To *Doh* or not to *Doh*: The Comparative Effectiveness of Sightsinging Syllable Systems

Mary Jo Lorek & Randall G. Pembrook

The selection of a proper syllable system to facilitate and improve sightreading has been a topic of debate among sightsinging instructors for years. While teachers and students seem to have preferences for various systems (preferences which they typically will defend vehemently when challenged), often the favored approach is a function of which system was learned first. Killian (1991) mentions the "this works for me school of thought."[1] Rogers (1996) labels the state of sightsinging instruction in the United States as being a " . . . pedagogical hodge-podge."[2] At times the choice is based on a careful review of the perceived merits inherent in a given system. Another manner of selecting a system, wherein choices are based on information attesting to the effectiveness of a particular approach, has been virtually unavailable due to the sparsity of research relating to syllable systems. In 1959, Bentley insisted that "proof must be given from experiments in which each system is used with equal enthusiasm by equally good teachers, and with groups of pupils of comparable age, ability and interest."[3] Unfortunately, few studies of this type have been reported.

The selection of an effective syllable system is an overwhelming task considering the number of choices. In the spring of 1989,

---

[1]Janice N. Killian, "The Relationship Between Sightsinging Accuracy and Error Detection in Junior High Singers," *Journal of Research in Music Education* 39 (1991): 217.

[2]Michael Rogers, "The Jersild Approach: A Sightsinging Method from Denmark," *College Music Symposium* 36 (1996): 149.

[3]Arnold Bentley, "Fixed or Movable Do?" *Journal of Research in Music Education* 7/2 (Fall, 1959): 167.

Pembrook and Riggins (1990) surveyed nearly 300 college sightsinging teachers and found that at least ten "syllable" systems were in use.[4] These included well-known systems such as fixed and movable *doh* but also some less popular systems such as inflected letter names, in which singers sing the note name (e.g., "G-sharp"), and "fixed number" systems, in which 1 stands for middle-c, 2 for c-sharp, and so forth.

The related literature concerning sightsinging syllable systems seems to divide into two large categories and one smaller one. These include: (1) articles describing the history and use of various systems; (2) articles proposing the strengths and weaknesses inherent in the systems; and (3) data-based studies detailing the results when various systems are compared.

The first type of literature includes articles by Dean (1965), Collins (1979), Shephard and Jordan (1884), Killam (1987), and the aforementioned study by Pembrook and Riggins (1990). Jerry Dean's thesis (1965) entitled "A History of Solmization and a Comparative Study of American Pitch Verbalization Methods"[5] includes a survey of 120 colleges and universities throughout the United States, in an effort to determine syllable system preferences, particularly with respect to neutral syllable systems. His findings indicated that note names, neutral syllables, scale degree numbers and movable *doh* systems were employed frequently and approximately equally while fixed *doh* systems (with and without chromatic inflection) were used less regularly.[6] An important dissertation in this area by Irma Collins (1979) reported movable *doh*, numbers, and "lah" to be the most frequently used systems.[7] Shephard and Jordan (1984) suggested that Western listeners and by extension, Western sightsingers

---

[4]Randall G. Pembrook & H. Lee Riggins, "Send Help! Aural Skills Instruction in U.S. Colleges and Universities," *Journal of Music Theory Pedagogy* 4/2 (1990): 231-241.

[5]Jerry M. Dean, A History of Solmization and a Comparative Study of American Pitch Verbalization Methods. Unpublished Master's thesis, University of Texas, Austin, 1965.

[6]Ibid., 43.

[7]Irma H. H. Collins, Current Attitudes and Trends in the Teaching of Sight Singing in Higher Education, Unpublished doctoral dissertation, Temple University, 1979.

may be predisposed to systems like movable *doh* or numbers which are based on hearing by scale degrees.[8] They state:

> Judgements of relations between tones in a sequence derived from the major diatonic scale by systematic distortions (either equalizations or stretching of intervals) reveal that Western listeners interpret the tones relative to that musical scale even when they try to judge the purely physical relations between the tones.[9]

