Hallam, 2013; Hallam et al., 2012). This dimension allows achievement to be differentiated with greater accuracy. However, this is not a fully independent dimension of defining performance, but rather an ancillary dimension

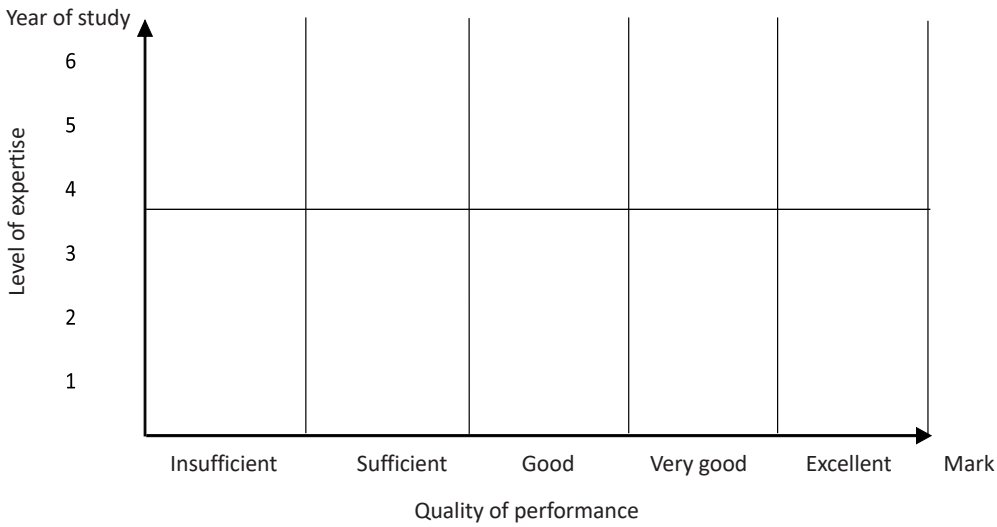

Figure 2. Two dimensions of achievement evaluation: level of expertise attained and quality of performance on the example of music school students.

Based on: Hallam (2013).

for the level of performance, which allows for a more precise characterisation of musical achievement. As a result, as part of the classification of performance levels in the system of the school's class years, Hallam suggested that the mark obtained in a school examination could be a measure of the quality of performance, as well as a measure of achievement in playing an instrument.

The distinction between two achievement criteria in instrument playing (the level and quality of performance criteria) allows both existing research to be interpreted and further research to be planned. By examining the level of the performance, we measure a more comprehensive phenomenon that, in addition to proper musical skills or abilities, to some extent (we do not know exactly to what extent) may include other phenomena, which are not instrument playing skills, e.g. showing the differences in the age or amount of time devoted to music practice of the performers. In the case of research on the quality of performance, the research group is often more homogeneous in terms of age and the number of years of study. Such conditions allow for a more accurate measurement of the differences in the achievements of instrument playing.

Factors determining the evaluation of achievements in playing an instrument

A more specific description of the level of performance and quality of performance (Hallam, 2013; Lehmann et al., 2007) still leaves the question of the detailed criteria for evaluating musical achievement and the factors that influence such an evaluation. For example, what is really taken into account when evaluating whether a given student fulfils curriculum requirements? How can we measure the overall level of musical skills? What are the real differences between persons presenting a better or worse level of quality? Out of the many factors influencing the measurement of performance (McPherson and Thompson, 1998), we will focus on

those that are relevant to the selection of evaluation criteria by the music expert. We will then discuss the principles of performance evaluation and present suggestions for sets of criteria that may be used by researchers or teachers in their work.

Two groups of factors that influence the measurement of performance should be taken into account when selecting evaluation principles and criteria: factors relating to the performer and factors relating to the circumstances of the performance (McPherson and Thompson, 1998; Miklaszewski, 2006). On the one hand, we should consider factors relating to the performer: the type of instrument, the type of tasks and repertoire performed by the person whose achievements are being determined. On the other hand, the decision on the criteria used and the subject of the evaluation of achievement depend on the situation and the specific purposes of the performance evaluation, e.g. whether it is a festival, an audition to a music school, a scientific study or a music competition. The criteria and rules of evaluation can be treated as the third group of factors affecting the subject of the measurement, which specifies the meaning of musical achievement in a given measurement.

