

**'How Would you Score Yourself?'**

**The Effect of Self-assessment Strategy Through Robots on Children's Motivation and Performance in Piano Practice**

Song, Heqiu; Tsiakas, Konstantinos; Ham, Jaap; Markopoulos, Panos; Barakova, Emilia I.

**DOI**

[10.1007/s12369-023-01080-3](https://doi.org/10.1007/s12369-023-01080-3)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

International Journal of Social Robotics

**Citation (APA)**

Song, H., Tsiakas, K., Ham, J., Markopoulos, P., & Barakova, E. I. (2023). 'How Would you Score Yourself?': The Effect of Self-assessment Strategy Through Robots on Children's Motivation and Performance in Piano Practice. *International Journal of Social Robotics*, 16(2), 327-340. <https://doi.org/10.1007/s12369-023-01080-3>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



# 'How Would you Score Yourself?': The Effect of Self-assessment Strategy Through Robots on Children's Motivation and Performance in Piano Practice

Heqiu Song<sup>1</sup> · Konstantinos Tsiakas<sup>2</sup> · Jaap Ham<sup>3</sup> · Panos Markopoulos<sup>1</sup> · Emilia I. Barakova<sup>1</sup>

Accepted: 15 November 2023  
© The Author(s) 2023

## Abstract

This research examines how to design social robots to support self-regulated learning skills for piano practice. More specifically, a social robot is used to provide feedback to children and initiate self-assessment. To assess the impact of this approach on children's motivation and performance, we conducted an experiment in a music school where 50 children practiced with both a self-assessment and a non-evaluative robot. Results showed that when the children interacted with the self-assessment robot they had higher motivation and better performance than when they interacted with the non-evaluative robot. Furthermore, interaction effects were found between the robot conditions, the children's learning stages, and their gender regarding their motivation and rhythm performance. Overall, the study demonstrates a positive influence of robot-initiated self-assessment on children's musical instrument practice and provided insights for personalized child-robot interaction design.

**Keywords** Social robot · Music instrument practice · Self-regulated learning · Self-assessment strategy · Motivation · Performance

## 1 Introduction

The increasing interest in and utilization of robotics in various aspects of daily life has led to their implementation in education settings. Research has demonstrated that the incorporation of robotics in education can improve cognitive and

learning outcomes [6]. While the specific roles and functions of robots in education may vary depending on the context, their primary purpose is to enhance these outcomes [15]. Previous studies have shown that social robots can serve as effective social support for children during music learning [72], and can also provide feedback to support the development of self-monitoring skills during educational activities [33].

In order to become an experienced musician, self-regulated learning skills and continuous practice are required [45]. Self-regulated learning helps to develop skills such as the ability to set goals, plan and organize tasks, monitor progress, and evaluate and adjust one's own learning strategies. When students are able to self-regulate their learning, they are more motivated, more independent, and better able to take ownership of their own learning. High-skilled and long-term players have developed self-awareness, the use of appropriate strategies, the ability to evaluate oneself, etc. [28, 41]. Arguably, encouraging self-regulation can help more players to become highly skilled.

Researchers that self-assessment is an important part of learning music and a self-regulated learning activity [3, 59]. Boud and Falchikov (1989) define self-assessment as: 'the involvement of learners in making judgments about their own

---

✉ Emilia I. Barakova  
e.i.barakova@tue.nl

Heqiu Song  
h.song@tue.nl

Konstantinos Tsiakas  
k.tsiakas@tudelft.nl

Jaap Ham  
j.r.c.ham@tue.nl

Panos Markopoulos  
p.markopoulos@tue.nl

<sup>1</sup> Department of Industrial Design, Eindhoven University of Technology, 5612 AP Eindhoven, The Netherlands

<sup>2</sup> Department of Industrial Design Engineering, Delft University of Technology, 2628 CE Delft, The Netherlands

<sup>3</sup> Department of Industrial Engineering and Innovation Science, Eindhoven University of Technology, 5612 AP Eindhoven, The Netherlands

learning, particularly about their achievements and the outcomes of their learning [11]. In addition to helping students develop as musicians, external feedback (e.g., from teachers or peers) is crucial for the self-evaluation process [73]. In music education, teachers can encourage the use of formative assessment techniques like self and peer evaluation to facilitate their assessments toward themselves. This can be particularly useful in a classroom setting where students can provide feedback to one another and to themselves under the right conditions.

Developing the skill of playing a musical instrument requires long-term practice, but young musicians may struggle to maintain their motivation for repetitive practice. Research has shown that inexperienced young players have not yet developed the ability to evaluate their own progress or use effective strategies in their practice [41]. Video recordings of these players' at-home practice sessions revealed that over 90% of the time was spent simply playing a piece from start to finish without using any particular performance-enhancing strategies [41]. The previous study has also shown that a significant number of learners commonly selected very low-level techniques for enhancing their performance and that these techniques are typically adopted by people with low levels of experience and are generally ineffective [28]. As a result, players with little experience tend to use ineffective techniques, whereas more skilled musicians are able to use more effective methods when learning an instrument. This difference in technique may be due to the varying levels of self-assessment skills that children with different levels of experience have. Children with lower levels of expertise may not be able to accurately analyze their own abilities and progress, which can prevent them from using the proper techniques in learning.

Moreover, research has explored the role of personality-related traits [14] and gender differences [29] in predicting music-practicing behavior. Hallam (2017) discovered statistical differences between boys and girls in the choice of systematic practice tactics, concentration, and quick correction of mistakes [29]. Statistically significant gender disparities were observed concerning the adoption of systematic practice strategies, with girls demonstrating a higher inclination towards systematic approaches. Boys, on the other hand, perceived themselves as having higher levels of concentration. Additionally, girls reported a greater tendency to promptly correct errors compared to boys [29].

Furthermore, research has shown that as musicians progress in their abilities, their sense of self-efficacy also increases. Schunk [67] suggested that allowing students to self-assess their own development can help them recognize their own competence, which in turn, can promote their self-efficacy. Previous research suggests that self-assessment treatments can be an effective way of enhancing students' self-efficacy. This is crucial since one of the main indicators of student

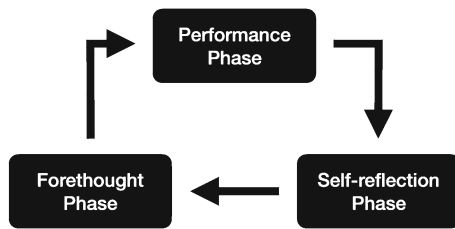
achievement is self-efficacy [30, 70]. Combining this with past research that showed social robots may increase children's motivation for practicing musical instruments [72], it is reasonable to assume that implementing a self-assessment strategy using a social robot will enhance students' motivation and performance. Therefore, this study undertakes an empirical experiment with children to determine whether a social robot that facilitates self-assessment may positively affect children's motivation and performance.

In this paper, we described a within-subject experiment conducted to compare children's motivation and performance practicing in different robot conditions (i.e., self-assessment robot and non-evaluative robot). 50 children practiced with both a self-assessment robot and a non-evaluative robot in random orders. Their motivation was measured by a questionnaire while performance was measured through evaluations of audio recordings. The results implied the positive effect of robot-initiated self-assessment on children's practice and design insights for personalized child-robot interaction in musical education.