In addition to documenting the wide array of systems in use, Pembrook and Riggins found that numbers were used for student sightsinging most frequently by the 279 college sightsinging teachers who responded to the survey. Movable *doh* and neutral syllable systems were in a virtual tie as the second most-used system.[10]

In the second category of literature, authors have attempted to evaluate the inherent qualities of syllable systems to determine the most effective approach. In 1959, Bentley proposed several arguments for movable *doh*, concluding that it is "a most useful means of aural training and teaching music reading."[11] Walt Multer (1978) offered a detailed discussion of the advantages and disadvantages of fixed and movable *doh*.[12] He felt that every musician should understand and use both. Consequently, according to Multer, it does not really matter whether one system is preferred since it is incumbent to know and use fixed and movable *doh*.[13] In 1991, Timothy Smith presented an exposition of fixed and movable systems, and their respective advantages and disadvantages. Smith proposed that movable *doh*, with *doh* as the tonic in minor was the superior system in the greatest number of areas.[14] This point was rebutted

---

[8]Roger N. Shepard and Daniel S. Jordan, "Auditory Illusions Demonstrating that Tones Are Assimilated to an Internalized Musical Scale," *Science* 226 (December, 1984): 1333-1334.

[9]Ibid., 1333.

[10]Pembrook & Riggins, 234.

[11]Bentley, 167.

[12]Walt Multer, "Solmization and Musical Perception," *Theory and Practice* 3/1 (1978).

[13]Ibid., 47.

[14]Timothy A. Smith, "A Comparison of Pedagogical Resources in Solmization Systems," *Journal of Music Theory Pedagogy* 5/1 (1991): 1-23.

by Houlahan and Tacka (1992) in an ensuing issue. They defended the movable *doh* system with *la* as the tonic in minor contending that this system helps develop a better understanding of aural relationships before depending on visual input and theoretical rules.[15]

Steve Larson (1993), also joining the debate about *doh*- versus *la*-based minor, examined the cognitive processes of sightsinging and transcribing to point out that certain syllable systems are "better suited" than others for specific purposes (i.e., singing in major, singing in minor, transcribing, etc.). He suggests that a syllable system should not be chosen cavalierly.[16]

Articles which have compared the effectiveness of syllable systems based on students' sightsinging capabilities while using them are few in number, particularly when compared to descriptive, historical, and philosophical articles. In 1946, Buchanan compared the scores of two groups of students—one using fixed *doh* and one using movable *doh*. He concluded that amateur singers can be expected to exhibit more improvement using movable rather than fixed syllables.[17] Other pertinent published, empirical studies could not be found.

In light of this paucity of data, a series of studies was initiated at the researchers' university to test the comparative effectiveness of the most frequently used sightsinging syllable systems as reported by Pembrook and Riggins. Every effort was made in those studies to ensure that potential problems identified by Bentley in 1959 [equivalency of groups, competence of instructors, and an enthusiasm for the approach][18] were adequately addressed. The remainder of the paper details four studies that compared solfege effectiveness in various conditions.

---

[15]Michael Houlahan & Philip Tacka, "The Americanization of Solmization: A Response to the Article by Timothy A. Smith, 'A Comparison of Pedagogical Resources in Solmization Systems,'" *Journal of Music Theory Pedagogy* 6 (1992): 137-151.

[16]Steve Larson, "The Value of Cognitive Models in Evaluating Solfege Systems," *Indiana Theory Review* 14/2 (1993): 115.

[17]Walter Buchanan, Comparison of Fixed and Movable Solfege in Teaching Sight Singing From Staff. Unpublished doctoral dissertation, University of Michigan, Ann Arbor, 1946.

[18]Bentley, 167.

STUDY #1

The researchers felt that the initial study should compare some type of movable system (e.g., movable *doh* with *doh* as tonic in major and *la* as tonic in minor or scale degree numbers) to a fixed system (e.g., fixed *doh* or inflected note names) to some type of "control" if possible. Because three sections of freshman ear training classes were available, a decision was made to compare all three approaches in the initial study. Though the Pembrook and Riggins survey had indicated that numbers were used more frequently than movable *doh*, movable *doh* was selected as the methodology from the first category for two reasons. First, the survey conducted in 1989 had indicated that of all the teachers using a single approach, movable *doh* was the most popular methodology[19] and second, historically, a review of the literature seemed to indicate that the solfege debate has focused on fixed *doh* and movable *doh* rather than fixed *doh* and numbers.