When looking for performance achievement criteria that may be relevant to the performer, Gary McPherson identified five different types of tasks determining a musician's skills (see McPherson and Thompson, 1998). He pointed out potential tasks, such as: (a) performing rehearsed repertoire, (b) sight reading, (c) improvising, (d) playing by ear, (e) playing from memory. According to the author, each of these five types of tasks requires the selection of individual, specific criteria for evaluating the quality of performance.

In addition to selecting the type of task to be evaluated, other decisions taken prior to the evaluation of musical performance are also important. They concern the musical repertoire constituting the basis of the evaluation (McPherson and Thompson, 1998). In order to compare the achievement of persons playing different instruments, one should select a repertoire taking into account the differences in the scale and technical capabilities of each instrument. The choice of the music piece is also significant in determining detailed evaluation criteria, which will be further clarified.

Research confirms that factors relating to the situation of the evaluation affect the measurement of musical achievement (Gillespie, 1997; Manturzewska, 2006; Wapnick and Darrow, 2013). Although the performance evaluator usually has no impact on the purpose and situation of the evaluation, he or she can decide on evaluation criteria that may be more or less appropriate in the context of a given performance (Miklaszewski, 2006). The set of criteria can also vary depending on the type of available information about the performance, e.g. whether it will be a live performance or a recorded performance; will auditory or visual-auditory material be available? The evaluation of posture when playing an instrument will not be possible if the judges only have audio material.

Criteria and principles for the evaluation of musical performances

Being aware of such factors as the type of instrument, the type and characteristics of the music to be played and the performance context, it is possible to define performance evaluation criteria in order to improve the reliability and relevance of the assessment (Miklaszewski, 2006). Different evaluation rules are used, i.e. different types of evaluation criteria for a piece performed by a musician. The type of evaluation criteria used determines the extent of its subjectivity, as well as what its subject will be.

Some of the empirical research, as well as the practice of evaluation in teaching, bases

the evaluation of performance on the "overall impression of the whole performance" (McPherson and Thompson, 1998; Saunders and Holahan, 1997). Researchers also apply more detailed evaluation criteria, such as the technical and expressive level of the performance (Lehmann et al., 2007; Sloboda, 2000) or assess the quality of individual micro-skills presented in the work, such as the adequacy of the played notes, rhythmic correctness, appropriate dynamic changes, quality of articulation (Ciorba and Smith, 2009; McPherson and Thompson, 1998).

The most subjective evaluation of musical performance occurs when the applied principle of evaluation consists of the "overall impression of the whole performance". In such a case, the subject and criteria of evaluation are difficult to determine. Many people evaluating the quality of musical performance in this manner (known in research as "competent judges") are not aware of what determines their judgements (Ciorba and Smith, 2009).

When trying to avoid such constraints, researchers of musical achievement (e.g. Parncutt and McPherson, 2002; Sloboda, 2000) propose that in assessing the success of a given musician, we should take into account his/her technical skills and expressiveness. The technical component of a performance refers to the mechanics of producing sound (DeLuca and Bolden, 2014; Sloboda, 2000). When assessing this component, we focus on the fluidity and coordination of movement, the technique or proficiency in performance. Expressive skills, on the other hand, manifest themselves in the so-called musical expression or the musical interpretation of the piece presented by the musician (Sloboda, 2000). We then talk about the communication function of musical expression – through interpretation, the musician gives the listener information about the structure of the musical composition (such as the distribution of accents, metric structure and information relating to emotions (e.g. joy, sadness, anger) which he/she noticed in the piece (Gabrielsson, 2003; Lehmann et al., 2007).

In so far as technical skills can be defined in a more objective manner, the expressive aspects of a performance are perceived subjectively and are much more difficult to measure (DeLuca and Bolden, 2014; Radocy, 1989; Sloboda, 2000). Whether the expressive aspects of a musician's performance are evaluated positively depends on whether the judges share his/her interpretation of the musical composition, or if such an interpretation is convincing (McPherson and Thompson, 1998). It is usually easier for judges to agree that a note has been played erroneously than to arrive at a consensus on the alleged lack of musical expression in a performance (Radocy, 1989)[1]. Therefore, the evaluation of expressive skills is often overlooked (DeLuca and Bolden, 2014) and frequently dismissed by researchers (especially in studies striving for an objective measurement of achievement, for example, by using a computer program; Lehmann and Ericsson, 1996).

