Original Research Article

National Association *for* Music Education

Journal of Research in Music Education
1–18

DOI: 10.1177/00224294211011962
jrme.sagepub.com
SAGE

Interval Identification Predicts Success in Melodic Dictation

Bryan E. Nichols[1] and D. Gregory Springer[2]

Abstract

The purpose of this study was to investigate possible predictive relationships between interval identification and melodic dictation performance on tasks where students identify short pitch spans after a brief tonicization. College musicians (N = 35) completed an interval identification test and a series of melodic dictation tasks. Results indicated that interval identification and melodic dictation tests reflected a battery of items ranging from very easy to very difficult with acceptable Cronbach's α levels. We conducted a two-stage hierarchical regression analysis to examine the extent to which interval identification served as a predictor of melodic dictation accuracy while controlling for selected music and demographic variables. Results indicated that interval identification served as a significant predictor of melodic dictation scores, contributing 28.9% of the variance in melodic dictation scores while controlling for musical experience variables. The analysis indicated a dictation task by interval ability interaction based on grouping by lower, mid-, and upper performing groups on the interval identification test. Issues in measurement of melodic dictation accuracy and strategies that affect the development of melodic dictation skills are discussed.

Keywords

melodic dictation, interval identification, pitch sequences, skill development, working memory

Melodic dictation is one of several fundamental skills typically taught in aural skills courses required of music majors at many colleges and universities (Buonviri, 2015a). Pembrook (1984) described dictation as "a process wherein the listener represents aural events, usually musical stimuli, by means of a standardized notational system

[1]The Pennsylvania State University, University Park, PA, USA
[2]Florida State University, Tallahassee, FL, USA

Corresponding Author:
Bryan E. Nichols, School of Music, The Pennsylvania State University, 105 Music Bldg. 1, University Park, PA 16802, USA.
Email: bnichols@psu.edu

peculiar to the field of music" (p. 1). The ability to dictate melodies accurately is claimed to be related to other skills, such as error detection of pitches while listening (Larson, 1977) and sight-singing (Norris, 2003). Results of descriptive studies have indicated that instructors of aural skills courses (a) preferred pitch systems that emphasized scale degree function (Buonviri & Paney, 2015; Paney & Buonviri, 2017; Taggart & Taggart, 1994), (b) believed that more instruction in aural skills pedagogy was needed (Buonviri & Paney, 2015), and (c) reported that written music theory and aural skills were typically taught in separate, independent courses (Taggart & Taggart, 1994). Because melodic dictation is a fundamental aural skill for developing musicians (Buonviri, 2015a), understanding the strategies that improve students' accuracy in these complex dictation tasks is an important goal for music educators in a variety of class and rehearsal settings.

Music theorists have long discussed the function of tonal centers due to the importance of a tonal "home" for structural and aesthetic elements of Western music. Krumhansl and Kessler (1982) explored the cognitive representation of pitch structures as a function of hierarchical relationships based on the structures of functional harmony and leading tones. A gradual decay in listeners' memory for the initial (primary) key occurs after a modulation (Farbood, 2016), although it is unclear how memory plays a role in the musician's maintenance of a tonal center for purposes like sight-singing, melodic dictation, or error detection. In typical Western tonal music, pitch sequences are defined by the relative position of each pitch to the tonic. Although it has been theorized that the tonic triad forms the primacy of pitch relationships (not the tonic itself; Parncutt, 2011), tonal music is defined by its melodic and harmonic properties in relation to a key.

Successful melodic dictation seems to be based on a variety of general cognitive abilities. For example, Buonviri (2014) observed that students who were consistently successful at melodic dictation were highly skilled at directing their attention to certain aspects of aural stimuli, prioritized dictation tasks intentionally while listening, and utilized a variety of problem-solving skills to complete the task. Additionally, instructors have reported that the ability to recognize pitch and rhythm patterns and the ability to use chunking to reduce the amount of tonal information are helpful in melodic dictation (Paney & Buonviri, 2014).

Researchers have also investigated other cognitive aspects and environmental variables that influenced musicians' success on melodic dictation tasks. To better understand why some students were less successful at melodic dictation tasks than others, Hoppe (1991) asked participants to sing the melody after notating it. Most participants who dictated melodies with mistakes perceived/heard the melodies correctly, as reflected in their accurate singing, but simply notated melodies incorrectly. Because most participants sung the melody correctly even after notating it, this finding suggests that those errors may be due to inaccurate understanding of musical pitches and modes rather than inaccurate hearing or perception.

Environmental variables nested within the dictation task may affect dictation accuracy, such as time delays/latencies between aural stimuli. For example, longer time delays between aural presentations resulted in less accuracy on melodic discrimination tasks due to loss of tonal information in short-term memory (Williams, 1975). Wuthrich and Tunks (1989), on the contrary, found that longer time delays between tones resulted in greater accuracy. Notably, there were different time delays examined in those studies that could have accounted for these differences. Williams (1975) tested for the effects of delays ranging from 0 to 15 s, and Wuthrich and Tunks (1989) tested delays ranging from 0 to 520 ms.

Researchers have examined various instructional strategies to determine whether they improve students' melodic dictation accuracy and have had mixed findings. College musicians who received high, medium, and low scores on melodic dictation tests did not differ significantly in the strategies used while completing dictation tasks (Powell, 2013). Similarly, Buonviri (2017) found no significant differences in music majors' dictation accuracy as a function of different listening strategies (required listening before writing, required writing while listening, and no specified strategy [control]), indicating that both strategies may be efficacious. In another study, music majors were more successful at "writing while hearing" and "listening before writing" than they were when "listening, singing, and writing" (Pembrook, 1986, p. 253). These mixed results highlight the complex nature of melodic dictation and the need for more systematic inquiry focused on melodic dictation instructional strategies.