## 2 Related Works

### 2.1 Self-regulated Learning and Motivation in Music Learning

Self-regulated learning (SRL), also known as self-regulation, is the process through which students independently activate and maintain cognition, attitudes, and behaviors that are methodically geared toward achieving learning objectives. For a long time, researchers in psychology and educational environments have been interested in how people control their own cognitive processes. For example, socio-cognitive theory highlights how social factors may positively or negatively influence how students consider their ability to request assistance over time [4]. There are several models of self-regulated learning, including those established by Boekaerts [9], Borkowski [10], Pintrich [61], Winne [78], and Zimmerman [84]. Boekaerts [7, 8] proposed a model of adaptable learning for the classroom that assigned appraisal a key role in the SRL process. In 2000, Boekaerts and Niemivirta proposed an expanded version of the adaptive learning paradigm [10], which stressed the SRL process' non-unitary feature. The attributes of an effective strategy user or information processor were established by Borkowski and his colleagues (see, for instance [10]). They suggest that effective information processing is the result of the successful implementation of cognitive, motivational, personal, and situational elements. At about the same time, Pintrich created a broad structure for SRL, which states that the four phases of SRL are forethought, monitoring, control, and reflection [60]. On the other hand, SRL is characterized as an event in the four-



**Fig. 1** Phases of self-regulated learning (Zimmerman, 1998, Zimmerman, 2000a)

stage model by Winne et al. [78, 80] and as a metacognitively directed behavior that enables students to adaptively control their use of cognitive techniques and strategies in the presence of a task, behavior is considered as an intrinsic aspect of learning [79]. As its name implies, Bandura's social cognitive theory [5] serves as the foundation for Zimmerman's (1989, 1990a,b, 1998, 2000a) social cognitive model of self-regulation [83, 84]. Zimmerman claims that self-regulation is iterative in nature and involves three phases: forethought, performance, and self-reflection (see Fig. 1, [84]). As was previously said, scientists gave varied definitions of SRL, but more notably, most of them agreed that SRL has a number of different stages and sub-processes.

Self-regulation concepts are currently being applied to academic learning and other types of learning, such as social and motor skills [10, 85]. Research indicated that learner skills and knowledge do not entirely explain student accomplishment [85], but rather components like self-regulation and drive were significant and provided some of the impetus for investigating academic self-regulated learning. Pintrich argues that self-regulatory behaviors control interactions between students and their surroundings and affect student success [60, 62]. Similar ideas have also been discussed by Zimmerman [83], who remarked that the social context or the environment plays a significant role in students' self-regulated learning (SRL). Learners evaluate their own performance, and such evaluations serve as the foundation for additional attempts to control their context, motivation, and behavior. In order to exercise good self-regulation, students must assess their own abilities, the learning environment, and any modifications that will improve their performance. The skills outlined above are particularly crucial for the development of musical skills in the context of music learning, which is a self-regulated learning activity.

For all facets of music learning and practice, SRL has the potential to significantly improve the success of musical skill learning [45]. In contrast, Ericsson suggested that the quantity of intentional practice is the primary difference between successful and failed learners during long-term instrument learning to attain a high-level accomplishment [24]. Studying to play an instrument involves long-term training to reach high proficiency levels. The majority of the research on musi-

cians' practicing has either tracked their actions [26, 49] or measured the amount of conscious practice they undertake [24]. Early theoretical applications of the SRL paradigm have focused on how much young musicians might engage in the types of SRL activities that unsupervised home practice needs [42, 43].

As previously discussed, the social context and surroundings are also essential for self-regulated learning [83]. Previous research has demonstrated the significance of parental involvement and peer support in children's music learning [16, 50]. Additionally, recent research has looked into how social robots may encourage youngsters to practice their musical instruments with social support [71, 72], suggesting that social robots might be successfully incorporated into children's musical instrument training. An earlier investigation regarding social robots in self-regulated learning environments has demonstrated the potential for using social robots as sources of social support in self-regulated learning [33].

## 2.2 Self-assessment in Music Learning

The case for the relationship between self-assessment and self-regulated learning (SRL) is compelling (e.g., [3, 59]. According to Panadero et al. [58], depending on the SRL assessments utilized, self-assessment treatments had a beneficial influence on students' SRL techniques with an effect size ranging from modest to medium. Researchers have provided a number of definitions for "self-assessment," all of which incorporate self-awareness and self-evaluation (e.g., [23, 57]).

However, the significance of self-evaluation has not been clearly shown. According to Andrade's assertion [3], in order to develop self-evaluation skills, external feedback may be required. And since the goal of feedback is to notify changes to procedures and outcomes that deepen learning and improve performance, the goal of self-assessment is to produce feedback that supports learning and enhances performance. This learning-oriented objective of self-assessment means that it ought to be formative: self-assessment is practically useless if there is no possibility for modification and correction. Recent debates over formative assessment in music have centered on the possible learning advantages of rubrics, self-reflection, and self and peer assessment along with concrete ways for fostering learning through assessment [18, 19, 68]. Earlier research has applied various techniques for implementing self-assessment strategies. The majority of these complied with common sense in terms of the link between self-assessment processes and the SRL phases, which is that self-assessment happens at the last stage (i.e., when the learning task is completed) [51].

Self-assessment has been employed in a variety of circumstances in the domain of music education. It has been used to

investigate creativity (for example, [63]), conducting ability (for example, [82]), and teacher preparation (e.g., [54]). Self-assessment has also been studied in the context of musical performance. Empirical studies show that social interactions (e.g., teacher-pupil-student) may improve the efficacy of self-regulation [17]. On the other hand, researchers have not yet paid as much attention to practice scenarios, despite the fact that practice is among the most crucial aspects of musical instrument learning. In the current study, we use a social robot (Pepper) to initiate a self-assessment session during children's musical instrument practice. To assess the benefits of the robot-initiated self-assessment, we also compared it to a control condition in which the robot exhibited identical behavior during musical practice except for the self-assessment strategy. In other words, we made a comparison between the self-assessment role and the non-evaluative role of the social robot.

### 2.3 Gender Difference in Learning

The impact of gender differences on learning has been a popular topic for researchers in the past few decades. Earlier studies have focused on learning styles [35], self-efficacy in flipped classroom model [53], reading ability [47], teacher's stereotype [39, 52], teacher perceptions of students' cognitive skills [12], motivation [38, 46], and so on. Additionally, researchers have also pointed out that compared to gender, gender identity explains better the use of learning styles [69].

In the domain of music education, gender differences also exist. First of all, there is already a difference in the choice of musical instruments before they start the learning process. Evidence showed that boys tend to choose drums, trombones, and trumpets, while girls prefer to learn flute, violin, and clarinet [1]. There could be instrument gender stereotypes they got from their parents [20, 21]. Multiple studies have found distinctive ways boys and girls engage with music, such as competence beliefs [22] and performance in school [2], in which girls have better scores. However, a study by McPherson et al. concluded different results, in which they found no significant difference between genders in competence beliefs, but differences in importance, usefulness, and task difficulties [44]. Moreover, researchers have also reported that girls adopted more systematic strategies during practice, boys concentrated better, and girls corrected their errors more often [29]. And the difference changes over time [77], for instance, a significant interaction effect was found between gender and levels of expertise on their choice of practicing strategies [29].

### 2.4 Research Questions and Hypotheses

After analyzing the literature mentioned above, the following research questions were addressed in this study:

1. How does a robot that initiates self-assessment, compared to a non-evaluative robot, impact piano practicing?
2. How does a robot that initiates self-assessment influence the motivation and performance of children of different genders in different learning stages?

Based on earlier studies, self-assessment has the potential to benefit motivation and learning outcomes during the learning process. Additionally, in an earlier study that compared children's motivation and performance while practicing with different conditions (i.e., alone, evaluative robot, and non-evaluative robot), we observed an interaction effect between children's learning stages and robot conditions [72], children in different learning stages tended to have different preferences for practice accompanies. Therefore, we propose the following hypotheses: (H1) Children will exhibit higher motivation and performance while practicing with the self-assessment robot than with the non-evaluative robot. (H2) An interaction effect will be observed between robot conditions and children's different learning stages.

