MUSIC EDUCATION RESEARCH
2024, VOL. 26, NO. 5, 634–649
https://doi.org/10.1080/14613808.2024.2365719

Routledge
Taylor & Francis Group

OPEN ACCESS

Check for updates

The instrumental practice goal realization inventory: from intention to evaluation

Zofia Mazur-Socha[a], Mariola Łaguna[a] and Peter Gollwitzer[b,c,d]

[a]Department of Psychology, The John Paul II Catholic University of Lublin, Lublin, Poland; [b]Department of Psychology, New York University, New York, NY, USA; [c]Department of Psychology, University of Konstanz, Konstanz, Germany; [d]Department of Political and Social Sciences, Zeppelin University Friedrichshafen, Friedrichshafen, Germany

ABSTRACT

This article reports on the development and validation of the Instrumental Practice Goal Realization Inventory (IPGRI) designed to assess the process of self-directed study, beginning with setting the intention to practice and ending with the evaluation of one's performance. This new tool is based on the theoretical model of action phases. The IPGRI consists of 12 items and four scales delineating the phases of goal realization: Practice Intention, Planning, Acting, and Evaluating Practice. We describe the conceptualization and development of the measure. The evaluation of its psychometric properties is based on two samples: 171 piano students and 235 students learning various instruments. They participated in three measurements, each spaced two weeks apart. Exploratory and confirmatory factor analyses performed on the two samples indicate that the four-factor structure of the IPGRI is a good representation of the realization of instrumental practice goals. A high reliability (internal consistency and test-retest stability) of the scales was observed. We also provide evidence for convergent, discriminant, construct, and known-group validity. The IPGRI is a short and multidimensional measure with good psychometric properties. Therefore, it may be used both in research and practice in the psychology of music and music education.

ARTICLE HISTORY
Received 3 June 2023
Accepted 1 June 2024

KEYWORDS
Musical practice; goal realisation; model of action phases; measurement

Introduction

Practice is a central aspect of every musician's development, regardless of their musical achievement levels (Jørgensen and Hallam 2016; Sloboda et al. 1996). Managing the process of learning a musical instrument is demanding; musicians engage in effortful, long-lasting, and monotonous exercises (Ericsson, Krampe, and Tesch-Roemer 1993), carefully monitoring their performance to improve it, and shielding their practicing from competing goals. Thus, it is important for both conducting relevant research and teaching an instrument to have access to a reliable and valid method of capturing key processes related to learning a musical instrument. This study aims to develop and validate the Instrumental Practice Goal Realization Inventory (IPGRI). This new measure is based on the Rubicon model of action phases, a psychological theory that pertains to both goal setting and goal implementation (Gollwitzer 1990).

CONTACT Zofia Mazur-Socha zsmazur@gmail.com
Supplemental data for this article can be accessed online at https://doi.org/10.1080/14613808.2024.2365719.

Since the 1990s, learning a musical instrument has frequently been conceptualised using a deliberate practice framework (Ericsson, Krampe, and Tesch-Roemer 1993), assuming that expert performance is a consequence of prolonged activities undertaken with the explicit purpose of acquiring or improving skills. Numerous studies have confirmed that a vast amount of practice is required to achieve a mastery level of skills, and that elite level performers accumulate over 10,000 h of training (Ericsson, Krampe, and Tesch-Roemer 1993; see: Jørgensen and Hallam 2016 for a review). Although this framework is valuable in explaining the highest levels of musical expertise, it may not be sufficient to thoroughly understand the typical development of well-performing musicians (McPherson and Renwick 2011). For example, research presented that strategies during practice can be more determinative of performance quality than the time of practicing (Duke, Simmons, and Cash 2009).

Many music studies utilised the social cognitive perspective and self-regulated learning (SRL) theory as a means of understanding music performance learning (McPherson and Renwick 2011; McPherson et al. 2017a, 2017b). This approach highlights that social experiences, personal beliefs and learning opportunities interact to promote positive musical performance. SRL, in particular, focuses on processes in the self-enhancing cycles of learning, namely on the forethought, actual performance, and self-reflection phases. Cognitive processes and personal beliefs before the performance (i.e. the forethought phase) are expected to enhance subsequent learning. These are followed by processes during learning (i.e. the performance phase) that enhance concentration and performance. The individuals' responses to a learning episode (i.e. the self-reflection phase) lead to forethought, shaping subsequent efforts.

While SRL provides a comprehensive and widely appreciated explanation of musical skill acquisition (see: Varela, Abrami, and Upitis 2016, for a review), understanding the wilful practice of an instrument can be enhanced if one makes a distinction between goal setting and goal realisation (Keller, Gollwitzer, and Sheeran 2020). Such a distinction is important, when one wants to develop interventions for behavioural change. To change the behaviour of individuals who have not yet decided to realise a goal, one needs to focus on the 'why' aspect – motivation, rationale, and reasons. In turn, changing the behaviour of people who have already decided to pursue the goal in question requires focusing on the 'how' aspect – strategies that help attain the goal. Additionally, highlighting the importance of forming implementation intentions (i.e. 'if-then' plans that specify when, where, and how one wants to realise one's goals) may further help design tools to facilitate musical performance. Psychological research has demonstrated that the making of such plans increases the effectiveness and success of goal striving in many domains (e.g. academic achievement, health, and interpersonal; Gollwitzer and Sheeran 2006).

Although research on skill acquisition in music performance is constantly advancing, measurement tools used in this area still have several limitations. Many of them do not consider both behaviour and cognition (McPherson et al. 2017a, 2017b). Researchers often apply case-study designs, interviews, and observation methods (Varela, Abrami, and Upitis 2016), which permit tracking the development of single participants in great detail, but are less useful in studying differences between groups of people. Some questionnaires do not have the necessary amount of items per subscale (e.g. only two items; Hamann et al. 1998) or have reliability coefficient below .60 (Lim and Lippman 1991). Furthermore, currently used self-report measures of practice are quite extensive, as they aim to capture many aspects of motivation or behaviour (e.g. consist of 30–50 items; Hallam et al. 2012). The development of a measure brief enough to limit exhaustion effects, with a clear theoretical basis and good psychometric properties, would therefore benefit research and practice in educational settings. Such a tool could be applied by music teachers in school, community music, and private teaching contexts in screening and progress-monitoring efforts, informing further interventions to enhance students' motivation to practice their instruments. We propose that the psychological model of action phases may provide the needed theoretical grounds for a new measure of goal realisation regarding instrumental practice.

The Rubicon model of action phases

In the late 1980s, Heckhausen and Gollwitzer (1987) presented a theoretical model of goal pursuit, namely the Rubicon model of action phases, which provides a comprehensive temporal perspective with four distinct phases: the predecisional phase, the postdecisional or preactional phase, the actional phase, and the postactional phase. This model was confirmed by much empirical research (e.g. Brandstätter et al. 2015; Büttner et al. 2014; Gollwitzer and Sheeran 2006; Keller, Gollwitzer, and Sheeran 2020; Parks-Stamm, Gollwitzer, and Oettingen 2010; Webb, Christian, and Armitage 2007) and used in explanations in different sub-disciplines of psychology (e.g. motivation, emotion, cognition, learning and performance; Silva et al. 2018; Webb et al. 2012). However, there was no research that applied the model of action phases to the context of music learning.