Other instructional strategies that are intended to improve dictation performance may unintentionally function as distractors, resulting in negative dictation performance. For instance, college musicians were less accurate at melodic dictation when they sang a preparatory solfège pattern before dictating the target melody (compared to an immediate dictation without singing the solfège pattern; Buonviri, 2015b). Similarly, music majors in another study were less accurate in their melodic dictation when they were asked to sing the entire melody before dictating it (Buonviri, 2019). In both of these cases, Buonviri (2015b, 2019) attributed the decrease in dictation accuracy to the fact that singing unintentionally functioned as a distraction during melodic dictation tasks. Additionally, music majors who received instructions intended to focus their attention on various aspects of the melody about to be dictated demonstrated *less* dictation accuracy compared to those in a control condition without attention-focusing instructions (Paney, 2016). Taken together, these findings highlight how certain instructional strategies (e.g., singing a preparatory pattern or listening to unnecessary instructions) may inadvertently function as distractions, which may have negative effects on students' dictation accuracy.

Certain intrinsic factors within musical excerpts themselves may also impact melodic dictation accuracy. For instance, music majors' performance on melodic dictation tasks indicated varied levels of difficulty among diatonic, chromatic, and atonal musical styles (Larson, 1977). Various types of phrase expansion methods (e.g., exact repetition, tonal sequence, and rhythmic alteration), however, had no significant effects on music majors' dictation accuracy (Clements, 2003). In another study (Madsen & Staum, 1983), college students (music and nonmusic majors) were able to discriminate successfully among identical versus similar melodies, but changes in mode (major vs. minor) were more likely to cause discrimination errors than changes in meter (simple vs. compound). Similarly, Halpern's (1984) participants (collegiate musicians and nonmusicians) rated melodies as highly similar when they reflected changes in mode (e.g., major to minor), but they rated melodies as very dissimilar when melodies with identical pitches were presented with different rhythmic values. Sink (1983) found that when melodic and rhythmic information were presented simultaneously, listeners paid less attention to rhythmic information. Together, these findings may highlight a perceptual distinction between pitch and rhythm among listeners when they are presented with complex listening tasks that include both pitch and rhythmic information. How musicians acquire pitch discrimination skills and the ability to identify and recall specific sequences has important implications for music learning in a variety of settings.

Whereas music theorists have examined the role of the tonal center in musical works, performers employ tonicization—establishing a tonal center by way of hearing or performing pitches—for the purpose of playing or singing music. For musicians, tonicization is necessary for performing pitch tasks such as reading music at sight or performing music in rehearsal or concerts. Researchers must establish a tonal center for study tasks that require it, and in these cases, tonicization is a part of the procedure, not a variable in the design. There are many ways to establish a tonal center that include offering a single pitch as starting note or as tonic reference pitch and offering two or three pitches (a triad) presented either harmonically or melodically. Authors of singing accuracy studies have reported the use of one starting pitch to help children begin a song task (e.g., "Jingle Bells" in Nichols, 2016; Nichols & Wang, 2016). Previous research suggests that providing a tonal context improves pitch matching in adolescent boys (Demorest & Clements, 2007) and that tonal contextualization improves interval perception (Graves & Oxenham, 2017). Demorest et al. (2018) used a practice item prior to pitch matching to effectively provide contextualization. In the other studies (Buonviri, 2015b, 2019), participants were oriented to key by hearing a I-V7-I chord progression. Sight-singing research that also utilized tonicization reported that more accurate sight-singers more frequently retonicized vocally during practice (Killian & Henry, 2005). A presentation of the tonic chord followed by the starting pitch has also been used to establish the tonal center (e.g., Demorest, 1998; Killian & Henry, 2005).

Another tonicization strategy has been used when measuring performance on working memory span tasks. In those cases, researchers presented a tonic-dominant dyad melodically before testing musicians' ability to identify pitches in melodic triads (e.g., G–D–G–D in Nichols et al., 2018; Wöllner & Halpern, 2015). In the first of those studies, researchers noted a floor effect in a task in which college musicians were asked to recall the last pitch from a series of triads by either playing them serially on a piano or by notating them on a staff (Nichols et al., 2018). When participants used the piano, they were able to identify—and thus were more able to recall—the last pitch of each triad, but when they were not using the piano, participants did not notate the stimuli successfully. Both jazz and classical musicians were universally low in performance by comparison to the other tasks. In that study, participants heard the tonic–dominant dyad twice (G–D–G–D) and then were given melodic triads that began outside the tonic–dominant interval but always ended within the tonic–dominant range (i.e., the last pitch of the triad was always the first, second, third, fourth, or fifth scale degree). Performance on the task was unexpectedly low, which leads us to explore what series of tasks may provide easy, medium, and difficult levels of item difficulty in these types of working memory/dictation tasks. Furthermore, because the ability to dictate melodies successfully requires the ability to identify pitches in relation to other pitches, we wanted to examine the possible predictive relationships between interval identification and melodic dictation. The purpose of this study was to investigate possible predictive relationships between interval identification and melodic dictation performance on tasks where students identify short pitch spans after a brief tonicization. As a secondary purpose, we were interested in investigating what differences in melodic dictation performance may be evident among participants who scored high, medium, and low on an interval identification test. The research questions were (1) To what extent does interval identification performance predict melodic dictation performance when controlling for selected demographic variables related to prior musical experience? and (2) What differences in melodic dictation performance are evident among high, middle, and low performers on an interval identification test?