In musical performance, we can distinguish more specific parameters of performance quality, such as: note accuracy, rhythmic accuracy, phrasing, articulation, effective dynamics, appropriate tempo or appropriate sense of style (McPherson and Thompson, 1998). A researcher or teacher who takes into account detailed performance criteria, should ensure that they are consistent with the above-mentioned factors relating to the performer, i.e. relevant to the technical capabilities of the instrument, the type of task and musical repertoire as well as being adequate given the context and

[1] However, the question of whether the notes are played in a correct manner is related to the intonation of playing certain instruments. This quality is often included in the subjective aspects of performance evaluation (Miksza, 2005; Zdzinski, 1996), which means that not in every case will the evaluators reach a consensus as to an erroneously played note.

purpose of the audition (McPherson and Thompson, 1998; Miklaszewski, 2006).

To summarise, the rules of evaluation can be listed starting from those in which we follow the most general criteria and those which include the most detailed and specific evaluation criteria. Music experts have long discussed whether performance evaluation should be based on the evaluator's intuition or on more specific criteria (McPherson and Thompson, 1998). Studies show that the evaluation of overall impression is not equivalent to the summary evaluation of more detailed performance aspects (Ciorba and Smith, 2009). Hence, it may be beneficial for teachers and researchers to take into account the different evaluation principles in evaluating the performance itself. Thus, performance evaluation could include both an overall impression, a technical and expressive assessment, and an assessment of detailed criteria (Table 1).

Since the selection of particular evaluation criteria depends on many characteristics of the performer and the evaluation situation, numerous examples of criteria can be found in the music literature (see a review in: Miklaszewski, 2006). Although research attempts aimed at ordering the performance criteria used in practice have already been made in Poland (Kaleńska-Rodzaj, 2014b), further research on the criteria for evaluating musical performance is needed (Miklaszewski, 2006). According to Kacper Miklaszewski (2006), prospective researchers are expected to identify and empirically verify criteria, which would contribute to the construction of objective performance evaluation methods and achievement in playing an instrument.

Scales and tools for evaluating achievement in playing an instrument

The presented overview of factors determining the evaluation of instrumental performance that relate to the performer and the repertoire, as well as the evaluation situation, the criteria and the evaluation rules, does not, however, specify how to measure success in playing an instrument. Numerous scales and measuring tools have been developed to harmonise and align the rules for assessing musical achievement. They will be discussed in this section, with an indication of their constraints.

Typical tools for evaluating performance are based entirely on the personal impression of the evaluator as to the quality or character of the performance (Ciorba and Smith, 2009). According to this approach,

Table 1
An example of detailed performance criteria for playing the piano

General quality of the performance	Level of performance	Category	Criteria
	Technical level	Sound	The scale of the timbre, the richness of articulation, the use of the pedal, the richness of the rhythmic formulas used, the scale, the type and beauty of the voice.
		Text	Accuracy of pitch, precision in rhythm, proficiency in performance of passages, scales and other technical elements.
	Expression (interpretation)	Convention	Style, realisation of articulation or agogic indications, understanding of the composer's distinctive conception, compliance or non-compliance with the accepted rules of interpretation.
		Proposal	Artistic impression, artistic value, concept of composition, maturity.

Based on: Miklaszewski (2006).

the evaluator declares whether the performers' overall skills are above average, average, below average or poor (Saunders and Holahan, 1997). The assessment may concern the overall impression of the entire performance or the specific dimensions of the performance, e.g. impressions on the technique or intonation. These quality and performance tools have one major drawback: they do not provide information on specific indications of the level of performance. We do not know the nature of the student's performance that influenced the judges to evaluate a performance as above-average or average (Saunders and Holahan, 1997).