The predecisional phase is a period of deliberation over wishes and desires; it helps choosing which of them one wants to realise (Gollwitzer 1990). Having many wishes and desires people are forced to be selective. The predecisional, still motivational phase ends with the transformation of a certain wish into a goal intention–the decision to pursue the wish as a goal. The person has now moved from motivation to volition, which is accompanied by feelings of commitment and a sense of obligation. Goal intentions take the following form: 'I intend to achieve X', and the 'X' delineates the specific behaviour or outcome the person wants to achieve (Gollwitzer 1990; Gollwitzer 1999).

After formulating a goal intention, people are confronted with the question of how to promote the achievement of the chosen goal as they move to the preactional phase. Usually, there are many different activities that people have to complete to attain a goal. Moreover, many goals are too complicated or distant to be realised in the here and now with one single act. For these reasons, people engage in planning. In the preactional phase, the task is to plan out in advance one's goal-directed actions, specifying when, where, how and how long to act to achieve the desired end state. In other words, one needs to formulate 'implementation intentions' (Gollwitzer 1999).

The actional phase starts with action initiation and is characterised by striving toward goal attainment (Gollwitzer 1990). Volitional strength (i.e. the person's goal commitment) determines whether people start to act and influences how much effort they put into their goal-directed actions. The obstacles experienced on the way to the goal can make one even more committed to the goal and put more effort into the activity. Thus, the effort may increase over time as people strive to maintain their goal commitment.

The final, postactional phase is associated with answering the question of whether striving for the goal is having a satisfying outcome (Gollwitzer 1990, 1999). During this phase, the individual can stop acting and wait for the consequences of goal striving to see if the aspired-to end state actually occurs. Additionally, one assesses whether the value of this outcome is consistent with the desirability of goal attainment that was anticipated.

The model of action phases can be applied to explain goal realisation with respect to instrumental practice. In this context, we define: (1) *practice intention* (IN) as the extent to which a student intends to put effort into the practice of one musical piece, and learn it systematically and conscientiously (this definition is in line with the theory of deliberate practice, which states that attempts at learning a musical instrument should be long-lasting, effortful, and repetitive; Ericsson, Krampe, and Tesch-Roemer 1993), (2) *practice planning* (PL) as the quality of the student's plans to practice a piece, which entails the specific content of practice, the practice strategies, as well as the specification of when, where, and how the student is going to practice, (3) *action phase of practice* (AC) as the intensity of the systematic effort to learn the musical piece and the diligence in practicing it, and (4) *performance evaluation* (EV) as the self-evaluation of the effects of practice with respect to the public performance of the piece.

The SRL theory describes constructs such as goal setting and time management as parts of the forethought phase of learning (Hatfield, Halvari, and Lemyre 2017). In contrast, incorporating the model of action phases, we propose to analyze IN and PL as separate and consecutive phases in the learning process. While SRL focuses on general tendencies to set goals (e.g. using questions such as

'I always set concrete long-term goals for myself'), we aim to measure intentions to practice a specific musical piece in the near future. Additionally, the SRL suggests that the performance phase of learning includes using psychological skills such as self-observation, arousal regulation, imagery, concentration, and self-control. Here we propose to limit the number of constructs evaluated, and to measure AC as the degree of effort put into practice.

For many instrumental or vocal music students, musical practice takes the form of self-directed study related to 'homework' assigned by the teacher (e.g. repeated performance of highly structured tasks). The home practice results are usually systematically verified at the individual lessons with the teacher, at the formal end-term exam, or at a music competition. Preparation for public performances, such as the end-term musical performance examination, can cause students to prioritise their practice (Hallam 2016) and mobilise energy to plan out their learning time. When students sign up for a competition, this requires increased practice time and a stronger dedication to learning.

Current research

Given the limitations of currently available measures of instrumental practice, the present article describes the construction and validation of a new assessment tool, the Instrumental Practice Goal Realization Inventory (IPGRI). The IPGRI targets the process of goal realisation (Gollwitzer and Moskowitz 1996; Heckhausen and Gollwitzer 1987; Keller, Gollwitzer, and Sheeran 2020) in practicing a specific musical piece (e.g. piano concerto no. 1 in E minor, op. 11 by Frederic Chopin). This new tool distinguishes four aspects related to practicing a musical piece: IN, PL, AC, and EV.

Experts recommended using different samples of participants in the development process of a new measure (Boateng et al. 2018). While the data from one sample would provide the basis for item selection, the next step in the scale development would be to test whether the structure extracted in the first sample is still adequate for another sample. Using data from a single sample for both exploratory and confirmatory factor analyses carries the risk that the model can be over fitted (Fokkema and Greiff 2017). To prevent this risk, we repeated the confirmatory factor analysis on data from two independent samples of young musicians. The first sample (Study 1) consisted of pianists, while the second sample (Study 2) consisted of musicians playing various instruments. This also allowed for a broader generalisation of the results obtained.

We designed these two studies to provide convincing evidence of the reliability and validity of the new instrument. To begin, Study 1 focused on item selection and an initial analysis of reliability and validity; a simple cross-sectional design was used for this study. Study 2 had a longitudinal design, as we aimed to determine the new measure's test-retest reliability. Moreover, Study 2 also included the end-term instrumental exam and the grade obtained on the exam, thus providing further data on the validity of the IPGRI.

Study 1: scale development

The primary aim of this study was to develop and select the items and to explore the factor structure of the IPGRI. Another aim was to indicate the reliability of the scale and preliminary validity using confirmatory factor analysis and analysis of convergent and discriminant validity.

Method

Participants

We calculated the minimum number of participants, by considering that there should be at least five times as many observations as the number of items to be analyzed in the factor analysis (Hair et al.

2009). Since the initial number of items in the IPGRI was 21, the minimum number of participants for the current analysis was considered to be 105.

One hundred and seventy one (63.2% female) piano students from 22 Polish music schools of second grade[1], took part in this study. The participants age ranged from 13 to 22 years (M = 16.83, SD = 1.51), and their experience in learning the piano ranged from 4 to 15 years (M = 9.37, SD = 2.24).

Procedure

The researchers selected second-grade music schools from different Polish cities. They contacted the schools' administration via email and phone to obtain consent for distribution and collection of a questionnaire by the school counsellors. Participation was voluntary, and the students received no compensation or course credit for it. The data were collected anonymously using paper-and-pencil forms. We collected the data shortly before the musical performance examination, a period of intensive instrumental learning, to make it easier for students to answer questions about practicing.