The main goal of our experiment is to evaluate the effects of the self-assessment session during robot-assisted music practice. However, based on the literature, individual differences, including personality traits and gender have an effect on the use of learning strategies during music practice. Therefore, the second research question was raised as an exploratory question based on the findings of our earlier studies to investigate the effect of the player's learning stage. The long-term goal of this research is to design support for learning strategies (self-assessment and monitoring) considering the effect of individual differences in music practice based on personality, learning stage, age, gender, and others.

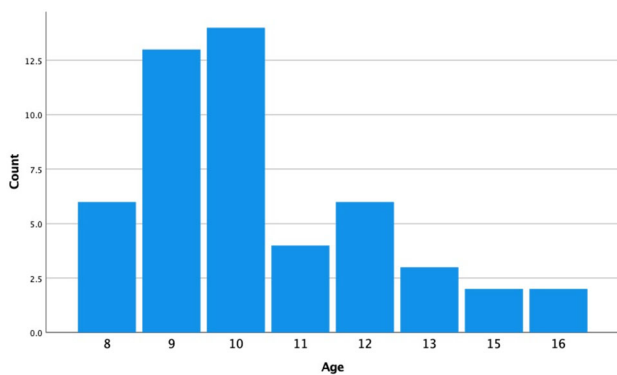
## 3 Methods

To answer the research questions, we conducted a within-subjects empirical study with children in a real-world setting, i.e., a music practice room in a local music school, with the role of the robot as the within-subjects factor. A detailed description of the methodology of the study can be found in the following subsections.

This study was conducted in three music schools in Eindhoven, Den Haag, and Veghel in The Netherlands to reach enough participants. As the most popular musical instrument among children and the easiest to adjust the difficulty levels for the children, the piano was chosen to be the musical instrument used in the study.

### 3.1 Participants

Fifty children ( $N = 50$ , 21 male and 29 female), aged from eight to sixteen years old, participated in the experiment.



**Fig. 2** Age distribution of the participants

They were all currently taking piano lessons at the music schools. The learning stages of the children varied from one month to ten and a half years, which is indicative of their level of expertise. All participants were able to understand and finish the questionnaire by themselves.

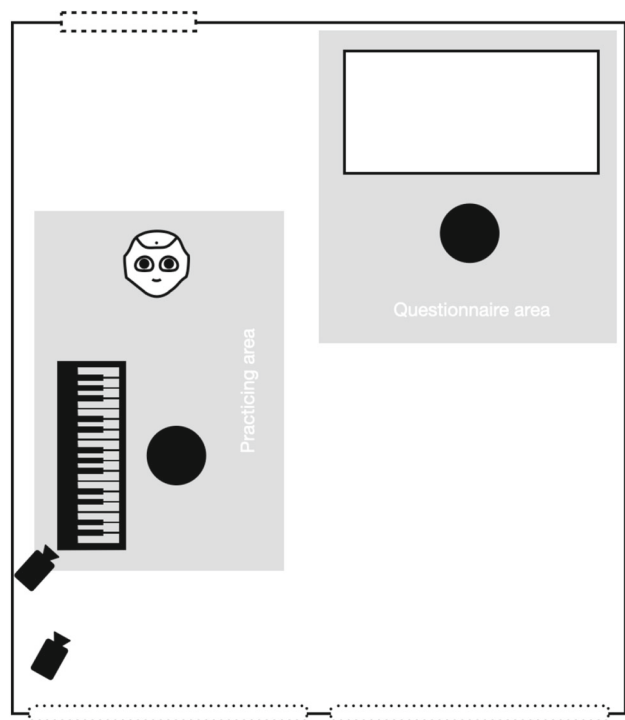
To answer the second research question, children were divided into three different learning stage groups during later data analysis: beginners ( $n = 25$ ), who have learned the piano for less than two years; developing players ( $n = 16$ ), who have learned the piano for more than two years but less than four years; advanced players ( $n = 9$ ), who have learned the piano for over four years. The division was made with suggestions from two piano teachers, both of whom have been teaching piano for over 15 years.

### 3.2 Experimental setup

Children practiced with the robot in the music school's practice room (see Fig. 3). Even though the robot completed a few of the interactions automatically (using the voice recognition function of the robot), the Wizard of Oz approach was used to conduct the experiment. Through the webcam, the researchers could control the robot from a different room. All of the sessions were recorded using a camera to document what happened while the children practiced with the robot.

### 3.3 Music Materials

To ensure that we would measure the effect of practicing with the robot on the children's performance rather than their prior experience, we provided each child with new music scores that matched their ability for each session (two music pieces were randomly assigned for the two sessions). These music notes were also chosen to be short enough for the children to learn and practice within a limited time frame. We created a total of six pieces of music, two for each group of children at different stages of learning (i.e., beginners, developing play-



**Fig. 3** Experiment room set up

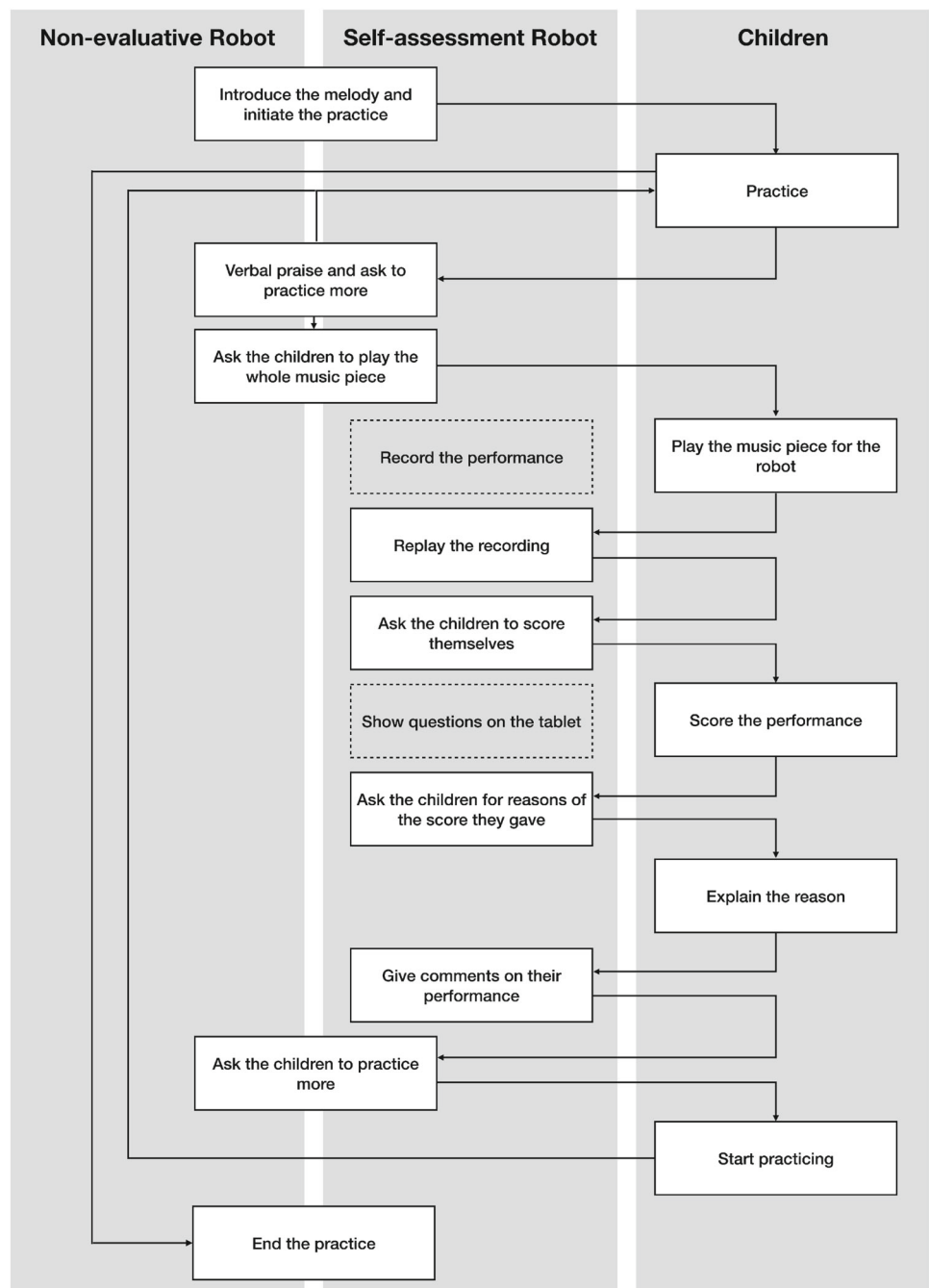
ers, and advanced players). These pieces were reviewed and approved by two experienced music teachers to be suitable for the three different levels.