When evaluating musical performance, a Likert's rating scale is sometimes used (Ciorba and Smith, 2009). It allows the level of agreement among judges in reference to different categories of the performance to be estimated. By using these scales, judges rate on a continuum (e.g. a 5-point scale, where 1–*I strongly disagree* and 5–*I strongly agree*) to indicate the extent to which they agree with a particular statement regarding a given aspect of the performance (e.g. "Performance was rhythmically correct"; Saunders and Holahan, 1997). These scales are often characterised by a high degree of reliability (assessed as the level of agreement among competent judges), but there are doubts as to their relevance. As in the case of the previously described tools for evaluating overall impression, these scales do not provide specific descriptions of evaluation criteria. Also, the performance standard to which the assessment refers is not precisely defined (for example, it is unclear what a "rhythmically correct performance" really means).

The so-called criteria-specific music performance rating scales (Saunders and Holahan, 1997) are considered more useful diagnostic tools in evaluating achievement in playing an instrument. Unlike the previously mentioned types of tools, these scales describe each level of the evaluated skill. The evaluators are obliged to indicate which of the provided criteria best describes the level of the performance. In this way, they note what they actually heard in the performance, not what they liked or disliked, or the extent to which they agreed or disagreed that the performance was closer to an indeterminate standard.

We can also use the so-called multidimensional assessment rubrics to evaluate performance (Ciorba and Smith, 2009; Cooper and Gargan, 2009; DeLuca and Bolden, 2014). Their advantage is that they include a description of each skill level and allow for an overall evaluation of the playing of a given instrument, taking into account many aspects of the performance (Ciorba and Smith, 2009). For example, the scales include criteria for evaluating the technical and expressive component of a performance (DeLuca and Bolden, 2014). The use of rubrics allows judges to evaluate the performance more reliably (DeLuca and Bolden, 2014; Jonsson and Svingby, 2007). They are helpful in evaluating the performance of persons playing different instruments, at different stages of education (Ciorba and Smith, 2009; Latimer, Bergee and Cohen, 2010).

The literature also provides scales which include criteria for the evaluation of playing on a given type of instrument (McPherson and Thompson, 1998). For example, the Watkins-Farnum Performance Scale[2] is used to evaluate achievement in playing wind instruments. The Brass Performance Rating Scale includes criteria for evaluating brass instruments. The Clarinet Performance Rating Scale is helpful in evaluating clarinet performance (McPherson and Thompson, 1998; Zdzinski, 1991). There are also scales for evaluating achievement in playing string instruments (Zdzinski and Barnes, 2002). They include the Music

[2] The Polish adaptation of the Watkins-Farnum Performance Scale was prepared by Miklaszewski.

Performance Assessment Scale developed by Julia Kaleńska-Rodzaj (2014b) for the evaluation of the performance of violinists and violists[3] and the Piano Performance Scale developed by Teresa Manasterska (1980) and many others (Ciorba and Smith, 2009).

Some of the tools for evaluating musical achievement are designed to measure specific types of skills. For example, the Watkins-Farnum Performance Scale measures sight reading skills (Zdzinski, 1991). This scale includes not only criteria for evaluating instrument playing, but also provides proposals for the musical pieces that can be played by the participants of a study (a series of exercises with increasing difficulty). Kevin Watson, on the other hand, has developed the Jazz Improvisation Performance Achievement Measure (Watson, 2010).

The Watkins-Farnum Performance Scale, as well as the method of measuring the technical aspects of performance used in the study by Laura Stambaugh and Steven Demorest (2010), refer to those performance parameters whose evaluation can be more objective (e.g. note accuracy, rhythmic accuracy). These methods allow the calculation of the sum of errors made in given performance dimensions, observed by the listener in specific sections of the musical text (e.g. errors that occurred in each bar measure of the work).

There have also been attempts to objectively measure achievement in performing musical pieces, e.g. by using computer programs (Zdzinski, 1991). The evaluation of a performance with the use of computer technology is highly reliable, but is usually limited to only several aspects of the performance. The pitch and rhythm accuracy or articulation are most frequently taken into account (Zdzinski, 1991). Other aspects of the performance, such as expression, are often omitted in the evaluation. An auxiliary tool, which includes criteria for the subjective dimensions of a performance, is a supplement designed by Stephen Zdzinski (1996) – the Performance Rating Scale Supplement. It is used to evaluate those performance aspects that are not taken into account by methods based on an objective point conversion system, namely, musicality, intonation/tone quality or technique.