Measure – initial items generation

We developed a pool of items that corresponded to each phase of goal realisation (Gollwitzer 1990; 1999) in musical practice, based on psychological theory and using the expertise of the first author who is a musician. As both quantity and quality of practice are important in learning an instrument or voice (Jørgensen 2008; McPherson and Renwick 2011), our items address the regularity of learning but also its diligence. To capture the sense of obligation that is characteristic of goal intentions, we used statements beginning with 'I will' or 'I intend'. Regarding the planning aspect of practice, we aimed to address basic components of implementation intentions, by specifying where, when and how one plans to work on a chosen goal. Additionally, we added a question about 'what to practice', as there are various things in musical pieces that can be practiced, for example, rhythm, intonation, and dynamics. Considering evaluation, we aimed to measure students' self-evaluations of their play, when they perform in front of others, such as teachers or colleagues. Standards that students hold regarding their assessment of performance are commonly drawn from their experiences when listening to professional musicians or their colleagues.

Three experts in psychometrics (one of them having experience of learning a musical instrument) provided feedback on whether the items are easy to understand. We used this feedback and only included items that were marked as clearly formulated. The pool of 21 items presented to participants included five items for IN, six for PL, six for AC, and four for EV.

Participants were instructed to think of and write down one musical piece (e.g. etude, sonata) that they had been practicing for at least one month in the past. The piece should also be one that they are working on for their instrument class at the music school. Next, they answered the items concerning this musical piece (see Appendix A) on a 5-point Likert scale ranging from 1 = *strongly disagree* to 5 = *strongly agree*.

Data analysis

We used principal component analysis (PCA) with Oblimin rotation to select items for the IPGRI and explore its structure, while confirmatory factor analysis (CFA) with maximum likelihood estimation was applied to test whether the selected items fit well with two alternative theoretical models, a model containing only one factor versus a model containing four factors.

Psychometric analyses of the measure entailed the inspection of descriptive statistics for items and for scales, as well as inter-item and inter-factor correlations. We estimated reliability with Cronbach's α coefficient and composite reliability (CR) index. To assess convergent validity, we used information regarding factor loadings, the values of the average variance extracted (AVE) and CR. To evaluate discriminant validity, we compared the square root of the AVE value of each scale with the value of its correlations with other scales (Hair et al. 2009). Pearson correlations between the IPGRI scales and the final exam grade offer additional evidence of construct validity. We used one-way between-subjects analysis of variance (ANOVA) to establish known-group validity.

Results and discussion

Item selection and PCA

We applied PCA to examine factor loadings of items created for the particular scales. The analysis also allowed for exploring the structure of the IPGRI to verify whether it corresponds with our theoretical assumptions. We used the Oblimin rotation to allow correlations between factors.

Regarding the initial item pool, the Kaiser-Meyer-Olkin test of sampling adequacy (KMO) was .84, indicating a good degree of non-unique covariance among the items (Kaiser 1974). Bartlett's test of sphericity (BTS) was statistically significant ($\chi^2 = 1621.998$, $df = 210$, $p < .001$), indicating that singularity is not an issue. The determinant was close to 0, showing that multicollinearity is not an issue either. These indices suggest that the item pool was suitable for further analysis. The results of the initial PCA yielded a five-factor solution with eigenvalues > 1, which accounted for 65.56% of the variance (Appendix A). Inspection of the factor-item content of the pattern matrix indicated that the fifth factor consisted of only two items. We thought that each scale should consist of at least three items; thus, we eliminated these two items from further analysis. As a result, we obtained a four-factor structure, representing items designed for the IN, PL, AC, and EV factors. We then selected the three items with the highest factor loadings for each factor. The obtained solution (Table 1) accounted for 72.96% of the variance and was characterised by adequate sampling (KMO = .78), no singularity (BTS $\chi^2 = 816.211$, $df = 66$, $p < .001$), and no multicollinearity (Determinant = .007). Items loaded strongly onto the scales they were expected to load on and did not load onto other factors. Thus, this four-factor solution was identified as the preferred latent structure and measurement model for the 12-item IPGRI.

IPGRI item properties

Descriptive statistics and correlations between the 12 items are presented in Table 2. The mean scores of the items ranged from 2.15 to 4.36 on a five-point scale, showing no extremely

Table 1. PCA Pattern Matrix for IPGRI Scales.

Goal realisation aspect	Item no.	Item content	Factor loadings (λ)			
			ξ1 IN	ξ 2 PL	ξ 3 AC	ξ 4 EV
IN	IN1	I'm going to practice this piece diligently	.919	.034	−.006	.025
	IN2	I will make an effort to practice this piece regularly	.858	−.010	−.079	−.026
	IN3	I intend to work systematically on this piece	.857	−.067	.039	.007
PL	PL1	I have planned how I will practice this piece	−.039	.889	−.030	−.078
	PL2	I have made plans for what I specifically want to practice in this piece	.049	.866	−.028	−.060
	PL3	My plans for when and where I'm going to practice this piece are clear	.052	.736	.028	.154
AC	AC1	I've been making an effort to learn to play this piece	.061	.073	.834	−.108
	AC2	I've been practicing this piece thoroughly	−.028	−.073	.821	.114
	AC3	I've been systematically practicing this piece	.024	−.083	.741	.066
EV	EV1	The quality of my performance was similar to the professional quality known to me from recordings or concerts	−.043	−.069	.102	.884
	EV2	The quality of my performance was good compared to that of my peers	.124	.063	.018	.860
	EV3	I am satisfied with my performance	−.083	−.007	−.218	.715

Note. IN = Practice Intention Scale; PL = Practice Planning Scale; AC = Acting on Practice Scale; EV = Evaluating Performance Scale.

Table 2. Descriptive Statistics and Correlations of Items Representing the Four IPGRI Scales.

			Correlation											
Item	*M*	*SD*	IN1	IN2	IN3	PL1	PL2	PL3	AC1	AC2	AC3	EV1	EV2	EV3
IN1	4.22	0.89	1											
IN2	4.11	0.95	.74***	1										
IN3	4.36	0.92	.69***	.64***	1									
PL1	3.43	1.19	.27***	.30***	.30***	1								
PL2	3.67	1.14	.31***	.35***	.37***	.66***	1							
PL3	3.11	1.20	.30***	.29***	.31***	.50***	.54***	1						
AC1	4.14	0.88	.29***	.32***	.23**	.18*	.23**	.15	1					
AC2	3.78	0.86	.30***	.30***	.23**	.30***	.31***	.27***	.54***	1				
AC3	3.73	1.05	.23**	.33***	.31***	.24**	.28***	.33***	.43***	.59***	1			
EV1	2.15	1.07	.02	.02	.01	.13	.17*	.28***	.06	.23**	.16*	11		
EV2	3.21	1.08	.11	.11	.12	.13	.15*	.19*	.14	.26**	.24**	.63***	1	
EV3	3.07	1.10	.09	.03	.06	.13	.14	.26**	.16*	.39***	.30***	.52***	.48***	1

Note. IN = Practice Intention Scale; PL = Practice Planning Scale; AC = Acting on Practice Scale; EV = Evaluating Performance Scale; *M* = mean; *SD* = standard deviation; *$p < .05$; **$p < .01$; ***$p < .001$.

low or high mean values. There were moderate to strong positive and statistically significant correlations between items belonging to the same scale (*r* ranged from .43 to .74). Items belonging to different scales showed very weak to weak positive correlations. No items correlated higher than .80, indicating low multicollinearity between items (Tabachnick and Fidell 2013).