### 3.4 Experimental Procedures

To address the research questions, this study used a within-subject experimental design, in which all participants practiced with the robot under two different conditions (i.e., with a non-evaluative robot and a self-assessment robot).

Upon arrival at the experiment location, a researcher provided the participants (i.e., children) with an explanation of the process and reviewed the music notes to ensure that they were suitable for the children's skill level. Their parents were asked to sign the informed consent form for the study. Following that, the participants and their parents were invited to meet the robot in the experiment room for the introduction session, in which the robot would introduce itself and play a game with the participants to let them get familiar with the robot. After the introduction session, the robot asked everyone except the participants to leave the experiment room so that they could start practicing together. Then, the practice started. Each participant practiced for two sessions with the robot that behaved differently in random orders. In the self-assessment robot condition, the robot started the practice by introducing new music notes and offering teaching videos. Then the children started practicing. After each time they finished one of the music notes, the robot gave compliments

**Fig. 4** Procedure of the practicing session



and encouraged them to practice more (e.g., “Well done! I’m still in your melody, could you please play more for me? Thank you!”). During each session (around 15 min), the robot initiated a self-assessment process for the children three times (see Fig. 4). First of all, the robot asked the children to play the whole music piece, in the meantime, the robot also recorded the performance of the children. After they finished the performance, the robot played the recording of their performance and asked them to score themselves on its tablet from three different perspectives, which are pitch, rhythm, and tempo, along with a justification of the scores. After the

children answered the questions, the robot offered comments on their performance and asked them to practice more, which is the end of the self-assessment process. At the end of each session, the robot asked the children to fill in a questionnaire that measured their motivation. Children took a ten to fifteen-minute break before they started the second practice session. There are three purposes of the break: firstly, children can get a rest from the practice physically; secondly, it gives children time to relax mentally so that they can have a fresh mind for the new robot condition; thirdly, researchers could take this time to change the shirt of the robot into another color. In

**Table 1** Motivation questionnaire questions

Source	Dimension	Question
FunQ	Autonomy	I knew what to do
		I did this activity because I had to. ( <i>r</i> )
		I did this activity because I wanted to
	Delight	I was happy
		I had fun
		I want to do something like this again
Stress	I felt angry. ( <i>r</i> )	
	I felt sad. ( <i>r</i> )	
	I felt bad. ( <i>r</i> )	
SIMS	Interest	I could focus easily
		I think this practice is important
		I did this activity because I wanted to

*r*: reversed items

addition, parents were asked to fill in a form that gathers demographic information (e.g., gender, age, and how long they have been learning piano.) about their children.

As for the non-evaluative robot condition, in order to balance the amount of interaction between children and robots with the self-assessment robot condition, the robot also asked the children to play the whole music piece three times during each session. However, in this condition, the robot only gave comments and compliments after the children played the music piece and asked them to practice more. This role was adapted from the design of a previous study [71], which designed a non-evaluative role and an evaluative role for a social robot.

### 3.5 Measurements

Children's motivation was measured by a questionnaire while their performance was evaluated by two experienced music teachers.

To create a reliable measurement of children's motivation during practice with the robot, we adapted questions from the FunQ questionnaire, an instrument intended to measure children's fun during a learning activity [74] and the Situational Motivation Scale (SIMS), which is a brief measure of situational motivation based on self-determination theory [27]. We thus derived a questionnaire with four dimensions, which are autonomy, delight, and stress from the FunQ questionnaire, and interest from the SIMS. All questions were adapted to match children's perceptions. This same questionnaire has been utilized in earlier studies which verified its reliability [72]. The autonomy, delight, and stress components were adapted from the FunQ questionnaire and used as indicators of motivation in the current questionnaire. The items in the interest dimension of the questionnaire were adopted from the SIMS, and were not included in the other

three dimensions. Each dimension consisted of three questions. Four items (one from the autonomy dimension and all three items in the stress dimension) were presented reversely and analyzed oppositely. The questions were modified into 5-point Likert questions and made simpler by adapting the language to the children's capacity, based on the findings of Mellor and Moore [48] and de Leeuw [37]. The average of the children's responses was calculated to construct a reliable (*Cronbach's alpha* = 0.85) measure for children's motivation during practice with a robot confirming the findings of an earlier study that used the same questionnaire [72] regarding the reliability of the scale.

To assess performance, we selected the children's final recordings from the two practice sessions. All the clipped recordings were sent to two independent music teachers, who gave evaluations on all the recordings. Evaluation forms and explanations of each performance factor were sent to the music teachers with the recording clips to ensure the reliability and validity of the measurement. By averaging the scores from the two music teachers' evaluations of children's performance, we were able to obtain a less biased assessment of children's performance when practicing in different robot conditions.

### 3.6 Data Analysis

In this part, we discuss the analysis we did on the questionnaire data and the data we collected to answer the study questions. The questionnaire findings were utilized to answer the motivation part of the questions. We used SPSS 28.0.1.0 to analyze the questionnaire data.

Performance data were scored by two experienced music teachers. That is after we edited and sent the recording clips of the children's last performance, two music teachers independently evaluated each performance from four perspectives (rhythm, pitch, tempo, and general impression) by giving a score from one to ten on each perspective. We took the average of the scores from the two music teachers as the measure of the performance of each child on four different aspects: rhythm (*Cronbach's alpha* = 0.88), pitch (*Cronbach's alpha* = 0.90), tempo (*Cronbach's alpha* = 0.86), and general impression (*Cronbach's alpha* = 0.93).

Because we used a within-subject experimental design, paired sample t-tests were performed to compare the two robot conditions with respect to children's motivation and performance. Repeated measures ANOVA was conducted to compare the influence of robot-initiated self-assessment on the motivation and performance of children in different learning stages.



## 4 Results

Detailed results of the current study are listed and explained in this section. In order to answer the research question and before we performed any of the analyses mentioned in the data analysis section, we first tested the normality of the questionnaire and performance data by Kolmogorov-Smirnov and Shapiro-Wilk test, the results showed that they all followed a normal distribution ( $p < 0.05$ ). Afterward, we performed a paired sample t-test and repeated measure ANOVA to answer the research questions as below.