Method

In this study, collegiate musicians completed (a) an interval identification test and (b) a melodic dictation task. Specifically, we aimed to address the floor effect found on a previously used task in which student musicians were asked to identify the last pitch in a melodic triad (Nichols et al., 2018). In that task, musicians listened to a tonic–dominant dyad presented melodically (G4–D5–G4–D5) for tonicization, followed by a melodic triad that began outside the dyad and ended within the dyad. In the triad presentation, the last pitch was always between G4 and D5 (i.e., inside the dyad used for tonicization). We defined this as melodic dictation according to Pembrook (1984), where the listener is asked to represent musical stimuli by notating what is heard, although we acknowledge this differs from tasks mirroring the aural skills class exercise in which student musicians must notate as much of a melody as possible. Therefore, this task was intended to represent authentic musical stimuli where musicians identify a short range of pitches to contextualize what is heard for evaluation or error detection.

Participants

Participants[1] ($N = 35$; music majors, $n = 29$; other majors, $n = 6$) were musicians recruited from ensembles in the school of music of a large research university in the northeastern United States. Of these participants, 26 were male, and nine were female. The mean age was 21.0 years ($SD = 1.2$), and the academic classification included freshman ($n = 6$), sophomore ($n = 1$), junior ($n = 16$), senior ($n = 6$), master's/doctorate ($n = 5$), and "other" ($n = 1$). Musicians reported varying primary instruments, including brass ($n = 16$), percussion ($n = 2$), piano ($n = 1$), strings ($n = 3$), voice ($n = 5$), and woodwinds ($n = 8$). Approximately half the participants had piano experience ($n = 18$), and among them, the mean years of piano lessons was 6.39

(*SD* = 4.85). Thirty-three reported classical music lesson experience (*M* = 7.0 years, *SD* = 4.4), and of these, one participant reported 10 years of jazz lessons, and one participant reported 1 year of jazz lessons. Of all participants, jazz lesson experience ranged from zero to 15 years (*M* = 2.0, *SD* = 2.8). Thirty-two participants reported experience in classical ensembles (*M* = 9.6 years, *SD* = 4.1), and 28 participants reported experience in jazz ensembles (*M* = 5.2 years, *SD* = 4.0).

Stimuli

Interval identification test. We developed a shortened version of a previously designed test for use (Stambaugh & Nichols, 2020). The previous test included 33 items that were presented aurally to participants. In that study, a Cronbach's α coefficient of .88 indicated adequate internal consistency for a piano timbre interval identification test among a population of undergraduate music education majors. Next, we used an application of the Spearman-Brown formula to determine that as few as 11 of the original 33 items could be used to maintain an acceptable reliability level of .70. Therefore, in the current study, we included 11 ascending and 11 descending intervals from the original test, incorporating a pitch range of C2 to B5, which represented items of low, medium, and high difficulty levels in the previous study and included both ascending and descending melodic presentations. Using Qualtrics software on a Macbook Pro laptop with built-in speakers, participants heard 1,500 ms of white noise followed by Pitch 1 for 750 ms and then Pitch 2 for 750 ms. Tones used in the interval identification test used the "grand piano" timbre in Finale software. We randomly presented the order of the intervals to each participant, ranging from m2 to M7, and participants answered from a multiple-choice list of all possible intervals.

Melodic dictation test. Consistent with tonicization protocols used by Nichols et al. (2018), the tasks for the melodic dictation test were presented after hearing a tonic–dominant dyad twice using the pitches G4 and D5 presented melodically (G–D–G–D). Item difficulty measures were not reported in the Nichols et al. study; thus, one goal was to replicate a task resulting in a floor effect for musical working memory in the original study where melodic triad presentations began outside the dyad and ended within the dyad. The purpose of the current test was not to investigate working memory but, rather, to explore item difficulty in the context of other melodic tasks. To accomplish this goal, we generated a series of tasks that was theorized to progress from easy to difficult. The first three tasks required participants to dictate one pitch, two pitches, and three pitches, respectively. For the fourth task, participants heard a triad performed melodically and were asked to notate the last note of the triad. We included a distractor task using the standard operation span approach (Engle, 2002) in which we asked participants to state aloud whether the triad was major or minor before notating a response—the task that generated the floor effect observed by Nichols et al. The current test included eight different subtests, each containing three items (see Figure 1). For each item, we first presented the G–D–G–D dyad melodically; then we asked participants to identify pitches by notating them on staff paper. Participants heard identical stimuli.

Task 1 (Single Pitches)

Subtest 1: Notate one pitch within G-D dyad. (3 items)
Subtest 2: Notate one pitch outside G-D dyad. (3 items)

Task 2 (Two Pitches)

Subtest 3: Notate two pitches within G-D dyad. (3 items)
Subtest 4: Notate two pitches where the first is outside the G-D dyad, and the second is within. (3 items)

Task 3 (Three Pitches)

Subtest 5: Notate three pitches within G-D dyad. (3 items)
Subtest 6: Notate three pitches where the first is outside the G-D dyad, and the last is within. (3 items)

Task 4 (Distractor Task)

Subtest 7: Notate the last pitch of a triad where the first is outside the G-D dyad, and the last is within (includes distractor task). (3 items)
Subtest 8: Same as Subtest 7, except multiple triads heard per item. First item had 2 triad presentations, second item had 3 triad presentations, and third item had 4 triad presentations. Participants waited until last triad was heard before notating responses. (3 items)

Figure 1. Melodic dictation test format divided by task and subtest.