Initial screening indicated that very few incoming freshmen at the authors' university were entering with command of a particular syllable system for singing. Therefore, subjects were randomly assigned to one of three groups. Subjects in group 1 (movable *doh*) were instructed to use solfege wherein *doh* indicated the tonic in major and *la* in minor. Subjects in group 2 (fixed *doh*) used *doh* for C regardless of the key, *di* for C-sharp, and so forth. Subjects in group 3 were required to do all vocalizing during sightsinging class on a neutral syllable. The neutral syllable "lah" was used because of its tendency to encourage proper breathing and singing mechanics. Therefore, regardless of the key or the specific pitches, subjects in group 3 always sang on "lah."

To insure the validity of the study, various factors that could account for group differences at the outset of the study were analyzed in an attempt to attain balanced groups. All subjects were tested on a sightsinging pretest. Responses, which were cassette-recorded, were analyzed by the ear training faculty and divided into high, medium, and low accuracy groupings. Equal numbers

---

[19]Pembrook and Riggins, 234.

of subjects from each category were assigned to the three experimental groups with a total of 18 students per group (N=54 subjects). Very few subjects indicated a predilection for a particular syllable system. Those that did were randomly distributed among the three groups. Students were also equally divided among groups with respect to high school grade point average in an attempt to balance for intelligence and academic ability.

Attempts were also made to guarantee equivalency of instruction and equal motivation among instructors. First, two teachers were selected who believed in the superiority of movable *doh* and fixed *doh*, respectively. The third instructor felt that syllable systems were not essential in teaching college students to sightsing. That individual taught the neutral syllable section (group 3). All were professors who had taught ear training for several years at the college level. Graduate teaching assistants were not used.

All three instructors used the same texts (Ottman's *Music for Sight Singing* , Berkowitz, Frontrier and Kraft's *A New Approach to Sight Singing*, and Hall's *Studying Rhythm*) and supplementary materials (e.g., choral scores and melodies composed by the instructors). Lesson plans were scripted so that instructors practiced the same examples the same number of times. However, the exact feedback of the instructors was not outlined since it would be based on student responses and therefore could not be anticipated.

Classes met for 50 minutes, twice a week during a 15 week semester. The semester during which the initial study was conducted was the first in a four semester undergraduate ear training sequence, including dictation, sightsinging and error detection. The university did not have a computer lab for additional practice at the time of the initial study. Each instructor missed only one of the 30 scheduled classes. All classes met from 9:00 until 9:50 a.m. on Tuesdays and Thursdays. This prohibited information about the testing examples from being circulated from class to class on testing days.

Students were tested three times for written dictation and error detection and three times for sightsinging during the semester. Only the scores on the sightsinging tests were originally used to evaluate the effectiveness of the three teaching techniques. Students in all classes were expected to use the appropriate syllable system dur-

ing testing, but were evaluated by the classroom teacher only for pitch and rhythmic accuracy, including maintaining a steady beat with no hesitations. Grading of syllables would have created an extra scoring deficit variable for two of the three groups. All sightsinging exams were taped and transcribed for objective evaluation. Students were allowed to study the sightreading examples for up to 2 minutes before attempting them, although most students in all groups began before the preparation time had expired. Examples simulated but did not duplicate examples from the sightsinging textbooks.

At the end of the semester, results from students who had missed no more than three classes and had taken all three sightsinging exams were analyzed. This included 17 from the fixed *doh* class, 15 from the movable *doh* class, and 13 from the neutral syllable class. Results indicated no significant difference between the three groups.[20] Averages of the percentage correct of the three sightsinging tests were calculated, and means for three groups were as follows: fixed *doh* = 91.51%, "lah" = 90.53%, and movable *doh* = 89.78%. Slight differences in the averages were attributed statistically to chance rather than the effect of the main factor—syllable system. Out of curiosity, one additional analysis of overall scores was conducted, which included students' dictation test results. Again, no significant difference was found between the treatment groups.[21]

STUDY #2

The first study was repeated for three reasons. First, it was felt that this initial research required replication and a larger data pool before results could be considered valid. Second, the researchers decided that even more controls should be used to insure group equivalency. Finally, it was speculated that while no appreciable difference due to syllable system could be identified after 15 weeks (1 semester) it was possible that the systems could affect students to a greater degree over the course of an entire year.