Before we started testing the hypotheses, we first performed a paired sample t-test to examine whether the order effect influenced the results. By comparing children's motivation and performance during the first and second sessions (regardless of the robot conditions), we found that the order effect was compensated by randomizing the order of robot conditions (motivation:  $t(49) = 0.01$ ,  $p = 1.00$ ,  $d = 200.77$ ; general performance:  $t(49) = 1.34$ ,  $p = 0.19$ ,  $d = 0.32$ ). Furthermore, children's age can be expected to correlate with the learning stage, as children at later learning stages have had more years of training and, therefore, are also likely to be older. To test the relationship between children's age and learning stages, we performed a correlation between children's age and learning stages. Results showed a positive correlation ( $r(49) = 0.60$ ,  $p < 0.001$ ). However, earlier research has shown a correlation between learners' self-assessment ability and their learning stage, instead of their age [41]. Therefore, as shown in the research questions, we chose to only focus on children's learning stage which is independent of children's age.

### 4.1 Effects on Motivation and Performance

To answer the first research question (*How does a robot that initiates self-assessment, compared to a non-evaluative robot, impact piano practicing?*), we made the comparison between two conditions (i.e., non-evaluative robot and self-assessment robot) children experienced by conducting a paired sample t-test on their motivation and performance. Results showed significant differences between the two robot conditions (i.e., non-evaluative robot and self-assessment robot). More specifically, children had higher motivation ( $t(49) = 5.55$ ,  $p < 0.001$ ,  $d = 0.25$ ), rhythm performance ( $t(49) = 4.97$ ,  $p < 0.001$ ,  $d = 1.20$ ), pitch performance ( $t(49) = 5.76$ ,  $p < 0.001$ ,  $d = 1.16$ ), tempo performance ( $t(49) = 5.82$ ,  $p < 0.001$ ,  $d = 1.33$ ), and general impression ( $t(49) = 6.87$ ,  $p < 0.001$ ,  $d = 1.01$ ) when they practiced with the self-assessment robot (which initiated a self-assessment process for the children) (see Table 2, motivation:  $M = 4.81$ ,  $SD = 0.23$ ; rhythm:  $M = 7.11$ ,  $SD = 1.75$ ; pitch:  $M = 7.32$ ,  $SD = 1.86$ ; tempo:  $M = 7.35$ ,  $SD = 1.71$ ; general impression:  $M = 7.39$ ,  $SD = 1.83$ ) than the non-evaluative robot (which

did not initiate a self-assessment process for the children) (motivation:  $M = 4.61$ ,  $SD = 0.27$ ; rhythm:  $M = 6.22$ ,  $SD = 1.94$ ; pitch:  $M = 6.31$ ,  $SD = 2.05$ ; tempo:  $M = 6.18$ ,  $SD = 2.03$ ; general impression:  $M = 6.35$ ,  $SD = 1.99$ ). Thereby, results indicated that the self-assessment strategy initiated by the robot can help children better with their motivation and performance during practice.

### 4.2 Interaction Effects for Different Learning Stages and Gender

In order to answer the second research question (*How does a robot which initiates self-assessment influence the motivation and performance of children in different learning stages?*), as mentioned in the method section, we first divided the children into three learning stage groups, which are beginners ( $n = 25$ ), developing players ( $n = 16$ ), and advanced players ( $n = 9$ ). Then, we used the learning stage group and gender as between-subject factors in the repeated measure ANOVA to check the interaction effect between these factors and robot conditions. According to the earlier studies mentioned in Sect. 2.3, gender could also be a factor that can influence children's practice. Additionally, there are studies that found gender differences in child-robot interaction during music practice [71, 72].

#### 4.2.1 Motivation

Primarily, the results of repeated measure ANOVA confirmed the results we got from paired sample t-test with the significant main effect toward robot conditions on children's motivation (see Table 3,  $F = 20.66$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.41$ ) and performance (rhythm performance:  $F = 24.01$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.39$ ; pitch performance:  $F = 16.31$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.38$ ; tempo performance:  $F = 22.97$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.43$ ; and general impression:  $F = 18.65$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.52$ ). With learning stage groups as the between-subject factor, results provided no evidence for an interaction between learning stage groups and robot conditions on children's motivation ( $F = 0.61$ ,  $p = 0.55$ ). We also conducted additional exploratory analysis, which showed that if we add gender as another between-subject factor, results present evidence for a three-way interaction effect between robot conditions, learning stage groups, and gender (see Fig. 5,  $F = 5.50$ ,  $p = 0.01$ ,  $\eta_p^2 = 0.28$ ). As we can see in Fig. 5, no matter what gender they are, developing players always have higher motivation in the self-assessment robot condition. However, the male advanced players tended to have lower motivation than the beginners and developing players in the non-evaluative robot condition, and similarly higher motivation while practicing with the self-assessment robot. The female advanced players seemed to have similarly low motivation when practicing with the

**Table 2** Means and standard deviations of children's motivation, rhythm performance, pitch performance, tempo performance, and general impression practicing with a self-assessment robot and a non-evaluative robot

Conditions	Motivation		Rhythm		Pitch		Tempo		General impression	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Self-assessment	4.81	0.23	7.11	1.75	7.32	1.86	7.35	1.71	7.39	1.83
Non-evaluative	4.61	0.27	6.22	1.94	6.31	2.05	6.18	2.03	6.35	1.99

**Table 3** Repeated measure ANOVA results of children's motivation, rhythm performance, pitch performance, tempo performance, and general impression practicing with a self-assessment robot and a non-evaluative robot

Motivation		Rhythm		Pitch		Tempo		General impression	
<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
20.66	0.00***	24.01	0.00***	16.31	0.00***	22.97	0.00***	18.65	0.00***

\*\*\*  $p < 0.001$ 

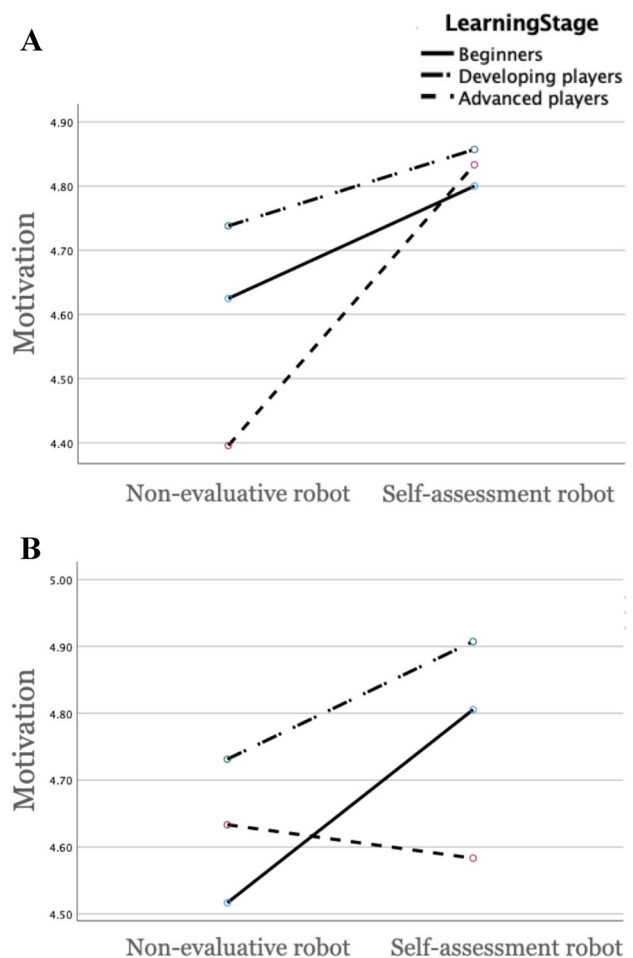
non-evaluative robot. But during the practice with the self-assessment robot, the female advanced players tended to have lower motivation while the beginners and developing players had higher motivation than the non-evaluative robot condition.