Procedure

Participants completed the tests individually in a quiet office after completing Institutional review board–approved consent procedures with a researcher or a research assistant present. First, participants completed the interval identification test using online Qualtrics software. We instructed participants to adjust the computer speakers to a comfortable volume during a practice item and informed them that they were allowed to listen to each interval one time only. Next, participants completed the melodic dictation test. We used PowerPoint slides that advanced automatically on a timer. Participants were instructed to "notate what you hear" and "please do not hum." Before each item, participants were presented with a tonicization slide, which included audio and visual cues of the G–D–G–D dyad. Next, each prerecorded task was played aloud (without notation), and a slide instructed participants to "Please write down the pitch(es) you heard." Participants wrote their responses on staff paper with a pencil. For the final task, the slide text instructed participants to state aloud "major" or "minor" upon hearing the triad before writing the pitches, which served as a distractor function. One of the researchers or research assistants monitored participants and verified that they spoke aloud during the distraction task. Finally, participants completed a questionnaire addressing demographic characteristics and their previous musical experience.

Results

Preliminary Analyses

Interval identification. We presented 22 melodic intervals, 11 ascending and 11 descending. We scored responses dichotomously as incorrect (0) or correct (1), and mean scores were calculated, resulting in a scale of 0 to 1. Participant means ranged from 0.36 to 1.0 (M = 0.76, SD = 0.18), and four participants achieved a perfect score. A Cronbach's α value of .834 indicated acceptable internal consistency reliability, and the deletion of no single item resulted in a gain in α. We considered an α higher than .70 to be "good" (Allen & Yen, 2001); thus, we chose to proceed with calculating difficulty levels. Test items ranged in difficulty from 0.17 to 1.0. Three of the 22 items had a 1.0 pass rate, meaning every participant scored it accurately; these three very easy items (Allen & Yen, 2001) included two major second intervals and one minor second interval.

Melodic dictation. We calculated a reliability value of Cronbach's α = .932 for the melodic dictation test. The deletion of no single item resulted in a gain in α. The difficulty levels of the individual items in this test ranged from .34 to .89 (M = .58, SD = .27), representing items of easy, medium, and difficult pass rates (Allen & Yen, 2001). Additionally, we examined test characteristics also at the task level (see Figure 1) by taking the mean of each subtest for analysis, and those difficulty levels ranged from .48 to .73. The performance means (and standard deviations) for each subtest were as follows: Subtest 1, M = .70 (SD = .35); Subtest 2, M = .73 (SD = .34); Subtest 3, M = .49 (SD = .38); Subtest 4, M = .48 (SD = .39); Subtest 5, M = .72 (SD = .22); Subtest 6, M = .56 (SD = .36); Subtest 7, M = .55 (SD = .31); Subtest 8, final pitches from two triad presentations, M = .59 (SD = .44); Subtest 8, three triads, M = .50 (SD = .38); Subtest 8, four triads, M = .54 (SD = .36).

Primary Analyses

We conducted a two-stage hierarchical regression analysis to examine the extent to which interval identification served as a predictor of melodic dictation accuracy while controlling for selected demographic variables. Before conducting the regression analysis, we tested for the assumptions of multiple regression analyses (Tabachnick & Fidell, 2007). Visual inspections of scatterplots suggested linear relationships, and normality plots of standardized residuals indicated normally distributed errors and homoscedasticity. Standardized univariate skewness and kurtosis values fell between the range of ±1.0 SD and were nonsignificant ($p >$

.05), suggesting univariate normality. Multivariate normality was established by nonsignificant results of Small's (1980) omnibus test of multivariate normality, $\chi^2(8) = 14.058$, $p = .081$. Durbin-Watson test results indicated a value of 2.628, signifying no concerns of autocorrelation. Tolerance statistics and variance inflation factors were also within recommended ranges (>0.5 and <10.0, respectively), indicating the absence of multicollinearity.

Pearson correlations among all variables are shown in Table S1 in the online supplemental material, and results of the regression analysis are summarized in Table 1. We entered the following independent variables into the first model as covariates: age, self-reported gender, number of years of private lessons on primary instrument, number of years on a secondary instrument, number of years of private piano lessons, and number of years of jazz ensemble experience. This model served as a significant predictor of melodic dictation accuracy and predicted 50.9% of the variance in melodic dictation scores, $F(6, 16) = 2.768$, $p = .048$, $R^2 = .509$, adjusted $R^2 = .325$. Within this model, years of private piano lessons served as a significant positive predictor of melodic dictation scores ($\beta = .520$, $r_{st} = .763$), and all other variables were nonsignificant.

In the second model, we added the mean interval identification score as an independent variable. The purpose was to isolate the effect of interval identification while controlling for the influence of demographic covariates included in the previous model. Results indicated that this model significantly increased in predictive power ($\Delta R^2 = .289$, $p < .001$) and explained 79.9% of the variance in melodic dictation scores, $F(7, 15) = 8.506$, $p < .001$, $R^2 = .799$, adjusted $R^2 = .705$. Within this second model, the only significant predictor of melodic dictation was interval identification scores ($p < .001$). Both the β weights and structure coefficients indicate that interval identification ($\beta = 1.143$, $r_{st} = .955$) served as the strongest predictor of melodic dictation accuracy.

Next, we focused on differences in melodic dictation performance between high, medium, and low scorers on the interval identification test. We divided the participants into three groupings based on their interval identification test scores (one third represented in each scoring group—high, medium, and low). Analysis of variance results indicate that performance on short melodic dictations varied by task type, $F(3, 32) = 16.832$, $p < .001$, $\eta_p^2 = .345$. Bonferroni-adjusted pairwise comparisons indicated that participants performed more accurately dictating one pitch than two ($p < .001$), dictating two pitches than three ($p < .001$), and dictating one pitch than the final note from a triad with a distractor task ($p = .006$). However, the analysis also indicated a task by ability interaction based on grouping by lower, mid-, and upper performing groups, $F(6, 29) = 2.224$, $p = .047$, $\eta_p^2 = .122$, where there was no significant difference on melodic dictation between participants in the low- and mid-performing groups ($p = .08$) on interval identification (see Figure 2). For these college musicians, overall performance ranged from 0.13 to 1.0, $SD = 0.27$. Performance on each task, in the order they were presented, was .72 (one pitch), .49 (two pitches), .64 (triad), and .55 (triads with distractor task).