Confirmatory factor analysis

To initially confirm the postulated structure of the IPGRI, we conducted a CFA, comparing two alternative solutions. Model 1 considered four correlated factors, namely IN, PL, AC, and EV. Model 2 considered one single factor of goal realisation. No cross-loadings or correlated error terms were included in either of the models.

We examined commonly used goodness-of-fit indices with their thresholds: the comparative fit index (CFI; ≥ .95), root mean square error of approximation (RMSEA; ≤ .06), standardised root mean residual (SRMR; ≤ .08), and the Akaike information criterion (AIC; the lower, the better the fit) served as criteria for an acceptable model fit (Weston and Gore 2006). We compared alternative models using the χ^2 difference test ($\Delta\chi^2$) and the difference in CFI (ΔCFI), assuming that an absolute difference in CFI that is higher than .01 (ΔCFI > .01) would indicate a significant difference in model fit (Cheung and Rensvold 2002).

Only Model 1 obtained an acceptable fit to the data (Table 3). Tests of differences confirm that Model 1 fits the data significantly better than Model 2 (*p*-values for $\Delta\chi^2$ exceeded .05 and ΔCFI > .01).

Reliability, descriptive statistics, and correlations between scales

As a single estimate may understate reliability, we applied two different reliability coefficients – Cronbach's α and CR (based on Model 1 CFA results). For either reliability estimate, a value of .70 or higher suggests good reliability (Hair et al. 2009). In interpreting reliability indicators, we considered that internal consistency estimates are affected by the number of items (Cortina 1993); our scales consisted of only three items.

Cronbach's α ranges from .76 to .87, and CR from .75 to .89 (see Appendix B for details). Both estimates meet the recommended threshold of .70, indicating adequate to high internal consistency.

The descriptive statistics (Appendix B) showed that most participants had scores above the midpoint of the scales, which is 3 (as answers range from 1 to 5), except scores for EV (mean was slightly below 3). Inspection of the standard deviations indicated that the scores were similarly dispersed. There were no extremely low mean values. Analyses of skewness (*SKE*) and kurtosis (*K*) indicated that there were no severe departures from normality ($SKE < |2|$, $K < |7|$; West, Finch, and Curran 1995).

Cross-sectional correlations between scales (Appendix C) were positive and mostly statistically significant, ranging from .31 to .41, except the correlation between IN and EV, which was not statistically significant.

Table 3. Goodness-of-fit Indices of Alternative Measurement Models of IPGRI.

								Model comparison				
Model	χ^2	*df*	*p*	RMSEA	SRMR	CFI	AIC	Models	$\Delta\chi^2$	Δ*df*	*p*	ΔCFI
Study 1												
M1. Four correlated factors	59,463	48	0.124	0.037	0.049	0.985	143,463					
M2. Single factor	429,566	54	0.001	0.202	0.152	0.524	501,566	M2 vs M1	370.103	6	0.001	0.461
Study 2, Time 1												
M1. Four correlated factors	90,702	48	0.001	0.062	0.045	0.964	174,702					
M2. Single factor	577,696	54	0.001	0.204	0.145	0.558	649.696	M2 vs M1	486.994	6	0.001	0.406

Note. *df* = degrees of freedom; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; CFI = comparative fit index; AIC = the Akaike Information Criterion.

Validity

To examine the convergent validity of scales, we ensured that the following requirements were met: (1) all factor loadings were statistically significant; (2) AVE values for each construct exceeded .50; and (3) CR values exceeded .60 (Hair et al. 2009).

All of these requirements were met, confirming the convergent validity of all scales. CFA results indicated that factor loadings in Model 1 were significant: for IN they were between .86 and .92, for PL between .74 and .89, for AC between .74 and .83, and for EV between .72 and .88. Appendix B shows that all AVE values were above .50 and all CR indicators were above .60.

To assess the discriminant validity of the IPGRI scales, we compared the square root of AVE for each scale with its correlations with the other scales (Hair et al. 2009). If the square roots of AVE are higher than the correlations, the scales represent distinct constructs. The square roots of AVE were .83 for IN, .76 for PL, .73 for AC, and .74 for EV. These values exceeded the values of the correlations between scales (see Appendix B and Appendix C), indicating that discriminant validity was achieved.

Study 2: reliability and validity of the IPGRI

The aim of Study 2 was to further examine the reliability and validity of our new measure. The longitudinal design of Study 2 allowed us to examine reliability more critically, as it included three measurement points separated by two weeks. The stability of the IPGRI could thus be tested over a shorter and longer time period (in two weeks between the first and second measurement and four weeks between the first and third measurement). We also examined the IPGRI among musicians playing different musical instruments to explore whether the measure is also applicable to a more diverse group of musicians.

We hypothesised that there should be a positive relationship between practicing before a music performance examination and the grade obtained in the exam, as the positive relationship between undertaking deliberate practice and musical achievements is well established in the literature (Ericsson, Krampe, and Tesch-Roemer 1993; Platz et al. 2014). We also expected a positive correlation between the self-evaluation of performance and the external evaluation – the grade obtained in a music performance exam.

To establish known-group validity, we hypothesised that students who had recently passed a music performance examination would score higher on AC and EV than students who had a relatively long time left until the exam. According to Lewin, what motivates people to perform an activity is also the psychological distance to the relevant critical event; the distance can refer to physical space but also time (Lewin, 1933 as cited in Deckers 2018). Thus, the further away in time the exam, the weaker the motivational force that induces students to practice their instruments. When students see the exam approaching, they should start to devote more time and effort to practicing, which may ultimately lead to better performance.

Method

Participants

We estimated the minimum sample size for Study 2 following the same procedures as in Study 1, recruiting at least five times as many participants as the number of variables to be analyzed in the factor analysis (Hair et al. 2009). Taking into account the 12 IPGRI items and expecting a typical attrition rate during the three study waves, we planned to recruit around 200 participants.