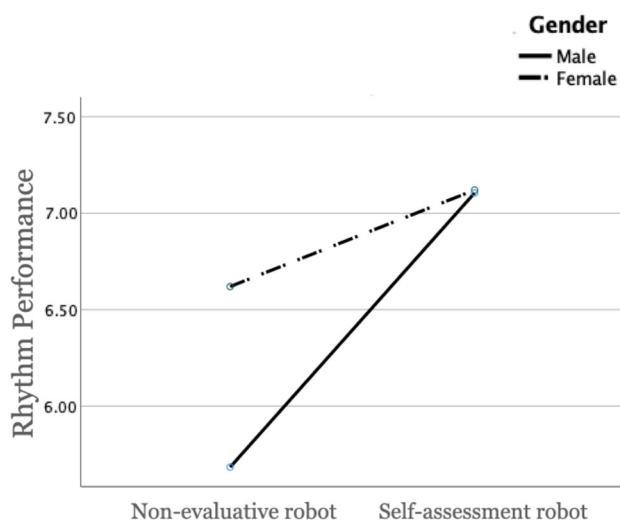
#### 4.2.2 Performance

The same analysis was conducted on the performance data. Similar to what we found on children's motivation, results provided no evidence for an interaction between robot conditions and learning stage groups on children's performance ( $F = 1.27$ ,  $p = 0.29$ ,  $\eta_p^2 = 0.13$ ). Additionally, we found an interaction effect between robot conditions and genders on children's rhythm performance (see Fig. 6,  $F = 5.74$ ,  $p = 0.02$ ,  $\eta_p^2 = 0.13$ ). As can be seen in Fig. 6, male players tended to have lower rhythm performance than female players when they practiced with the non-evaluative robot, while their rhythm performance was similarly higher in the self-assessment robot condition.

## 5 Discussion

To investigate whether a social robot can initiate self-regulated learning and thereby help children practice musical instrument playing, we designed a within-subject empirical experiment in a real-world practice room. By comparing children's motivation and performance in two conditions, which are practicing with a robot that initiates self-assessment and practicing with a robot that only provides social support, we were able to investigate and verify the hypotheses.

**Fig. 5** Children's motivation in different robot conditions (i.e., self-assessment robot and non-evaluative robot), learning stage groups (i.e., beginners, developing players, and advanced players), and genders (i.e., male (A) and female (B))



**Fig. 6** Children's rhythm performance in different robot conditions (i.e., self-assessment robot and non-evaluative robot) and genders (i.e., female and male)

### 5.1 Self-assessment in Music Instrument Practice

RQ1 (*How does a robot that initiates self-assessment, compared to a non-evaluative robot, impact motivation and performance in piano practicing?*) explored the impact of applying self-assessment strategy through robots in children's musical instrument practice. The results supported our hypothesis (H1) which posited that a robot that initiates the self-assessment strategy for children's practice would improve their motivation and performance compared to the robot that does not offer the self-assessment strategy. The results of the study supported our hypothesis, demonstrating the importance of self-assessment in children's musical instrument practice. Additionally, the design of the robot used in this study offers insight for robot designers in child-robot interaction design.

As mentioned before, self-assessment has been used and researched in a number of different music contexts [54, 63], which has already shown the importance of self-assessment in music learning and performance. It could be because music learning requires much practice to reach high-level skills and self-regulated learning, which holds significant potential for increasing the efficiency of musical skill acquisition across all aspects of music performance instruction [45]. Our results provide evidence regarding the value of self-regulated learning, especially self-assessment, in musical instrument practice.

### 5.2 Differences in Learning Stage and Gender

Furthermore, we also investigated how a robot that initiates self-assessment influences the motivation and performance

of children in different learning stages. Contrary to our expectations, the results did not reveal a direct interaction effect between the robot conditions (i.e., self-assessment robot and non-evaluative robot) and the children's learning stages (i.e., beginners, developing players, and advanced players). Even though earlier studies have shown that it is important for researchers and teachers to develop different self-assessment interventions for students in different expertise levels [56, 64], this comparison has not been made before. Based on our results, we can conclude that the self-assessment strategy is beneficial for children's learning motivation no matter which learning stage they are in. Furthermore, earlier research has shown that as students grow in their expertise they also become more aware of how much more they still need to learn and how better others are [34]. This implies that even without external judgment from the self-assessment robot, the advanced players should already be able to assess themselves and get discouraged. In such a case, the compliments and positive feedback from the robot could help them rebuild their confidence and motivation.

Furthermore, results provided evidence for an interaction effect found between the type of robot (i.e., self-assessment robot and non-evaluative robot), the learning stage (i.e., beginners, developing players, and advanced players), and the gender (i.e., female and male). Combining the evidence from an earlier study [72], we argue that these results can be explained by gender differences between male players and female players. In that study, results showed that female players tended to have higher persistence with the non-evaluative robot than the male players with the non-evaluative robot, which is an indicator of their motivation level. The results implied that female players focused longer on practice and followed the instructions from the robot better than the male players during practice. As seen in Fig. 5, male players tended to have higher motivation in the self-assessment robot condition compared to the other condition, while female players showed similar motivation levels in both robot conditions. A possible explanation could be the small number of advanced players in our sample ( $n = 9$ ). Additionally, researchers claimed that females are more likely to have higher intrinsic motivation ([25, 55, 66]), which could be another explanation for the results.

As for the performance of children, the results were similar to children's motivation. No significant interaction effect was found between the robot conditions and the learning stage, which implies that children in all learning stages have similar preferences for the robot-initiated self-assessment. Earlier research has also found no difference between the two learning stages (primary school and secondary school students) in their motivation improvement [31]. But we found an interaction effect between robot conditions and genders on children's rhythm performance. Male players tended to have lower rhythm performance than female players when they

practiced with the non-evaluative robot, while their rhythm performance was similarly higher in the self-assessment robot condition. The result can also be explained by the different persistence levels female and male players have, female players, tend to focus more on practice with the non-evaluative robot, while male players focus less [72]. Furthermore, Wolters found that females reported significantly more strategy use than males [81]. In this case, even though the robot was not providing the self-assessment strategy in the non-evaluative robot condition, female players were using more other strategies to help themselves with practicing than the male players, which turned out to be the difference in their performance in the non-evaluative robot condition. Those can become the reason for the performance difference we found.

As the first study that employed a robot for initiating self-assessment in children's learning scenarios, we established its feasibility and its positive effects on children's motivation and performance. Earlier studies have investigated the impact of self-assessment in self-regulated learning [58] and how to measure self-assessment [76]. There is evidence showing that self-assessment initiated by students themselves can result in their motivation, and improvements in the effectiveness and quality of learning [13, 40, 75]. Our findings, which employed a social robot for self-assessment initiation confirm the positive effect of self-assessment in self-regulated learning and demonstrate the feasibility of using a social robot to help initiate self-assessment. Recently, human-robot interaction research has been paying increasing attention to personalization [32, 36]. In light of this trend, the differences between different learning stages and different genders we found may provide useful leads to robot designers but also researchers in this field. An alternative explanation for the individual differences found could pertain to the children's different capacities for self-assessment at different learning stages. Ross reports that self-assessment skills can be improved by training over time [65], the more they are used the more fluent they become. Children with longer periods of learning are likely to have gained a better awareness of their own abilities and knowledge thus their self-assessment skills will correspondingly improve in accuracy.

### 5.3 Limitations and Future Works

In this study, we conducted an experiment with fifty children enrolled in music schools. All participants were recruited voluntarily via the music schools. It should be noted, that the number of players in each learning stage group was not equal (beginners ( $n = 25$ ), developing players ( $n = 16$ ), and advanced players ( $n = 9$ )), which may introduce imbalance in the data during analysis.