Table 1. Hierarchical Regression Predicting Melodic Dictation Scores.

Model	R^2	Adjusted R^2	ΔR^2	Variable	*b*	*SE*	β	*r*	r_{st}
1	.509	.325	.509*	Constant	0.297				
				Age	0.027	.040	.132	−.023	−.032
				Gender	−0.238	.124	−.417	−.270	−.378
				Years of private lessons	−0.010	.012	−.173	.188	.263
				Years on second instrument	0.014	.010	.285	.515	.721
				Years of private piano*	0.025	.011	.520	.545	0.763
				Years of jazz ensemble	−0.015	.014	−.246	−.010	−.014
2	.799	.705	.289***	Constant	−0.872				
				Age	0.025	.027	.122	−.023	−.026
				Gender	0.019	.099	.033	−.270	−.302
				Years of private lessons	−0.008	.008	−.147	.188	.210
				Years on second instrument	0.009	.007	.185	.515	.576
				Years of private piano	0.008	.008	.171	.545	.610
				Years of jazz ensemble	0.001	.010	.017	−.010	−.011
				Interval identification***	1.143	.246	.761	.854	.955

Note. Melodic dictation = mean score on melodic dictation test; age = age in years; self-reported gender (male [coded as 1] or female [coded as 2]); years = number of years of experience; interval identification = mean score on the interval identification test. Structure coefficients (r_{st}) were calculated as the correlation between the predictor and the criterion divided by the multiple correlation coefficient (Courville & Thompson, 2001).
*$p < .05$. ***$p < .001$.

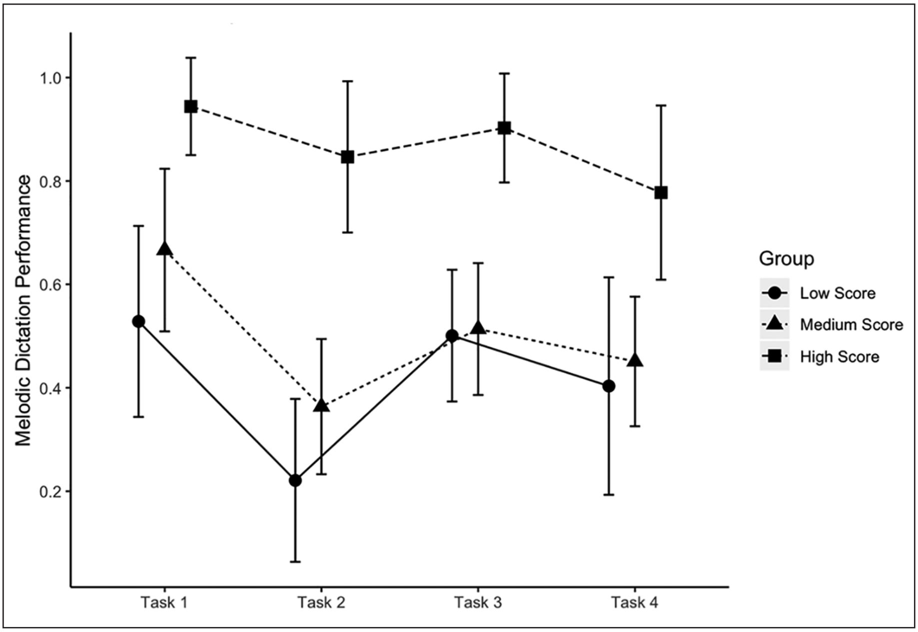

Figure 2. Melodic dictation by interval identification performance (brackets represent 95% confidence intervals).

Discussion

The purpose of this study was to investigate possible predictive relationships between interval identification and melodic dictation performance on tasks where students identify short pitch spans after a brief tonicization. The task required participants to notate one, two, and then three pitches after tonicization of the tonic and dominant pitches. In each task, participants first heard pitches that began within the tonic triad and then heard pitches that began outside of it (see Figure 1 for description of the tasks). Finally, participants notated the last pitch in a triad after completing a distraction task, stating whether the triad heard was major or minor in tonality. We theorized this presentation to present a sequence of tasks of increasing difficulty, but the results indicated the tasks were not progressively more difficult. Still, performance across tasks differed overall, and indices of item difficulty and internal consistency suggest this test to be an acceptable measure of melodic dictation.

Performance on these dictation stimuli could be theorized to differ from melodic dictation in applied settings, where musicians may attempt to recall or notate longer pitch sequences, including full melodies. Our goal was not to replicate the general experience of recalling and notating typical melodies with repeated listenings but to determine whether undergraduate musicians are proficient at identifying even the first note of a melodic fragment (thus our one-note presentation task following a brief tonicization). For these participants, performance differed significantly on the four tasks. This result might be explained by the use of a dyad to establish the key, where pitches above the first five scale degrees are more easily related to the dominant pitch (fifth scale degree). If only the tonic pitch—or even tonic chord presented harmonically—were presented, different outcomes could be expected.

In addition to the melodic dictation test, participants completed a measure of melodic interval identification. Cronbach's α coefficient indicated this test of 22 items was internally consistent; therefore, we explored a linear relationship between performance on the two tests. We chose a hierarchical model in which demographic variables related to prior musical experience were first entered as covariates, followed by interval identification, which significantly increased the prediction value. In addition to variables in the first level of the model, the model including interval identification predicted 79.9% of the variance in melodic dictation scores—28.9% of which could be uniquely predicted by interval identification. The contribution of interval identification to this model was significant and interpreted to represent a large portion of the variance explained by the model (Cohen et al., 2003). However, other musical variables not tested in this model could eclipse the contribution of interval identification or reduce the role of interval identification in predicting the melodic dictation of short melodic fragments.