---

[20] $F[2,42]=.21$, $p=.81$.
[21] $p=.73$.

Forty-two more subjects at the same institution were divided among three classes and received instruction using either movable *doh*, fixed *doh*, or "lah." To create three equal groups, students again were assigned to one of three classes based on high school grade point average and entering sightsinging ability as determined by an entrance test. In addition, attempts were made in study #2 to divide students according to their applied instrument type and their scores on a melodic perception test which assessed students' ability to detect minute changes in frequency. Though different instructors were used in this study because of re-assignment of duties within the ear training department, all three instructors were again experienced college ear training teachers with more than five years of experience. Instructors met weekly to review the scripted lesson plans and select and develop materials to augment the texts (Ottman's *Music for Sight Singing* and Benward's *Ear Training: A Technique for Listening*. Killam had identified these as the most frequently used materials for sightsinging and ear training courses in her survey.[22])

Students were again tested six times per semester—three times in sightsinging and three times in dictation. Procedures and grading methods were identical to those described in Study #1. A statistical analysis was again calculated, first for data from the fall semester and then for scores from the second term. Again, results indicated no significant difference in Fall semester scores as a function of syllable system.[23] Scores were nearly identical to the previous study. Means for the groups were as follows: fixed *doh*= 89.38%, "lah" = 88.62% and movable *doh* = 86.88%. Even though the order of the second and third groups switched (Fixed *doh*=92.52, movable *doh*=92.29, and "lah"=89.14), the means from the second semester still indicated no significant differences. Based on these two studies conducted over three semesters with 87 subjects during the first year of university ear training study it would appear that the selection of a particular, single syllable system is not critical to students' short-term sightsinging success.

---

[22]Rosemary Killam, et al, "Survey and Results: Most-Used Theory Texts in US Colleges and Universities," The University of North Texas, 1987.

[23]$F[2,39]=.34$, $p=.71$.

STUDY #3

After both of the initial studies indicated that fixed *doh*, movable *doh* and "lah" produced statistically equivalent results within the timeframe of the study, it was decided that a sightsinging syllable system for all music classes at the university would be adopted based on factors other than its immediate efficacy. Because survey results indicated that movable systems were more prevalent than fixed ones in ear training instruction at other institutions, it was decided that a movable system represented potentially the most valuable option for students who might be transferring to other undergraduate programs, or completing graduate music degrees elsewhere. Therefore, all members of the next entering freshman class received instruction using movable *doh*. By the fourth semester of instruction, when classroom material included modulating melodies and melodies exhibiting chromatic inflection, students began complaining about the "usefulness" of the movable *doh* system. Study #3 originated from those complaints.

Keeping in mind that the original studies had indicated no difference in the effectiveness of fixed *doh*, movable *doh*, and "lah" during the first year, a decision was made to allow students to abandon the movable *doh* syllable system while being tested on selected melodies during the fourth semester. Because these students had not been exposed to fixed *doh*, it was posited that an implementation of that system during the final semester might be frustrating because of the proactive interference of previously learned movable solfege syllables and, therefore, counterproductive for the students. However, instructors encouraged students to use the neutral syllable "lah" when chromatic and modulating melodies made movable *doh* syllables appear unwieldy. The researchers were curious to discover (1) if students would prefer this approach and, more importantly, (2) whether it would facilitate, inhibit, or have little effect on their sightsinging capabilities under those circumstances.

In previous studies, it was essential to create equivalent groups and instruction in order to eliminate group variance as a confounding variable. In Study #3 subjects sang half of the melodies on "lah" and half using movable *doh*; therefore, subjects served as their own control. Subjects were randomly assigned to sing "lah" or movable *doh* first. Using subjects as their own control on two melodies for

each testing, however, necessitated the creation of similar melodies to be sightread under those two conditions. All melodies used under the two conditions had the same number and types of intervals, the same types and number of chromatic alterations and were identical regarding rhythmic complexity and total length as measured by number of notes. Furthermore, melodies were alternated so that half of the subjects sang "lah" on one melody and half sang "lah" on the second melody.