Furthermore, although we tried to balance the amount/duration of interactions between the two robot conditions, the self-assessment robot still had more interactions

with the children than the non-evaluative robot during the self-assessment session. Even though the unequal amount of interactions between the two conditions could be one of the factors that affected the results, according to the results we got, children still had higher motivation and performance practicing with the self-assessment robot.

Based on our results, some directions for future research can be derived for the design of child-robot interaction. Interestingly, with the robot Pepper, there are more available (behavioral) functions that can be improved and used in the interaction design that were not used in our current study. For example, the limitations of speech recognition of Pepper restricted us from using it for difficult speech recognition, simple speech recognition in a noisy environment, or even music recognition. With an improvement of this function, the usefulness and fluency of Pepper in the context of music learning can be much better. Since there are also multiple sensors and cameras on Pepper, we might be able to develop various authentic interactions between the robot and children. Nonetheless, are more interactions better? In a previous study [72], we discussed the timing of interaction and based this timing on the specialties of the robot during children's musical instrument practice. Because of the noise robots make during their movement, we propose that it is better to interact with the children every time they finish practicing the piece once or every time the children pay attention to the robot. Otherwise, the interaction can be a disruption for the children in the practice.

## 6 Conclusion

This study is the first to examine child-robot interaction in self-regulated learning. We have demonstrated that robot-initiated self-assessment can help enhance children's motivation and performance during practice. However, it might work differently for children of different genders and at different learning stages. Such individual differences could be further explored in future studies. Extending our knowledge regarding the factors that influence the impact of social robots in self-regulated learning can help guide the personalization of robot behavior to the needs of learners.

**Author Contributions** All authors contributed to the study's conception and design. Material preparation and data collection were performed by Heqiu Song and Konstantinos Tsiakas. Data analysis was performed by Heqiu Song and Jaap Ham. The first draft of the manuscript was written by Heqiu Song and all authors revised the manuscript. All authors read and approved the final manuscript.

**Funding** Researcher Heqiu Song is funded by the China Scholarship Council.

**Data Availability** Only the materials used in this study are available within the article. Data are available only through request.

**Code availability** Not applicable.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethics approval** This study was approved by the ethics board of Eindhoven University of Technology (Reference number: ERB2019ID4).

**Consent to participate** All participants and their parents gave their consent for the data collection.

**Consent for publication** All listed authors have approved the manuscript before submission.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abeles HF, Porter SY (1978) The sex-stereotyping of musical instruments. *J Res Music Educ* 26(2):65–75
- Agak H (2002) Gender difference and academic achievement in music among form four students in Kenya 1991-1995. *Bulletin of the council for research in music education* pp 94–101
- Andrade H, Cizek GJ (2010) *Handbook of formative assessment*. Routledge, New York
- Bandura A (1997) *Self-efficacy: the exercise of control*. Macmillan, London
- Bandura A, Abrams K (1986) Self-regulatory mechanisms in motivating, apathetic, and despondent reactions to unfulfilled standards. Unpublished manuscript, Stanford University, Stanford, CA
- Belpaeme T, Kennedy J, Ramachandran A et al (2018) Social robots for education: a review. *Sci Robot* 3(21):eaat5954
- Boekaerts M (1992) The adaptable learning process: initiating and maintaining behavioural change. *Appl Psychol* 41(4):377–397
- Boekaerts M (1995) Self-regulated learning: bridging the gap between metacognitive and meta motivation theories. *Educ Psychol* 30(4):195–200
- Boekaerts M (1997) Self-regulated learning: a new concept embraced by researchers, policy makers, educators, teachers, and students. *Learn Instr* 7(2):161–186
- Boekaerts M, Niemivirta M (2000) Self-regulated learning: finding a balance between learning goals and ego-protective goals. In: *Handbook of self-regulation*. Elsevier, p 417–450
- Boud D, Falchikov N (1989) Quantitative studies of student self-assessment in higher education: a critical analysis of findings. *High Educ* 18(5):529–549
- Brandmiller C, Dumont H, Becker M (2020) Teacher perceptions of learning motivation and classroom behavior: the role of student characteristics. *Contemp Educ Psychol* 63(101):893
- Brown G, Harris L (2013) Student self-assessment
- Butkovic A, Ullén F, Mosing MA (2015) Personality related traits as predictors of music practice: Underlying environmental and genetic influences. *Personal Individ Differ* 74:133–138
- Chin KY, Hong ZW, Chen YL (2014) Impact of using an educational robot-based learning system on students' motivation in elementary education. *IEEE Trans Learn Technol* 7(4):333–345
- Creech A, Hallam S (2003) Parent-teacher-pupil interactions in instrumental music tuition: a literature review. *Br J Music Educ* 20(1):29–44. <https://doi.org/10.1017/S0265051702005272>
- Creech A, Hallam S (2011) Learning a musical instrument: the influence of interpersonal interaction on outcomes for school-aged pupils. *Psychol Music* 39(1):102–122
- Crochet LS, Green SK (2012) Examining progress across time with practical assessments in ensemble settings. *Music Educ J* 98(3):49–54
- DeLuca C, Bolden B (2014) Music performance assessment: exploring three approaches for quality rubric construction. *Music Educ J* 101(1):70–76
- Delzell JK, Leppla DA (1992) Gender association of musical instruments and preferences of fourth-grade students for selected instruments. *J Res Music Educ* 40(2):93–103
- Dibben N (2002) Gender identity and music. *Musical identities* pp 117–133
- Eccles J, Wigfield A, Harold RD et al (1993) Age and gender differences in children's self-and task perceptions during elementary school. *Child Dev* 64(3):830–847
- Epstein RM, Siegel DJ, Silberman J (2008) Self-monitoring in clinical practice: a challenge for medical educators. *J Contin Educ Heal Prof* 28(1):5–13
- Ericsson KA, Krampe RT, Tesch-Römer C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 100(3):363
- Gledhill R, Van der Merwe C (1989) Gender as a factor in student learning: preliminary findings. *Medical Education*
- Gruson LM (1988) Rehearsal skill and musical competence: Does practice make perfect. *Generative processes in music: the psychology of performance, improvisation, and composition* pp 91–112
- Guay F, Vallerand RJ, Blanchard C (2000) On the assessment of situational intrinsic and extrinsic motivation: the situational motivation scale (sims). *Motiv Emot* 24(3):175–213
- Hallam S, Cross I, Thaut M (2011) *Oxford handbook of music psychology*. Oxford University Press, Oxford
- Hallam S, Varvarigou M, Creech A et al (2017) Are there gender differences in instrumental music practice? *Psychol Music* 45(1):116–130
- Honicke T, Broadbent J (2016) The influence of academic self-efficacy on academic performance: a systematic review. *Educ Res Rev* 17:63–84
- Hortigüela Alcalá D, Hernando Garijo A, Pérez-Pueyo Á et al (2019) Cooperative learning and students' motivation, social interactions and attitudes: perspectives from two different educational stages. *Sustainability* 11(24):7005
- Irfan B, Ramachandran A, Spaulding S, et al (2019) Personalization in long-term human-robot interaction. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), IEEE, pp 685–686
- Jones A, Bull S, Castellano G (2018) "i know that now, i'm going to learn this next" promoting self-regulated learning with a robotic tutor. *Int J Soc Robot* 10(4):439–454
- Kruger J, Dunning D (1999) Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 77(6):1121
- Kulturel-Konak S, D'Allegro ML, Dickinson S et al (2011) Review of gender differences in learning styles: suggestions for stem education. *Contemp Issues Educ Res* 4(3):9–18