We interpret this predictive relationship to suggest that interval identification may be a contributing skill for melodic dictation of short fragments beyond that which can be attributed to demographic variables such as age, self-reported gender, and experience. Perhaps aural interval identification is related to the skill of pattern recognition, which has been shown to positively influence melodic dictation (Paney & Buonviri, 2014). This novel finding can be used to support the teaching and learning of interval identification, specifically the skill of identifying longer pitch sequences by ear. Indeed, musicians' ability to "know" the pitches that are heard is a critical skill for identifying or notating sequences. Moreover, the ability to identify pitches is critical for the musicianship skill of error detection, where musicians identify inaccurate pitches in relationship to the accurate ones. Teachers and conductors are dependent on this skill as they design pedagogical techniques for addressing the inaccuracies they detect.

The relationship between interval identification—commonly understood as a measurement of the distance between two pitches—and melodic dictation of short fragments leads us to examine the nature of interval identification and melodic dictation generally: Is performing a dictation of two pitches similar to the task of interval identification? Dictation generally requires referencing a tonal center, whereas interval identification does not, except in instances where musicians rely on scale degree locations to determine the interval rather than actual distance between pitches. The results of this study suggest a strong relationship between melodic dictation and interval identification, and future studies could include direct comparisons of interval identification and melodic dictation of two-note fragments. Identifying short pitch sequences like those used in this melodic dictation test is an exercise similar to longer dictation tasks in which pitches must be not only identified in reference to a key center but also related to interval identification tasks where two pitches must be identified in absence of a key center. The latter task of interval identification could be theorized to be easier because no tonicization is required, and thus, no key center is maintained in working memory. Interval identification could also be theorized to be more difficult because no cues or clues are gained from scale degree locations in a key center.

Melodic dictation success is dependent on the listener's ability to accurately determine the first pitch heard. If the first pitch is notated incorrectly, subsequent pitches in the sequence are likely to be inaccurately notated—perhaps maintaining the same pitch relationships (intervals). Our participants only heard each fragment once, representing a first-attempt approach where musicians were asked to identify pitches upon hearing them. Perhaps performance on either test—melodic dictation or interval identification—would be higher given repeated listening, and this must be acknowledged because in applied settings, musicians sometimes have an opportunity to "hear it again."

In a previous study, the melodic dictation task with a distractor yielded a floor effect for working memory compared to other musical and nonmusical tasks among college musicians (Nichols et al., 2018). In the present study, these test items were well distributed in terms of test difficulty, supporting their use in future studies as a measure for short melodic fragments and tonicization. Although working memory was not a part of the research questions that guided this work, it may be suggested that working memory capacity in the previous study was low because not all college-age musicians are capable of accurately detecting the majority of pitches in a short melodic fragment—at least when also employing a distractor task. When the task cannot be completed, variables such as working memory capacity cannot be measured.

The greatest limitation in explaining these melodic dictation tasks may be that reports on difficulty levels/performance generally is so heavily dependent on the specific items/intervals chosen. Different items and intervals and starting notes would have yielded different results. Still, we acknowledge that interval identification appears to be a strong predictor for melodic dictation where sequences begin close or further away from the tonic pitch—at least in short fragments. We chose to give the interval identification first and then the melodic dictation test without balancing for order effects; we believed the first test could prime participants for performance on the second, and this could be true for using the opposite order. Following conventional test design, we began with melodic dictation items believed to be easier. First taking the dictation test could theoretically affect performance on interval identification given the repeated reinforcement of certain pitches used to establish the key center. In this case, persistence in working memory of the G–D dyad might have implications for the difficulty levels of certain intervals relative to other intervals. We chose what we believed to be a conservative approach to give participants the interval identification test first so as not to prime the G–D dyad in participants prior to the interval test, and this choice may have affected our results.

Tonicization—potentially more than stimuli characteristics such as instrument timbre—may affect performance on melodic dictation. Possibly some participants did not perform accurately on some items because the tonicization method used here (tonic-dominant dyad) was unfamiliar to them, or possibly they would have benefited from a procedure whereby they exercised agency over the tonicization process: playing the pitches themselves on their own instrument or a piano, being encouraged to hum the note rather than silently audiating, or playing a tonicization for as long as made them comfortable. Our approach allowed for the establishment of a tonic home without suggesting a particular tonality (major vs. minor) because no third scale degree was included in the tonicization stimulus, which has also been used in similar studies involving the dictation of short melodic fragments (e.g., Nichols et al., 2018).

Given the variety of tonicization procedures used by researchers previously (e.g., Buonviri, 2015b, Demorest, 1998; Killian & Henry, 2005; Nichols, 2016; Nichols & Wang, 2016), researchers in the future must make a choice for what procedure to use for establishing the key center. Our purpose was not to compare tonicization methods but to quickly and effectively establish the key so that testing could be performed. Sometimes in music performance settings or in dictation classes, no tonicization is given. Still, the question of the capacity for musicians to maintain a tonal center is important. Did the presentation of a three-pitch sequence for the triad result in a retonicization of sorts whereby participants were oriented to a particular triad or even tonal center? This effect could be greater when a major triad is presented in a three-note sequence rather than a minor triad. Finally, what other features, such as melodic contour, decrease or increase interference with the tonal center in musicians reading or notating Western music?