The sample of participants at the first measurement time (T1) was composed of 235 (68.9% female) Polish students learning keyboards ($n = 83$, 35.3%), strings ($n = 58$, 24.7%), woodwinds ($n = 53$, 22.6%), plucked strings ($n = 15$, 6.4%), brass ($n = 14$, 6%), percussion ($n = 6$, 2.6%), and vocal ($n = 6$, 2.6%). They attended 12 different second-grade music schools, ranging in age from 13 to 28 years ($M = 17.42$, $SD = 2.33$), and learned instruments for 1 to 18 years ($M = 7.77$, $SD = 2.71$). At T1, most of the students were still preparing for their music performance exam (99%[2]). The second measurement (T2) was conducted two weeks after T1 and comprised 109 participants, with the majority still waiting for their exam (76%). The third measurement (T3) was four weeks after T1, with 65 students participating. By that time, most of them had already completed their music performance exam (86%).

Procedure

To recruit the participants, we have obtained a full list of second-grade music schools in Poland and contacted the administration of each school via email (sending 124 emails) to request consent to distribute a questionnaire through school counsellors. Fourteen school counsellors agreed to participate in the project, of whom 12 have arranged meetings with students. Prior to the study, the participants were told that among those who complete all three questionnaires, we will draw 10 prizes in the amount of approximately $14. Students received links to three online surveys at two-week intervals via the Remind: School Communication mobile application.

Measures

IPGRI

Participants completed the IPGRI developed in Study 1 (Appendix D). The instructions were slightly modified to align them with the longitudinal nature of the study – the participants were asked to remember the music piece they mentioned at the beginning of the study and to refer to the same piece throughout the study.

Time until the exam

The students were asked to write down the date of their externally graded music exam. We calculated the number of days from questionnaire completion to the exam. The scores at T1 ranged from −15 to 76 days ($M = 22.20$, $SD = 10.01$), at T2 from −12 to 31 days ($M = 6.88$, $SD = 7.95$), and at T3 from −29 to 10 days ($M = -8.10$, $SD = 8.18$).

Musical performance examination grade

At T3, we asked the students to report the grades they received on their end-term musical performance examination. The exam consisted of performing several music pieces, including the students had indicated in the questionnaire. The musical performances were graded by a team of instrumental teachers from the students' schools, as a part of their routine end-term activity. The grading scale ranged from 1 = *very poor* to 6 = *excellent*. The examination grades ranged from 3 to 6 ($M = 4.64$, $SD = 0.63$).

Data analysis

We applied CFA to further confirm that the data fits well with the theoretical model of the IPGRI. Interfactor correlations, reliability, convergent and discriminant validity were examined analogously to Study 1. Pearson correlations between the IPGRI scales and the end-term musical performance examination grade provided further evidence of construct validity. One-way between-subjects ANOVA was used to establish known-group validity examining differences in AC and EV scores between students who (a) were shortly before their exam, (b) had a relatively long time left until their exam, (c) had already taken their exam.

Results and discussion

Confirmatory factor analysis

We performed CFA on the data from the first measurement. Similar to Study 1, we compared two alternative solutions – Model 1 considered four correlated factors and Model 2 considered a single factor. Only Model 1 obtained an acceptable fit to the data (Table 3). Tests of differences confirmed that Model 1 fits the data significantly better than Model 2 (p-values for $\Delta\chi^2$ exceeded .05 and ΔCFI > .01).

Reliability and descriptive statistics for scales

Cronbach's α ranged from .76 to .96, and CR from .70 to .93 (Appendix B), indicating adequate-to-high internal consistency. The IN scale obtained the highest reliability ($M_\alpha = .90$, $M_{CR} = .90$), while the PL and EV scales the lowest reliability (for PL $M_\alpha = .80$, $M_{CR} = .73$, for EV $M_\alpha = .79$, $M_{CR} = .77$).

To evaluate test-retest reliability after the two and four-week interval, the Pearson correlation coefficients were calculated for each scale (Appendix C). Test-retest correlations over the two-week interval (between T1 and T2) were all statistically significant at a level of $p < .001$ and ranged from .54 for IN to .68 for EV, indicating moderate to high stability of the IPGRI scales. Over the four-week interval (between T1 and T3) all correlations were statistically significant at least at the level of $p < .05$, except the correlation of the PL scale. The magnitude of these correlations ranged from .20 for PL to .52 for EV. These results demonstrate weak to moderate stability of the scales over four weeks, with the exception of the PL scale, which was not very stable over this time period. For both time intervals, the EV scores were the most stable.

Inspection of the descriptive statistics (Appendix B) demonstrated that at all three measurement times, the average scores of the scales were above the midpoint, which is 3, except the mean scores for EV at T1, which were slightly below 3. Standard deviations indicated that the results were similarly dispersed; however, there was more variability in IN, PL and AC at T3 when compared to the other two measurements (T1 and T2). The skewness and kurtosis scores did not show any major deviations from the normal distribution.

Appendix C presents the cross-sectional and longitudinal correlations between scales. Most cross-sectional correlations (at T1, T2 and T3) were positive and statistically significant (ranging from .19 to .81; correlations between IN and EV were not statistically significant at all three measurement times, and correlations between EV and PL or AC were not statistically significant at T3). Analysis of the longitudinal correlations in the two-week intervals (between T1 and T2 and between T2 and T3) indicates that most of them were statistically significant, and showed weak to moderate positive relationships between the IPGRI scales (ranging from .26 to .68). Over a four-week interval (between T1 and T3), positive weak to moderate correlations between scales were found (ranging from .25 to .52); however, most of the correlations were not statistically significant.

Validity

In Study 2, all factor loadings in Model 1 were statistically significant, all AVEs were above .50, and all CR indicators were above .60 (Appendix B). This indicates that the convergent validity of the scales was established.

The square roots of the AVE at all measurement times ranged from .83 to .94 for IN, from .71 to .84 for PL, from .76 to .89 for AC, and from .76 to .78 for EV. These exceeded the correlations between the scales (Appendix B and Appendix C), indicating that discriminant validity was also established.

To provide further evidence for the construct validity of the IPGRI scales, we examined how the IPGRI scales correlate with the end-term exam grade received in the music performance examination (Appendix C). As expected, AC before the exam, at both T1 and T2 correlated statistically significantly and positively with the grade at T3 (.28 and .30 respectively), and EV also correlated positively with the grade at T3 (.29 for T1, .36 for T2, and .43 for T3). The grade achieved was not statistically significantly correlated with IN and PL.

We used one-way between-subjects ANOVA to examine differences in AC and EV scores between students from the T2 sample who had already passed their performance music performance examination ($n = 23$), who had between 1 and 8 days left until their exam ($n = 37$), and who had between 9 and 31 days left until their exam ($n = 37$). The three groups had equal numbers of participants ($\chi^2_{(2, 97)} = 4.04$, $p = .133$). The results showed statistically significant differences between these groups in AC ($F_{(2, 94)} = 4.40$, $p < .05$; $\eta_p^2 = .085$) and EV ($F_{(2, 94)} = 3.15$, $p < .05$; $\eta_p^2 = .063$). Bonferroni *post hoc* paired comparisons indicated that students who had already completed their exam had recently put more effort (AC) into practice ($M = 4.22$, $SD = 0.69$) than students who had more than 8 days left until their exam ($M = 3.65$, $SD = 0.70$, $p = .016$). Musicians who had taken their exam already tended to evaluate their performance (EV) as more satisfying ($M = 3.49$, $SD = 0.71$) than musicians who had more than 8 days left to prepare for their exam ($M = 2.86$, $SD = 0.89$, $p = .042$). These scores are consistent with our expectations and confirm the known group validity of the IPGRI scales. There were no statistically significant differences between the three groups for IN ($F_{(2, 94)} = 0.15$, $p = .864$; $\eta_p^2 = .003$) and PL ($F_{(2, 94)} = 0.59$, $p = .554$; $\eta_p^2 = .012$).