Thirty-seven students took part in the study. The testing included two melodies featuring chromatic patterns (one sung on "lah" and one sung on movable *doh* syllables) and two melodies featuring modulating passages (again sung using the two syllable systems). The location of the key change was indicated on the test paper for modulating melodies, so that students singing movable *doh* would know on which note to change syllable orientation. Students were given two minutes to review the passage as in Studies #1 and #2. All melodies sung by the subjects were cassette-recorded and then transcribed and analyzed at a later date for greater reliability of accurate assessment.

The mean for the melodies sung using movable *doh* was 85.41% and the mean for the melodies sung on "lah" was 84.16%. These scores were compared and again no statistically significant difference was found.[24] Sixteen of the 37 students being tested professed a preference for movable *doh* which they had used for nearly two years. Eleven stated no preference and six preferred neutral syllables while four had no opinion. The scores of the 22 who professed a preference were analyzed comparing their sightsinging accuracy when using the preferred and non-preferred systems. Results of the comparison revealed no significant difference in accuracy under the two conditions (preferred system mean=85.68%; non preferred mean=82.23%).[25] Students were statistically equal in accuracy when using the system of choice and when using an "imposed" system, though the difference in accuracy came closest to statistical significance of any comparison described in this paper.

---

[24] $t[36]=.90$, $p=.37$.
[25] $t[21]=1.69$, $p=.11$.

STUDY #4

Because the results of the previous studies had indicated no significant difference in effectiveness among the three broad categories of a movable system, a fixed system, and a control in developing sightsinging accuracy, the results from the Pembrook and Riggins study were analyzed for the next most logical direction for research. Their figures had indicated that numbers was the most frequently used sightsinging system. Therefore, a decision was made to compare two popular approaches within the movable category—numbers vs movable *doh*.

The ensuing incoming class (37 students) was divided into three classes with still more attention given to group equivalency. Classes were structured to insure statistically equivalent numbers with relation to gender, high school grade point average, entering sightsinging ability as determined by a pretest, applied instrument experience, and years of piano study. In spite of testing *two* systems, students were divided into *three* classes to provide good student/teacher ratios. Because it was anticipated that two fall classes would be combined during the second semester, these two sections were taught using numbers as the sightsinging syllable system, while the other class was conducted using movable *doh*.

The same textbooks were used as in Study #2 (Ottman and Benward/Kolosick). During this particular semester, students were tested five times during the semester, with only two sightsinging tests (because of time constraints of individual testing) and three dictation tests. Again, scores were not significantly affected by the use of a particular syllable system.[26] Means were 93.17 for 22 numbers students in two classes and 94.10 for 15 movable *doh* students in the third class.

DISCUSSION

The results of these four studies seem to indicate that within a college setting, with music majors over a one- or two-semester period,

---

[26] $t[35]=.49$, $p=.63$.

no particular voiced system, be it movable *doh*, numbers, fixed *doh*, or a neutral syllable, produces significantly higher sightsinging scores than another. The results must be viewed carefully, however. Regardless of the syllable system a student may be attempting to use, it is impossible to know what may be taking place at some cognitive level. The fact that the neutral syllable group consistently had the largest range of scores on tests and consequently the largest standard deviations may suggest that successful students in that group were using some type of subvocalized system and unsuccessful students, who had the lowest scores of the three groups in Study #2, were in fact singing on "lah" without placing the pitch to be sung on some type of movable or fixed grid. This observation is largely speculative and should be investigated in future research.

While some in the neutral syllable group admitted a non-vocalized use of syllables, it should be noted that something representing nearly the opposite also occurred. During the course of the four studies, it was not at all unusual to witness students in the fixed *doh*, movable *doh*, and numbers groups who sang the melodies correctly according to pitch and rhythm but did not vocalize the proper syllables while doing this. Apparently, for those students, the syllables represented an abstract series of sounds which were required while singing the test examples but which were not needed or used to locate the pitch.