Whether conceptualized as a tonic note or a tonic triad (per Parncutt, 2011), to what degree do musicians hear pitches in reference to a tonal center, and for how long is the tonal center maintained? Farbood (2016) suggested a quick decay for the initial tonal center after modulations in listeners, but less is known about musicians who are hearing, synthesizing, and making music. Some of our participants may not have scored well in a test of longer melodic dictations because they were not always successful in identifying the first pitch(es) of a pitch sequence. Previous research on melodic dictation may rest on the assumption that musicians are successful at identifying or contextualizing the initial pitches, which may not often be true. Such questions regarding tonicization are central to the topic(s) in this article and to many of studies cited in the review of literature. Further research should continue to explore the complex processes of melodic dictation as a central issue for music education.

There are broader implications for music teaching and learning based on the results of this study. Students may require greater support in establishing the key in aural skills exercises as well as identifying initial pitches in relation to a key center prior to developing melodic dictation skills for longer sequences. Furthermore, students may not be successful in melodic dictation tasks if they are unable to establish the initial pitches, which cannot be presumed to occur more easily than later pitches in a long sequence.

Finally, results of this study indicate that interval identification seems to be important as a foundational skill relied on for more complex tasks such as melodic dictation, and these data suggest additional implications for music educators. Given that nearly 30% of the variance in melodic dictation scores could be accounted for by interval identification scores, even while controlling for various demographic characteristics and factors related to prior musical experience, this finding underscores the need to hone interval identification abilities when preparing students for success during complex cognitive tasks like melodic dictation and error detection (Larson, 1977). Also, our finding that Task 3 (triads) may have been easier than Task 2 (dyads) overall suggests that developing interval identification or melodic dictation skills may not progress from identifying one note, then two notes, and then more but, rather, a more complex progression involving specific pitch sequences in relation to a tonal center. Music educators and music learners should therefore consider that the development of melodic dictation abilities likely does not follow a simplistic one-note, then two-note, and then three-note process. Rather, students likely experience success on complex dictation tasks due to various contextual cues, tonal cues (major vs. minor), or triad-specific identifiers that aid in melodic dictation processing. Future research might explore difficulty indices for specific melodic dictation sequences as well as tonicization strategies for musicians with varying levels of experience because this information would be helpful in further understanding the factors that influence the important skill of melodic dictation.

Acknowledgments

The authors wish to thank Scott Atchison for assistance in creating Figure 2.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Bryan E. Nichols https://orcid.org/0000-0002-8970-3284
D. Gregory Springer https://orcid.org/0000-0001-7923-726X

Supplemental Material

Table S1 is available in the online version of this article at https://doi.org/10.1177/00224294211011962.

Notes

1. Before recruiting participants, we conducted an a priori power analysis using G*Power software to determine a minimum sample size for a multiple regression analysis. Results of the power analysis indicated a minimum sample size of 34 (input parameters included an effect size of $R^2 = .2$, $\alpha = .05$, and a power level [$1 - \beta$] of .80). We also conducted a post hoc power analysis using the pwr package in R using our observed data, and results indicated an observed power level of .94.
2. Some authors (Courville & Thompson, 2001; Ziglari, 2017) have cautioned against the sole use of beta weights when interpreting a variable's contribution to a model and have

suggested that structure coefficients also be used when making these interpretations. Consistent with these recommendations, we have included both beta weights (β) and structure coefficients (r_{st}) because "the two sets of coefficients—β weights and structure coefficients—provide us with a more insightful stereoscopic view of dynamics within our data. Interpreting only beta weights . . . usually will not yield sufficient understanding of all the relevant dynamics in our data" (Courville & Thompson, 2001, p. 245). Structure coefficients are calculated as the correlation between the predictor (independent variable) and the criterion (dependent variable) divided by the multiple correlation coefficient (Courville & Thompson, 2001).

References

Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Waveland Press.

Buonviri, N. O. (2014). An exploration of undergraduate music majors' melodic dictation strategies. *Update: Applications of Research in Music Education*, *33*(1), 21–30. https://doi.org/10.1177/8755123314521036

Buonviri, N. O. (2015a). *Building better dictation skills*. Rowman & Littlefield.

Buonviri, N. O. (2015b). Effects of a preparatory singing pattern on melodic dictation success. *Journal of Research in Music Education*, *63*(1), 102–113. https://doi.org/10.1177/002249415570754

Buonviri, N. O. (2017). Effects of two listening strategies for melodic dictation. *Journal of Research in Music Education*, *65*(3), 347–359. https://doi.org/10.1177/0022429417728925

Buonviri, N. O. (2019). Effects of silence, sound, and singing on melodic dictation accuracy. *Journal of Research in Music Education*, *66*(4), 365–374. https://doi.org/10.1177/0022429418801333

Buonviri, N. O., & Paney, A. S. (2015). Melodic dictation instruction: A survey of Advanced Placement music theory teachers. *Journal of Research in Music Education*, *63*(2), 224–237. https://doi.org/10.1177/0022429415584141

Clements, P. A. (2003). *The effect of phrase expansion recognition on melodic dictation accuracy* [Publication No. 3085587] [Doctoral dissertation, University of Missouri-Kansas City]. ProQuest Dissertations and Theses Global.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: β is not enough. *Educational and Psychological Measurement*, *61*(2), 229–248. https://doi.org/10.1177/0013164401612006

Demorest, S. (1998). Improving sight-singing performance in the choral rehearsal: The effect of individual testing. *Journal of Research in Music Education*, *46*(2), 182–192. https://doi.org/10.2307/3345622

Demorest, S., & Clements, A. C. (2007). Vocal pitch production and perception among middle school male singers. *Journal of Research in Music Education*, *55*(3), 190–203. https://doi.org/10.1177/00224294075500302

Demorest, S., Nichols, B. E., & Pfordresher, P. (2018). The effect of focused instruction on young children's singing accuracy. *Psychology of Music 46*(4), 488–499. https://doi.org/10.1177/0305735617713120

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Farbood, M. M. (2016). Memory of a tonal center after modulation. *Music Perception*, *34*(1), 71–93. https://doi.org/10.1525/mp.2016.34.1.71

Graves, J. E., & Oxenham, A. J. (2017). Familiar tonal context improves accuracy of pitch interval perception. *Frontiers in Psychology*, *8*, Article 1753. https://doi.org/10.3389/fpsyg.2017.01753

Halpern, A. R. (1984). Perception of structure in novel music. *Memory & Cognition*, *12*(2), 163–170. https://doi.org/10.3758/bf03198430

Hoppe, K. M. (1991). *The melodic dictation strategies of musicians, and common pitch and rhythm errors* [Publication No. 9128254] [Doctoral dissertation, University of Texas]. ProQuest Dissertations and Theses Global.