General discussion

Two studies allowed for the development and validation of the IPGRI (Appendix D). Exploratory PCA conducted on the initial item pool resulted in the selection of 12 items with the highest factor loadings, representing four dimensions of goal realisation in musical practice: IN, PL, AC, and EV. Examination of the goodness of fit of two alternative models indicated that the four-factor model adequately represents the structure of the data collected in two samples. In contrast, the single-factor model does not fulfil the requirements for model fit. Thus, the CFA supported evidence that the IPGRI captures the four dimensions mentioned above, based on the model of action phases (Gollwitzer 1990; Heckhausen and Gollwitzer 1987). The IPGRI scales show a high internal consistency in terms of Cronbach's α and CR indicators, even though the scales consist of a small number of items. Their scores are also stable at the two and, for the most part, at the four-week time interval.

Positive weak to moderate correlations between scales suggest that the IPGRI addresses one phenomenon, with interdependent, yet distinct aspects. Firstly, IN is connected with PL in the sense that the more students intend to practice, the more they plan their practice. This connection is evident in the samples of Studies 1 and 2, and students' intentions predict practice planning reported two and four weeks later. Second, correlations of both IN and PL with AC indicate that the more students intend and plan to practice an instrument, the more they practice regularly and with effort. Additionally, IN and PL predicted AC two weeks later. Third, PL as well as AC predicted EV measured two weeks later. The more students plan their practice, the more they find their recent performance satisfactory. The more they report that they have been practicing systematically and putting effort into their learning, the more they rate their recent performance as satisfactory. Fourth, there are also some 'reversed' relationships between students' AC at a given time and their IN and PL two weeks later. The more participants report their recent practice as systematic and effortful, the more they indicate intentions and plans for systematic and conscientious practice in the future. This is consistent with the model proposed by Zimmerman (2002) presenting learning as a self-regulation cycle. In that model, 'self-reflections from prior efforts to learn affect subsequent forethought processes' (68). For instance, self-dissatisfaction can lead to setting less ambitious goals and plans for learning. Thus, we can expect that students' reflections about the regularity and intensity of their practice will influence their motivation.

Both studies initially revealed that the IPGRI fulfils the requirements regarding the convergent and discriminant validity of its scales. Standardised factor loadings are statistically significant and high, the average percentage of variation explained (variance extracted) among the items of each construct is sufficient, and the internal consistency of all of the scales is acceptable. Moreover, the discriminant validity achieved indicates that the related concepts delineated in the goal realisation process as captured by the IPGRI scales, are indeed distinct and sufficiently different from each other.

We found positive and consistent relationships between the grades obtained on the musical performance exam and the AC and EV scores obtained two or four weeks prior to the exam, providing additional evidence for construct validity. Based on how systematically and diligently students practice and how they evaluate their performance, it is possible to predict the grades they will later receive on their final music performance exam. The more reliably and systematically they practice, the higher their later grades; also, the higher their grades, the higher their performance self-evaluation.

We also demonstrated known-group validity of the IPGRI: students who had recently passed their end-term musical performance exam reported practicing more systematically and thoroughly, and they were more satisfied with their public performance than students who had more than eight days left until their exam. The differences in AC and EV scores between these groups are consistent with the notion that a short psychological distance to a performance event leads people to engage more intensively in related activities (Lewin, 1933 as cited in Deckers 2018). Students who have more time left until their exam are less inclined to practice and, as a result, are less satisfied with their performance.

In summary, there is substantial evidence that the IPGRI is a reliable and valid instrument. Consistent with the Heckhausen and Gollwitzer model (Gollwitzer 1990; Heckhausen and Gollwitzer 1987; Keller, Gollwitzer, and Sheeran 2020), the phenomena of practice are accurately captured, taking into account the four distinct aspects of setting an intention, planning, acting, and evaluating performance.

Limitations and further research

We aimed to develop a simple and short questionnaire which captures the key aspects of goal realisation in musical self-directed study. Our questionnaire does not measure the desirability and feasibility of practice (Heckhausen and Gollwitzer 1987), and thus, does not focus on the deliberative processes that occur prior to setting a goal to practice. Future research may propose other measures that include items measuring such constructs (e.g. 'To what extent is practicing this piece important to you?').

We also need to consider the limitations of using retrospective self-judgements provided as indicators of the goal realisation aspects. Having initial evidence that IPGRI is reliable and its scales are short enough to be used multiple times, subsequent studies can explore its usefulness and applicability in intensive longitudinal diary studies. By measuring aspects of regular daily practice, we can more accurately capture the phenomenon of practicing an instrument.

Moreover, self-assessment of items requires self-reflection ability (which might be underdeveloped, especially in young students) and is associated with social desirability problems (Maruyama and Ryan 2014). Since practice time and quality might differ at different stages of education (Hallam et al. 2012; Jørgensen and Hallam 2016), we recommend that future studies also investigate age differences in practice intention, planning, acting, and evaluation. In addition, our samples consisted of musicians who had been learning their instruments for relatively long time: for nine years (in Study 1) and eight years (in Study 2) on average, which demands high motivation. It is possible that younger and less experienced students, who invested less time in learning, would have lower intentions to practice. This issue can be clarified in future studies.

The constructs measured in IPGRI are potentially related to existing models of self-regulated learning, deliberate practice, motivation, and self-efficacy in musical practice. Therefore, there is a need to assess convergent validity in future research by comparing the components included in the IPGRI with other similar constructs to further validate the instrument.

Practicing reported at the final measurement was not related to the musical performance exam grade, although practice reported at the previous measurements was positively correlated with the end-term exam grade obtained. The reason for this difference might be that the time interval from the music performance exam was different across measurements. At the first measurement, most students were awaiting their end-term exam, while during the third measurement, most students had already passed their exam. When students have an exam in the near future, they may prioritise their practice and learn more (Hallam 2016). After the external evaluation of performance, students' engagement in practice may vary across students to a larger extent. Some learners may take a break from practicing and start their vacations, while others may start preparing for upcoming concerts or change their musical pieces. Additionally, the size of the study sample at the third measurement was smaller than at the other measurements. In conclusion, in future studies on music learning, it seems worthwhile to control the time intervals until the external evaluations or public performances, as these may have a joint effect on practice. Finally, it would be valuable to start where we left off, and further examine the validity of the IPGRI. For instance, there is a good reason to analyze measurement invariance by testing whether the IPGRI model fits well to data from musicians playing different instruments. Examining the role of age, playing experience, gender, and personality traits in goal realisation would also add to the currently available data on the validity of the IPGRI.