Most of the above-mentioned tools are not yet available in Poland. However, many translations have been made in Poland of psychometric tools relating to the measurement of skills and broadly understood musical achievement. Some of them have not been published yet[4].

Conclusions

We tried to show that the evaluation of the performance of a musical piece is a complex matter. However, both research into musical achievement, as well as the daily practice of teaching instrument playing, cannot avoid the definition and measurement of achievement. Although many studies address the issue of predicting success in music studies, the problem of finding a suitable criterion for defining achievement in playing an instrument, as well as the relevant measurement methods, is still an ongoing issue (Manturzewska, 2014). A better understanding of the complexity of music performance achievement and the dependence of evaluation on the methods of measurement allows us to make more informed decisions at the stage of evaluating the performance of musical

[3] The Kaleńska-Rodzaj scale can also be used for a group of people playing other instruments. However, as the author points out, in order to verify this, further research is needed with various groups of musicians.

[4] Readers looking for diagnostic tools in the field of music education available in the Polish language are referred to the electronic database of diagnostic instruments and tools in the field of music psychology conducted by the Department of Psychology of Music at the Fryderyk Chopin University of Music: http://psychologia.chopin.edu.pl/bazaprac/narzedzia-diagnostyczne/

pieces, whether in scientific research, the educational practice of music schools or music competitions.

This article discussed the indicators of achievement in playing an instrument that are currently used and presented in the literature, i.e. the evaluation of musical performance and the number of performances. By taking a closer look at the issue of musical performance evaluation, we presented the limitations of such an approach in defining musical achievement. We also made a distinction between the quality and level of performance (Hallam, 2013; Lehmann et al., 2007). Furthermore, we discussed the factors that are taken into account when evaluating musical achievement in scientific research and which actually affect evaluation. Reflection on this particular issue may contribute to the improved relevance and reliability of evaluation performed by future researchers and teachers, who will be more careful in selecting appropriate criteria when evaluating a particular instrument, a specific type of music and tasks, in the context of a particular evaluation situation, relating to the circumstances and goals of the performance. We also presented the tools used to measure achievement in playing a given instrument. The available methods differ in their construction and have different limitations. The decision to select a particular measurement method is important in that, to a certain extent, it constitutes the subject of the measurement of achievement in playing an instrument.

We hope that the review of the definitions of musical achievement and evaluation methods will inspire teachers and researchers to look for increasingly effective and cost-effective methods of evaluating instrument playing. Further research on the issue of music performance achievement is necessary, and the issue of measurement is important, not only for the development of scientific knowledge, but also for evaluation practice in schools and during competitions.

Literature

Bergee, M. J. (2007). Performer, rater, occasion, and sequence as sources of variability in music performance assessment. *Journal of Research in Music Education*, *55*(4), 344–358. DOI: 10.1177/002242940831751

Bonneville-Roussy, A., Lavigne, G. L. and Vallerand, R. J. (2011). When passion leads to excellence: the case of musicians. *Psychology of Music*, *39*(1), 123–138. DOI: 10.1177/0305735609352441

Ciorba, C. R. and Smith, N. Y. (2009). Measurement of instrumental and vocal undergraduate performance juries using a multidimensional assessment rubric. *Journal of Research in Music Education*, *57*(1), 5–15. DOI: 10.1177/0022429409333405

Cooper, B. S. and Gargan, A. (2009). Rubrics in education: old term, new meanings. *The Phi Delta Kappan*, *91*(1), 54–55. DOI: 10.1177/003172170909100109

DeLuca, C. and Bolden, B. (2014). Music performance assessment exploring three approaches for quality rubric construction. *Music Educators Journal*, *101*(1), 70–76. DOI: 10.1177/0027432114540336

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, *31*(3), 221–272. DOI: 10.1177/03057356030313002

