Final Ritardandi and the Expression of Musical Emotion

Ronald S. Friedman
University at Albany, State University of New York

According to Juslin (2001), final ritardandi constitute part of the acoustic code employed by musicians to communicate two distinct emotional states: sadness and tenderness. To test this proposition, in two experiments, participants were exposed to a set of hymns that were modified in mode and/or tempo to primarily express either happiness, sadness, or tenderness. In addition, the inclusion of final ritardandi was experimentally manipulated such that these timing variations were either present or absent in the hymn stimuli. Participants were then asked to rate the emotions expressed by each variant of the hymns. In line with Juslin (2001), results revealed that when final ritardandi were included, expressively sad music was perceived as conveying more sadness, whereas expressively tender music was perceived as conveying more tenderness. Inclusion of ritardandi did not heighten the expression of happiness in music that was in a major key nor diminish perceived emotional expressivity, depending on the specific emotional state at issue.

As extensively reviewed by Juslin (2001, 2013; Juslin & Laukka, 2003), a considerable amount of research has been conducted to determine the particular combinations of acoustic cues associated with the perceived expression of distinct qualities of emotion, including anger, fear, happiness, sadness, and tenderness. However, the vast majority of these studies have assessed perceptions of emotional expressiveness in samples of music in which expressive cues, including interpretative variations implemented by performers, have been allowed to covary freely. For instance, in several pioneering studies, Juslin and his colleagues (Gabrielsson & Juslin, 1996; Juslin, 1997a) had instrumental musicians repeatedly play the same melody, each time shaping their performance to express one of a set of distinct emotions. A separate set of listeners subsequently rated the emotions expressed by each rendition. Results suggested that listeners were quite successful at decoding the emotional intentions of the performers. However, these experiments could not establish the extent to which any particular expressive cue was essential to the acoustic code used to communicate a given emotion state.
To address this limitation, researchers have also conducted factorial experiments in which a wide range of expressive cues (e.g., variations in articulation, sound level, or tempo) have been orthogonally manipulated using synthesized musical stimuli (e.g., Eerola, Friberg, & Bresin, 2013; Juslin, 1997b; Juslin & Lindström, 2010; Scherer & Oshinsky, 1977). This has enabled researchers to estimate the unique contribution of these cues to the communication of various emotional states. However, due to the vast number of distinct cues that may contribute to emotional expression, as well as the practical impossibility of manipulating every conceivable cue in a single factorial design, the additive and/or multiplicative effects of a great many expressive cues have yet to be systematically assessed using this “synthetic approach” (Eerola et al., 2013).

Among these insufficiently investigated expressive cues is the final ritardando, an interpretative variation that entails the gradual slowing of tempo at the boundaries of musical sections. Although their use may vary across musical styles (Gabrielsson & Juslin, 1996), final ritardandi are routinely implemented by performers in common practice music (e.g., Friberg & Sundberg, 1999; Palmer & Hutchins, 1996). Here, they reflect and may serve to communicate the performer’s interpretation of the hierarchical structure of the music (Penel & Drake, 2004). In addition, they have been hypothesized to heighten musical tension, intensifying listeners’ pleasure when a musical passage ultimately comes to a close (Huron, 2006; cf. Kleinsmith, Friedman, & Neill, 2016; 2017). Most pertinent to the matter at hand, Juslin (2001; p. 315; Figure 14.2) has suggested that final ritardandi also form part of the acoustic code used by musicians to convey two distinct feeling states: sadness and tenderness. In support of this proposition, he cites work by Gabrielsson and Juslin’s (1996) study does not establish whether final ritardandi contribute to listeners’ perceptions of expressed sadness and tenderness independent of other musical cues with which they tend to covary (e.g., legato articulation, low sound level).

The present study was designed to rectify this limitation and to more directly test whether final ritardandi uniquely contribute to the musical communication of sadness and tenderness, two of the predominant emotions expressed by music (Juslin, 2013). Experiment 1 focused on the role of final ritardandi in conveying sadness. Here, participants were asked to rate the emotion expressed by each of a set of hymns that were modified in mode (major vs. minor) and tempo to sound expressively happy versus sad. Final ritardandi were implemented in half of the hymn stimuli. In line with Juslin’s (2001) summary of the acoustic codes associated with different emotions, it was predicted that expressively sad hymns would sound even sadder when they ended with a final ritardando. Experiment 2 partially replicated the design of the first experiment and additionally included hymn stimuli that were set at the same, relatively slow tempo as those meant to convey sadness, yet that were transposed to a major key, thereby imbuing them with core features associated with the expression of tenderness (Eerola et al., 2013). Again, consistent with Juslin’s research synthesis (2001), it was predicted that these hymns would be perceived as expressing even more tenderness when they ended with final ritardandi. Finally, it was predicted that hymns that included final ritardandi would not be perceived as expressing greater happiness, thereby confirming that these timing variations selectively contribute to the expression of specific types of emotion (sadness and tenderness), as opposed to heightening the perceived quantity of emotion, irrespective of its quality (cf. Bhata et al., 2011).