Another aspect of the study worthy of discussion relates to the "control group" which used "lah." As Butler and Lochstampfor (1994) point out "conditions are not ideal for mounting rigorous experimental studies of aural training methodology in the most realistic situation, which is when actual aural training classes are used in the study."[27] A true control group, wherein students were totally exempted from the curricular requirements of four semesters of ear training courses, was not implemented in these studies. Ethical considerations relating to purposefully creating a group of musicians whose aural skills development might be retarded as a result of that experimental methodology prevented the researchers from

[27]David Butler and Mark Lochstampfor, "Bridges Unbuilt: Aural Training and Cognitive Science," *Indiana Theory Review* 14/2 (Fall, 1993): 7.

implementing this design. Nonetheless, it would be interesting to see the results from a group of individuals who are restricted from developing aural skills in the traditional theory curriculum and instead are limited to other avenues such as sightreading in their ensembles, listening in applied lessons, composition or arranging classes, and so forth. The results of such a study might be greatly reinforcing to ear training teachers if the students were found to have comparatively inferior aural skills, or they might challenge the field to re-evaluate ear training pedagogies if that group of students developed aural skills equal to students in traditional programs.

Two of the three teachers in the first study could not continue as instructors for the second study and were replaced. Because the results of that second study replicated those from the initial one (even the rank of high to low means of the first semester of Study #2 was identical to Study #1), it was theorized that teacher effect was minimized by the standardization of materials used in class and the scripted/timed lesson plans. This factor would appear to bolster the validity of the conclusions indicating no difference in sightsinging development as a function of the syllable system employed.

FUTURE RESEARCH

Three future research studies seem evident from these initial investigations. First, additional research needs to be conducted where combinations of approaches are used and compared to other combinations or single approaches (e.g., a section using fixed *doh* and numbers vs. movable *doh* or a section using movable *doh* and numbers vs. numbers only). The research by Pembrook and Riggins indicates that most ear training teachers use more than one method.[28] Treatments featuring a combination of approaches might increase the external validity of future studies.

Second, the results of this series of studies must be evaluated within its context. All studies were conducted with music majors in their freshman or sophomore years. As entering freshmen, most were not proficient with any type of syllable system; those that

[28]Pembrook and Riggins, 234.

were had the option of placing out of the beginning courses by means of a proficiency test. This study did not attempt to answer the question of whether any one system (fixed *doh*, movable *doh*, numbers, or *la*) is most effective in developing sightsinging abilities if started at an early age. Some elementary music education methodologies (e.g., the Kodaly approach) strongly advocate the use of movable *doh* systems. Future research should address the efficacy of such a system when initiated in preschool or kindergarten and continued throughout the first two years of college. Educators interested in an answer to that longitudinal issue may not be persuaded by information from this study which was limited to college freshmen and sophomores. Nonetheless, those who teach undergraduates in college ear training classes may see applications in the current research, especially if their students exhibit limited previous solfege experience as the subjects in these four studies did.

Finally, future studies should attempt to gather more information regarding student attitudes toward solfege systems. Do novices object to using unfamiliar solfege syllables initially and then come to appreciate them later? Surprisingly in Study #3, when asked to sing "lah" in their fourth semester of instruction in movable *doh*, students who were the most vocal in criticizing the awkwardness of movable *doh* in chromatic and modulating passages were the most resistant to change. A few of the freshman assigned to the "lah" class in Study #2, and the numbers classes in Study #4 expr essed a desire to learn movable *doh*, although many students assigned to the "lah" class expressed relief in avoiding solfege. If syllable systems are equally effective in creating sightsinging accuracy, perhaps it is not imperative that instructors or department chairs choose the methodology. Students might choose based on personal preferences. The results in Study #3, when students chose their own system, came closest to statistically significant differences at traditional error levels.

All of these issues are important to the process of life-long development of aural skills. The pursuit of the research in this article was undertaken as an attempt to better understand the effectiveness of current methodologies and to stimulate thinking toward the development of a "most effective system," if indeed that is possible.