Gillespie, R. (1997). Ratings of violin and viola vibrato performance in audio-only and audiovisual presentations. *Journal of Research in Music Education*, *45*(2), 212–220. DOI: 10.2307/3345581

Hallam, S. (2013). What predicts level of expertise attained, quality of performance, and future musical aspirations in young instrumental players? *Psychology of Music*, *41*(3), 267–291. DOI: 10.1177/0305735611425902

Hallam, S., Rinta, T., Varvarigou, M., Creech, A., Papageorgi, I., Gomes, T. and Lanipekun, J. (2012). The development of practising strategies in young people. *Psychology of Music*, *40*(5), 652–680. DOI: 10.1177/0305735612443868

Jonsson, A. and Svingby, G. (2007). The use of scoring rubrics: reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. DOI: 10.1016/j.edurev.2007.05.002

Jordan-Szymańska, A. (2006). Psychologiczne uwarunkowania oceny wykonań muzycznych. In M. Chmurzyńska and B. Kamińska, *Ocenianie wykonań muzycznych* (pp. 45–75). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Kaleńska, J. (2008). Efekt Pigmaliona w praktyce. Obraz ucznia zdolnego u nauczycieli szkół muzy-

cznych a oczekiwania wobec jego funkcjonowania na scenie. In W. Limont, J. Cieślikowska and J. Dreszer (eds.), *Zdolności. Twórczość. Talent* (vol. 2, pp. 13–28). Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Kaleńska-Rodzaj, J. (2014a). Motywacja do nauki gry na instrumencie muzycznym w świetle modelu Jachuelynne S. Eccles. *Annales Universitatis Paedagogicae Cracoviensis*. *Studia Psychologica*, 1, 8–23.

Kaleńska-Rodzaj, J. (2014b). Skala oceny występu muzycznego – w poszukiwaniu kryteriów oceny estetycznej wykonań muzycznych. In R. Lewandowski and J. Kaleńska-Rodzaj (eds.), *Psychologia muzyki. Współczesne konteksty zastosowań* (pp. 111–135). Gdańsk: Harmonia Universalis.

Latimer, M. E., Bergee, M. J. and Cohen, M. L. (2010). Reliability and perceived pedagogical utility of a weighted music performance assessment rubric. *Journal of Research in Music Education*, *58*(2), 168–183. DOI: 10.1177/0022429410369836

Lehmann, A. C. and Ericsson, K. A. (1996). Performance without preparation: structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology: A Journal of Research in Music Cognition*, *15*(1–2), 1–29. DOI: 10.1037/h0094082

Lehmann, A. C., Sloboda, J. A. and Woody, R. H. (2007). *Psychology for musicians: understanding and acquiring the skills*. New York: Oxford University Press.

Manasterska, T. (1980). *Problemy konstrukcji Skali osiągnięć pianistycznych jako metody obiektywizacji oceny wyników nauczania w procesie kształcenia pianisty*. [Unpublished doctoral thesis.] Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Manturzewska, M. (1966). Zgodność ocen wykonawstwa pianistycznego wydawanych przez ekspertów muzycznych. *Biuletyn Psychometryczny*, 1, 111–115.

Manturzewska, M. (1968). Z badań nad ocenami wykonawstwa muzycznego wydawanymi przez ekspertów muzycznych. *Zeszyty Naukowe PWSM w Warszawie*, 3, 113–116.

Manturzewska, M. (1990a). A biographical study of the life-span development of professional musicians. *Psychology of Music*, *18*(2), 112–139. DOI: 10.1177/0305735690182002

Manturzewska, M. (1990b). Badania ocen osiągnięć muzycznych wydawanych przez nauczycieli szkół muzycznych. In M. Manturzewska and H. Kotarska (eds.), *Wybrane zagadnienia z psychologii muzyki* (pp. 295–304). Warszawa: Wydawnictwa Szkolne i Pedagogiczne.

Manturzewska, M. (2006). O trudnej sztuce oceniania wykonań muzycznych. In M. Chmurzyńska and B. Kamińska (eds.), *Ocenianie wykonań muzycznych* (pp. 9–44). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Manturzewska, M. (2014). *Psychologiczne wyznaczniki powodzenia w studiach muzycznych* (2nd ed.). Warszawa: Centrum Edukacji Artystycznej, Uniwersytet Muzyczny Fryderyka Chopina.