Implications for theory and practice

Although currently used theories allow for understanding many aspects of musical instrument students' motivation and practice behaviour, considering other theoretical perspectives may expand our knowledge, filling specific research gaps. For example, considering the psychological model of action phases (Gollwitzer 1990; Heckhausen and Gollwitzer 1987) allows for tracking the role of planning in goal realisation in musical practice. So far, practice planning has not been included as a separate aspect of music instrumental learning, and the distinction between processes related to intention, planning, action, and evaluation has not been highlighted. Use of the IPGRI could guide the development of innovative relevant behaviour change interventions. Psychologists may seek ways to more effectively support students in their practice efforts, as many musicians are not sufficiently engaged in practice (Twarowska 2012). Teachers in both institutional and private settings, or even students themselves, can use this newly developed tool to diagnose whether students have not yet set the intention to practice and need support in forming a strong intention, or whether they rather have problems with planning and initiating their practice, as well as arriving at a proper evaluation of their practice. Acquiring such knowledge should enable teachers to provide their students with the necessary support in practicing their instruments.

The IPGRI was designed to address the unique challenges of music education research. This instrument is brief and, therefore, easy to use by researchers who study learning and performance in music students and by teachers. It diminishes the risk of distracting effects apparent in long questionnaires (Katz et al. 2017). The IPGRI can be used multiple times in a short time period to monitor changes in the students' practice behaviour.

The IPGRI was validated in two independent and relatively large samples; piano students (Study 1) and students learning different instruments (Study 2). This is in line with recommendations for studies aimed at developing psychometric tools (Boateng et al. 2018). It allows for cross-testing of the factorial structure and verification of the psychometric properties of the different subscales. With three measurement times performed in Study 2, we were able to test relationships between action phases over time, and to provide further evidence of test-retest reliability and validity.

Overall, our results indicate that the IPGRI has strong psychometric properties. Quality of practice and musical achievements are key issues in the psychology of music and music education. This new measure may support future research, while also being useful to music students and their teachers.

Notes

1. In Poland there are music schools at the primary level, called 'music schools of the first grade', and at the secondary level, called 'music schools of the second grade' (both levels usually last six years). Children, adolescents and adults of various ages can attend such schools, as some music education institutions operate independently of the compulsory education system.

2. The percentages of students who were before or after their exams are calculated based on data from 205, 97, and 58 participants, respectively, for the three measurement times. Only these students provided information about the exam date, the rest did not know the exact exam date.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Narodowe Centrum Nauki: [Grant Number 2018/29/N/HS6/02641].

Notes on contributors

Zofia Mazur-Socha is a doctor of psychology at The John Paul II Catholic University of Lublin, Institute of Psychology and professional musician. She is an author of measurement tools used in music research and education. Her research interests include music education, self-regulation, systematic reviews and meta-analyses.

Mariola Laguna is a professor of psychology at The John Paul II Catholic University of Lublin, Institute of Psychology and head of the Social Personality Psychology Lab. Her research interests include motivation, goal realization, psychometrics.

Peter Gollwitzer is a professor of psychology at New York University, University of Konstanz, and Leuphana University in Lueneburg and head of the Motivation Lab. He is an expert in research on motivation and goal realization. His research interests include motivation, goal realization, self-regulation.

References

Boateng, G. O., T. B. Neilands, E. A. Frongillo, H. R. Melgar-Quiñonez, and S. L. Young. 2018. "Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer." *Frontiers in Public Health* 6: 149. https://doi.org/10.3389/fpubh.2018.00149.

Brandstätter, V., L. Giesinger, V. Job, and E. Frank. 2015. "The Role of Deliberative Versus Implemental Mindsets in Time Prediction and Task Accomplishment." *Social Psychology* 46 (2): 104–115. https://doi.org/10.1027/1864-9335/a000231.

Büttner, O. B., F. Wieber, A. M. Schulz, U. C. Bayer, A. Florack, and P. M. Gollwitzer. 2014. "Visual Attention and Goal Pursuit: Deliberative and Implemental Mindsets Affect Breadth of Attention." *Personality & Social Psychology Bulletin* 40 (10): 1248–1259. https://doi.org/10.1177/0146167214539707.

Cheung, G. W., and R. B. Rensvold. 2002. "Evaluating Goodness-of-fit Indexes for Testing Measurement Invariance." *Structural Equation Modeling* 9 (2): 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Cortina, J. M. 1993. "What is Coefficient Alpha? An Examination of Theory and Applications." *Journal of Applied Psychology* 78 (1): 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Deckers, L. 2018. *Motivation: Biological, Psychological, and Environmental*. 5th ed.. New York: Routledge.

Duke, R. A., A. L. Simmons, and C. D. Cash. 2009. "It's not how Much; It's how: Characteristics of Practice Behavior and Retention of Performance Skills." *Journal of Research in Music Education* 56 (4): 310–321. https://doi.org/10.1177/0022429408328851.

Ericsson, K., R. Krampe, and C. Tesch-Roemer. 1993. "The Role of Deliberate Practice in the Acquisition of Expert Performance." *Psychological Review* 100 (3): 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Fokkema, M., and S. Greiff. 2017. "How Performing PCA and CFA on the Same Data Equals Trouble." *European Journal of Psychological Assessment* 33 (6): 399–402. https://doi.org/10.1027/1015-5759/a000460.

Gollwitzer, P. 1990. "Action Phases and Mind-Sets." In *The Handbook of Motivation and Cognition: Foundations of Social Behavior*, edited by E. T. Higgins, and R. M. Sorrentino, 53–92. Guilford Press. https://nyuscholars.nyu.edu/en/publications/action-phases-and-mind-sets.

Gollwitzer, P. M. 1999. "Implementation Intentions: Strong Effects of Simple Plans." *American Psychologist* 54 (7): 493–503. https://doi.org/10.1037/0003-066X.54.7.493.

Gollwitzer, P. M., and G. B. Moskowitz. 1996. "Goal Effects on Action and Cognition." In *Social Psychology: Handbook of Basic Principles*, edited by E. T. Higgins, and A. Kruglanski, 361–399. New York: Guilford .

Gollwitzer, P., and P. Sheeran. 2006. "Implementation Intentions and Goal Achievement: A Meta-Analysis of Effects and Processes." *Advances in Experimental Social Psychology* 38 (6): 69–119. https://doi.org/10.1016/S0065-2601(06)38002-1.

Hair, J. F., W. C. Black, B. J. Babin, and R. E. Anderson. 2009. *Multivariate Data Analysis*. 7th ed. London: Pearson.