36. Lee MK, Forlizzi J, Kiesler S, et al (2012) Personalization in HRI: a longitudinal field experiment. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, pp 319–326
37. de Leeuw E (2011) Improving data quality when surveying children and adolescents: cognitive and social development and its role in questionnaire construction and pretesting. Tech. rep., Finland
38. McGeown S, Goodwin H, Henderson N et al (2012) Gender differences in reading motivation: Does sex or gender identity provide a better account? *J Res Reading* 35(3):328–336
39. McKown C, Weinstein RS (2002) Modeling the role of child ethnicity and gender in children's differential response to teacher expectations 1. *J Appl Soc Psychol* 32(1):159–184
40. McMillan JH, Hearn J (2008) Student self-assessment: the key to stronger student motivation and higher achievement. *Educ Horiz* 87(1):40–49
41. McPherson GE, Renwick JM (2001) A longitudinal study of self-regulation in children's musical practice. *Music Educ Res* 3(2):169–186
42. McPherson GE, Renwick JM (2011) Self-regulation and mastery of musical skills. *Handbook of self-regulation of learning and performance*. pp 234–248
43. McPherson GE, Zimmerman BJ (2011) Self-regulation of musical learning: A social cognitive perspective on developing performance skills. In: MENC handbook of research on music learning. Oxford University Press
44. McPherson GE, Osborne MS, Barrett MS et al (2015) Motivation to study music in Australian schools: the impact of music learning, gender, and socio-economic status. *Res Stud Music Educ* 37(2):141–160
45. McPherson GE, Osborne MS, Evans P et al (2019) Applying self-regulated learning microanalysis to study musicians' practice. *Psychol Music* 47(1):18–32
46. Meece JL, Glienke BB, Askew K (2009) Gender and motivation. *Handbook of motivation at school*. pp 425–446
47. Meissel K, Meyer F, Yao ES et al (2017) Subjectivity of teacher judgments: exploring student characteristics that influence teacher judgments of student ability. *Teach Teach Educ* 65:48–60
48. Mellor D, Moore KA (2014) The use of likert scales with children. *J Pediatr Psychol* 39(3):369–379. <https://doi.org/10.1093/jpepsy/jst079>
49. Miksza P (2011) A review of research on practicing: summary and synthesis of the extant research with implications for a new theoretical orientation. *Bull Council Res Music Educ* 190:51–92
50. Moore DG, Burland K, Davidson JW (2003) The social context of musical success: a developmental account. *Br J Psychol* 94(4):529–549. <https://doi.org/10.1348/00071260322503088>
51. Morijiri Y, Welch GF (2022) Decisions on the quality of piano performance: evaluation of self and others. *Front Psychol* 13:7090
52. Muntoni F, Retelsdorf J (2018) Gender-specific teacher expectations in reading—the role of teachers' gender stereotypes. *Contemp Educ Psychol* 54:212–220
53. Namaziandost E, Çakmak F (2020) An account of EFL learners' self-efficacy and gender in the flipped classroom model. *Educ Inf Technol* 25(5):4041–4055
54. Napoles J (2010) Differential effects of instructor feedback vs. self-observation analysis. *Bulletin of the Council for Research in Music Education* (183)
55. Niemivirta M (1997) Gender differences in motivational-cognitive patterns of self-regulated learning
56. Panadero E (2011) Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics. Unpublished doctoral dissertation) Universidad Autónoma de Madrid, Madrid, España
57. Panadero E, Brown GT, Strijbos JW (2016) The future of student self-assessment: a review of known unknowns and potential directions. *Educ Psychol Rev* 28(4):803–830
58. Panadero E, Jonsson A, Botella J (2017) Effects of self-assessment on self-regulated learning and self-efficacy: four meta-analyses. *Educ Res Rev* 22:74–98
59. Panadero E, Andrade H, Brookhart S (2018) Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *Austr Educ Res* 45(1):13–31
60. Pintrich PR (2000) The role of goal orientation in self-regulated learning. In: *Handbook of self-regulation*. Elsevier, p 451–502
61. Pintrich PR, Garcia T (2012) Self-regulated learning in college students: knowledge, strategies, and motivation. In: *Student motivation, cognition, and learning*. Routledge, p 129–150
62. Pintrich PR, Zusho A (2002) The development of academic self-regulation: The role of cognitive and motivational factors. In: *Development of achievement motivation*. Elsevier, p 249–284
63. Priest T (2006) Self-evaluation, creativity, and musical achievement. *Psychol Music* 34(1):47–61
64. Reitmeier CA, Vrchota DA (2009) Self-assessment of oral communication presentations in food science and nutrition. *J Food Sci Educ* 8(4):88–92
65. Ross JA (2006) The reliability, validity, and utility of self-assessment. *Pract Assess Res Eval* 11(1):10
66. Rozendaal J, Minnaert A, Boekaerts M (2003) Motivation and self-regulated learning in secondary vocational education: information-processing type and gender differences. *Learn Individ Differ* 13(4):273–289
67. Schunk DH (1996) Self-efficacy for learning and performance
68. Scott SJ (2012) Rethinking the roles of assessment in music education. *Music Educ J* 98(3):31–35
69. Severiens S, Dam GT (1997) Gender and gender identity differences in learning styles. *Educ Psychol* 17(1–2):79–93
70. Sitzmann T, Ely K, Brown KG et al (2010) Self-assessment of knowledge: A cognitive learning or affective measure? *Acad Manag Learn Educ* 9(2):169–191
71. Song H, Zhang Z, Barakova EI, et al (2020) Robot role design for implementing social facilitation theory in musical instruments practicing. In: Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. ACM, Cambridge United Kingdom, pp 253–260, <https://doi.org/10.1145/3319502.3374787>,
72. Song H, Barakova EI, Markopoulos P, Ham J (2021) Personalizing HRI in musical instrument practicing: the influence of robot roles (evaluative versus nonevaluative) on the child's motivation for children in different learning stages. *Front Robot AI* 8:699524
73. Taras M (2003) To feedback or not to feedback in student self-assessment. *Assess Eval High Educ* 28(5):549–565
74. Tisza G, Markopoulos P (2021) FunQ: measuring the fun experience of a learning activity with adolescents. *Curr Psychol*. <https://doi.org/10.1007/s12144-021-01484-2>
75. Topping K (2003) Self and peer assessment in school and university: reliability, validity and utility. *Optimising new modes of assessment: In search of qualities and standards* pp 55–87
76. Ward M, Gruppen L, Regehr G (2002) Measuring self-assessment: current state of the art. *Adv Health Sci Educ* 7:63–80
77. Webster PR, Zimmerman MP (1983) Conservation of rhythmic and tonal patterns of second through sixth grade children. *Bulletin of the Council for Research in Music Education* pp 28–49
78. Winne P, Hadwin A, Hacker D, et al (1998) Metacognition in educational theory and practice. *Studying as self-regulated Learning* pp 277–304
79. Winne PH (1996) A metacognitive view of individual differences in self-regulated learning. *Learn Individ Differ* 8(4):327–353
80. Winne PH, Perry NE (2000) Measuring self-regulated learning. In: *Handbook of self-regulation*. Elsevier, p 531–566

81. Wolters CA (1999) The relation between high school students' motivational regulation and their use of learning strategies, effort, and classroom performance. *Learn Individ Differ* 11(3):281–299
82. Yarbrough C (1987) The relationship of behavioral self-assessment to the achievement of basic conducting skills. *J Res Music Educ* 35(3):183–189
83. Zimmerman BJ (1989) A social cognitive view of self-regulated academic learning. *J Educ Psychol* 81(3):329
84. Zimmerman BJ (2000) Attaining self-regulation: a social cognitive perspective. In: *Handbook of self-regulation*. Elsevier, p 13–39
85. Zimmerman BJ, Schunk DH (2001) *Self-regulated learning and academic achievement: theoretical perspectives*. Routledge, New York

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.