McPherson, G. E. and Renwick, J. M. (2011). Self-regulation and mastery of musical skills. In B. Zimmerman and D. H. Schunk (eds.), *Handbook of self-regulation of learning and performance* (pp. 234–248). New York: Routledge.

McPherson, G. E. and Thompson, W. F. (1998). Assessing music performance: issues and influences. *Research Studies in Music Education*, *10*(1), 12–24. DOI: 10.1177/1321103X9801000102

Miksza, P. (2005). The effect of mental practice on the performance achievement of high school trombonists. *Contributions to Music Education*, *32*(1), 75–93.

Miksza, P. (2007). Effective practice: an investigation of observed practice behaviors, self-reported practice habits, and the performance achievement of high school wind players. *Journal of Research in Music Education*, *55*(4), 359–375. DOI: 10.1177/0022429408317513

Miklaszewski, K. (2006). Koncepcja kryteriów oceny wykonawstwa muzycznego. In M. Chmurzyńska and B. Kamińska (eds.), *Ocenianie wykonań muzycznych* (pp. 77–93). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Parncutt, R. and McPherson, G. (2002). *The science and psychology of music performance: creative strategies for teaching and learning*. Oxford: Oxford University Press.

Radocy, R. E. (1989). Evaluating student achievement. *Music Educators Journal*, *76*(4), 30–33. DOI: 10.2307/3401012

Rozporządzenie Ministra Kultury z dnia 29 września 2004 r. w sprawie warunków i sposobu oceniania, klasyfikowania i promowania uczniów oraz przeprowadzania sprawdzianów i egzaminów w publicznych szkołach i placówkach artystycznych (Dz. U z 2004 r. nr 214, poz. 2179 z późn. zm.).

Saunders, T. C. and Holahan, J. M. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. *Journal of Research in Music Education*, *45*(2), 259–272. DOI: 10.2307/3345585

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, *15*(8), 511–514. DOI: 10.1111/j.0956-7976.2004.00711.x

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science*, *14*(6), 317–320. DOI: 10.1111/j.0963-7214.2005.00389.x

Simon, S. H. (2014). Using longitudinal scales assessment for instrumental music students. *Music Educators Journal*, *101*(1), 86–92. DOI: 10.1177/0027432114539704

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, *10*(4), 397–403. DOI: 10.1016/S1364-6613(00)01531-X

Stambaugh, L. A. and Demorest, S. M. (2010). Effects of practice schedule on wind instrument performance: a preliminary application of a motor learning principle. *Update: Applications of Research in Music Education*, *28*(2), 20–28. DOI: 10.1177/8755123310361768

Twarowska, M. E. (2012). Uczeń w systemie – od rekrutacji do okresu poszkolnego. W: W. Jankowski (red.), *Raport o stanie szkolnictwa muzycznego I stopnia. Diagnozy, problemy, wnioski modelowe* (s.107–134). Warszawa: Instytut Muzyki i Tańca.

Wapnick, J. and Darrow, A. A. (2013). Sectional versus intact evaluations of four versions of a Chopin etude. *Journal of Research in Music Education*, *60*(4), 462–474. DOI: 10.1177/0022429412465318

Watson, K. E. (2010). The effects of aural versus notated instructional materials on achievement and self-efficacy in jazz improvisation. J*ournal of Research in Music Education*, *58*(3), 240–259. DOI: 10.1177/0022429410377115.

Zdzinski, S. F. (1991). Measurement of solo instrumental music performance: a review of literature. *Bulletin of the Council for Research in Music Education*, 109, 47–58.

Zdzinski, S. F. (1996). Parental involvement, selected student attributes, and learning outcomes in instrumental music. *Journal of Research in Music Education*, *44*(1), 34–48. DOI: 10.2307/3345412

Zdzinski, S. F. and Barnes, G. V. (2002). Development and validation of a string performance rating scale. *Journal of Research in Music Education*, *50*(3), 245–255. DOI: 10.2307/3345801

A preliminary version of this article was published in Polish in *Edukacja*, *139*(4), 2016.