Hallam, S. (2016). Motivation to Learn, edited by S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology,*. 2nd ed., 479–492. Oxford: Oxford University Press.
Hallam, S., T. Rinta, M. Varvarigou, A. Creech, I. Papageorgi, T. Gomes, and J. Lanipekun. 2012. "The Development of Practising Strategies in Young People." *Psychology of Music* 40 (5): 652–680. https://doi.org/10.1177/0305735612443868.
Hamann, K. L., K. V. Lucas, P. McAllister, and D. J. Teachout. 1998. "An Investigation Into the Factors Contributing to Individual Practice." *Journal of Band Research* 34 (1): 59–68.
Hatfield, J. L., H. Halvari, and P.-N. Lemyre. 2017. "Instrumental Practice in the Contemporary Music Academy: A Three-Phase Cycle of Self-Regulated Learning in Music Students." *Musicae Scientiae* 21 (3): 316–337. https://doi.org/10.1177/1029864916658342.
Heckhausen, H., and P. M. Gollwitzer. 1987. "Thought Contents and Cognitive Functioning in Motivational Versus Volitional States of Mind." *Motivation and Emotion* 11 (2): 101–120. https://doi.org/10.1007/BF00992338.
Jørgensen, H. 2008. "Instrumental Practice: Quality and Quantity." *The Finnish Journal of Music Education* 11 (1–2): 8–18.
Jørgensen, H., and S. Hallam. 2016. "Practicing." In *The Oxford Handbook of Music Psychology*, edited by Susan Hallam, I. Cross, and M. Thaut, 2nd ed., 449–462. Oxford: Oxford University Press.
Kaiser, H. F. 1974. "An Index of Factorial Simplicity." *Psychometrika* 39: 31–36.
Katz, B. A., N. Lustig, Y. Assis, and I. Yovel. 2017. "Measuring Regulation in the Here and now: The Development and Validation of the State Emotion Regulation Inventory (SERI)." *Psychological Assessment* 29 (10): 1235–1248. https://doi.org/10.1037/pas0000420.
Keller, L., P. M. Gollwitzer, and P. Sheeran. 2020. "Changing Behavior Using the Model of Action Phases." In *Handbook of Behavior Change*, edited by M. S. Hagger, L. S. Cameron, K. Hamilton, N. Hankonen, and T. Lintunen, 77–88. Cambridge: Cambridge University Press.
Lim, S., and L. G. Lippman. 1991. "Mental Practice and Memorization of Piano Music." *Journal of General Psychology* 118 (1): 21–30. https://doi.org/10.1080/00221309.1991.9711130
Maruyama, G., and C. S. Ryan. 2014. *Research Methods in Social Relations*. 8th ed.. New York: Wiley-Blackwell.
McPherson, G. E., P. Miksza, and P. Evans. 2017a. "Self-Regulated Learning in Music Practice and Performance." In *Handbook of Self-Regulation of Learning and Performance*, edited by D. H. Schunk, and J. A. Greene, 2nd ed., 181–193. New York: Routledge. https://doi.org/10.4324/9781315697048-12.
McPherson, G. E., M. S. Osborne, P. Evans, and P. Miksza. 2017b. "Applying Self-Regulated Learning Microanalysis to Study Musicians' Practice." *Psychology of Music* 47 (1): 18–32. https://doi.org/10.1177/0305735617731614.
McPherson, G. E., and J. M. Renwick. 2011. "Self-regulation and Mastery of Musical Skills." In *Handbook of Self-Regulation of Learning and Performance*, edited by B. Zimmerman, and D. H. Schunk, 234–248. New York: Routledge.
Parks-Stamm, E. J., P. M. Gollwitzer, and G. Oettingen. 2010. "Implementation Intentions and Test Anxiety: Shielding Academic Performance from Distraction." *Learning and Individual Differences* 20 (1): 30–33. https://doi.org/10.1016/j.lindif.2009.09.001.
Platz, F., R. Kopiez, A. C. Lehmann, and A. Wolf. 2014. "The Influence of Deliberate Practice on Musical Achievement: A Meta-Analysis." *Cognition* 5: Article 646. https://doi.org/10.3389/fpsyg.2014.00646.
Silva, M. A. V. da, T. M. São-João, V. C. Brizon, D. H. Franco, and F. L. Mialhe. 2018. "Impact of Implementation Intentions on Physical Activity Practice in Adults: A Systematic Review and Meta-Analysis of Randomized Clinical Trials." *PLoS One* 13 (11): e0206294. https://doi.org/10.1371/journal.pone.0206294.
Sloboda, J. A., J. W. Davidson, M. J. A. Howe, and D. G. Moore. 1996. "The Role of Practice in the Development of Performing Musicians." *British Journal of Psychology* 87 (2): 287–309. https://doi.org/10.1111/j.2044-8295.1996.tb02591.x
Tabachnick, B. G., and L. S. Fidell. 2013. *Using Multivariate Statistics*. 6th ed. Boston, MA: Pearson.
Twarowska, M. E. 2012. "Uczeń w systemie – Od rekrutacji do okresu poszkolnego [Student in the System: From Recruitment to After-School Period]." In *Raport o stanie szkolnictwa muzycznego I stopnia. Diagnozy, problemy, wnioski modelowe*[Report on the state of musical education of the first degree. Diagnoses, problems, model conclusions], edited by W. Jankowski, 107–134. Warszawa: Instytut Muzyki i Tańca.
Varela, W., P. C. Abrami, and R. Upitis. 2016. "Self-regulation and Music Learning: A Systematic Review." *Psychology of Music* 44 (1): 55–74. https://doi.org/10.1177/0305735614554639.
Webb, T. L., J. Christian, and Ch. J. Armitage. 2007. "Helping Students Turn up for Class: Does Personality Moderate the Effectiveness of an Implementation Intention Intervention?" *Learning and Individual Differences* 17 (4): 316–327. https://doi.org/10.1016/j.lindif.2007.03.001.
Webb, T. L., I. S. Gallo, E. Miles, P. M. Gollwitzer, and P. Sheeran. 2012. "Effective Regulation of Affect: An Action Control Perspective on Emotion Regulation." *European Review of Social Psychology* 23 (1): 143–186. https://doi.org/10.1080/10463283.2012.718134.
West, S. G., J. F. Finch, and P. J. Curran. 1995. "Structural Equation Models with Nonnormal Variables: Problems and Remedies." In *Structural Equation Modeling: Concepts, Issues and Applications*, edited by R. H. Hoyle, 56–75. Thousand Oaks: Sage.
Weston, R., and P. A. Gore. 2006. "A Brief Guide to Structural Equation Modeling." *Counseling Psychologist* 34 (5): 719–751. https://doi.org/10.1177/0011000006286345.
Zimmerman, B. J. 2002. "Becoming a Self-Regulated Learner: An Overview." *Theory Into Practice* 41 (2): 64–70. https://doi.org/10.1207/s15430421tip4102_2.