Killian, J., & Henry, M. (2005). A comparison of successful and unsuccessful strategies in individual sight-singing preparation and performance. *Journal of Research in Music Education*, *53*(1), 51–65. https://doi.org/10.1177/002242940505300105

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368. https://doi.org/10.1037/0033-295X.89.4.334

Larson, R. C. (1977). Relationships between melodic error detection, melodic dictation, and melodic sightsinging. *Journal of Research in Music Education*, *25*(4), 264–271. https://doi.org/10.2307/3345267

Madsen, C. K., & Staum, M. J. (1983). Discrimination and interference in the recall of melodic stimuli. *Journal of Research in Music Education*, *31*(1), 15–31. https://doi.org/10.2307/3345107

Nichols, B. E. (2016). Task-based variability in children's singing accuracy. *Journal of Research in Music Education*, *64*(3), 309–321. https://doi.org/10.1177/0022429416666054

Nichols, B. E., & Wang, S. (2016). The effect of repeated attempts and test-retest reliability in children's singing accuracy. *Musicae Scientiae*, *20*(4), 551–562. https://doi.org/10.1177/1029864916638914

Nichols, B. E., Wöllner, C., & Halpern, A. R. (2018). Score one for jazz: Working memory in classical and jazz musicians. *Psychomusicology: Music, Mind, & Brain*, *28*(2), 101–107. https://doi.org/10.1037/pmu0000211

Norris, C. E. (2003). The relationship between sight singing achievement and melodic dictation achievement. *Contributions to Music Education*, *30*(1), 39–53.

Paney, A. S. (2016). The effect of directing attention on melodic dictation testing. *Psychology of Music*, *44*(1), 15–24. https://doi.org/10.1177/0305735614547409

Paney, A. S., & Buonviri, N. O. (2014). Teaching melodic dictation in advanced placement music theory. *Journal of Research in Music Education*, *61*(4), 396–414. https://doi.org/10.1177/0022429413508411

Paney, A. S., & Buonviri, N. O. (2017). Developing melodic dictation pedagogy: A survey of college theory instructors. *Update: Applications of Research in Music Education*, *36*(1), 51–58. https://doi.org/10.1177/8755123316686815

Parncutt, R. (2011). The tonic as triad: Key profiles as pitch salience profiles of tonic triads. *Music Perception*, *28*(4), 333–336. https://doi.org/10.1525/mp.2011.28.4.333

Pembrook, R. G. (1984). *The interference of the transcription process and other selected variables on memory during melodic dictation* [Publication No. 8427320] [Doctoral dissertation, Florida State University]. ProQuest Dissertations and Theses Global.

Pembrook, R. G. (1986). Interference of the transcription process and other selected variables on perception and memory during melodic dictation. *Journal of Research in Music Education*, *34*(4), 238–261. https://doi.org/10.2307/3345259

Powell, J. (2013). *Melodic dictation: An examination of strategy and time usage by college music students* [Publication No. 1543369] [Master's thesis, Southern Methodist University]. ProQuest Dissertations and Theses Global.

Sink, P. E. (1983). Effects of rhythmic and melodic alterations on rhythmic perception. *Journal of Research in Music Education*, *31*(2), 101–113. https://doi.org/10.2307/3345214

Small, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Journal of the Royal Statistical Society, Series C*, *29*(1), 85–87. https://doi.org/10.2307/2346414

Stambaugh, L., & Nichols, B. E. (2020). The relationships among interval identification, pitch error detection, and stimulus timbre by pre-service teachers. *Journal of Research in Music Education*, *67*(4), 465–480. https://doi.org/10.1177/0022429419885931

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Pearson.

Taggart, C. C., & Taggart, B. F. (1994). Sightsinging systems: A survey of American colleges and universities. *Southeastern Journal of Music Education*, *6*, 194–209.

Williams, D. B. (1975). Short-term retention of pitch sequence. *Journal of Research in Music Education*, *23*(1), 53–66. https://doi.org/10.2307/3345203

Wöllner, C., & Halpern, A. (2015). Attentional flexibility and memory capacity in conductors and pianists. *Attention, Perception & Psychophysics*, *78*(1), 198–208. https://doi.org/10.3758/s13414-0150989-z

Wuthrich, C. E., & Tunks, T. W. (1989). The influence of presentation time asynchrony on music interval perception. *Psychomusicology*, *8*(1), 31–46. https://doi.org/10.1037/h0094233

Ziglari, L. (2017). Interpreting multiple regression results: β weights and structure coefficients. *General Linear Model Journal*, *43*(2), 13–22. https://doi.org/10.31523/glmj.043002.002

Author Biographies

Bryan E. Nichols is an assistant professor of music at Penn State University. His research interests include singing accuracy and participation and music teacher education.

D. Gregory Springer is an assistant professor of music education at Florida State University. His research interests include music perception, music performance evaluation, and music teacher education.

Submitted December 9, 2019; accepted March 10, 2021.