**Experiment 1**

**METHOD**

**Participants.** A minimum sample size was chosen based on an a priori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), which indicated that at least 34 participants would be required to achieve a statistical power of .80 to detect a medium-sized interaction effect ($d = .50$). Ultimately, 38 undergraduate students (22 female, 16 male) at the University at Albany were recruited for the study from a psychology research methods course. Approximately forty-four percent of participants reported having had some formal music training and 47% reported having played an instrument for at least one year ($M = 1.76$ years, $SD = 0.94$). (There were no effects of music training on emotion expression ratings. Therefore, this variable will not be discussed further). Participants who completed the experiment were compensated for their time by receiving extra exam credit in the course. Participants were recruited in class for a study on “responses to melodies” to take approximately 20-30 min and were run in groups of up to 7 at computer workstations in a small classroom.

In this, and in the subsequent experiment to be reported below, written informed consent was obtained from each participant. All studies were approved by the Institutional Review Board of the University at Albany.
Materials. Five different hymns (see Appendix) were selected from a digitized corpus (www.pdhymns.com). Hymns were employed as stimuli because they have a melodic and harmonic structure that is representative of the musical experience of individuals familiar with Western common practice tonal music (Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). Moreover, they are freely available on the internet in multiple formats (printed score, MIDI, Finale), allowing for easy access and modification. For the present purposes, all hymn selections were freely chosen by the author with the sole stipulation that they end with an authentic (V-I) cadence, “...the most strongly closural cadence [in]...common-practice tonal music...” (Temperley, 2011, p. 1) and are available in Finale format. Finale 2014 software was used to generate four versions of each hymn. In two versions, the hymn proceeded a tempo to the final chord. In the other two versions, a ritardando was implemented during the closing harmonic progression using Finale’s Human Playback Function, a suite of software utilities that permits the application of a range of naturalistic temporal and dynamic variations to the musical score. To maximize experimental control, no other expressive variations were added to the hymns. In addition, to manipulate whether the hymns were generally perceived as expressing happiness versus sadness, two versions of each hymn were heard in a major key (either A major or D major), as originally composed, and set at a relatively fast tempo (120 bpm), whereas the remaining two versions were transposed into a minor key (either A minor or D minor) and set at a relatively slow tempo (80 bpm). Sound files providing examples of all versions of one of the hymn stimuli, crossing the factors of Ritard (absent vs. present) and Cue Combination (major/fast vs. minor/slow) appear in the Supplementary Materials in the online version of this paper.

To measure perceptions of the emotion conveyed by the hymn stimuli, a variant of the Self-Assessment Manikin (SAM) technique (Lang, 1980) was employed. Here, participants viewed a graded series of 9 humanoid figures ranging in their facial expressions from very sad to very happy and were asked to choose the figure that best represented the emotion expressed by the music they just heard. Scores ranged from 1 to 9, with lower scores reflecting greater sadness (higher scores reflecting greater happiness) and 5 representing a neutral midpoint. The SAM is among the most popular self-report-based measurement tools in emotion research (Betella & Verschure, 2016). Its pictorial format is easy to process, promotes participant engagement, and may enhance construct validity by better capturing “...the unarticulated nature of emotions...” than do scales based on verbal labels (Rubin, Rubin, Graham, Perse, & Seibold, 2009, p. 338).

Procedure. Participants were each seated in front of a computer. After providing written informed consent, they were asked to put on a provided pair of headphones and to follow the on-screen instructions. Participants were subsequently presented with each of the hymn stimuli in 1 of 2 fixed-random orders, generated with the stipulation that two versions of the same hymn were never heard sequentially. After listening to each hymn, participants rated “how much sadness or happiness it expressed” on the aforementioned 9-point SAM scale, using the numeric keypad.

RESULTS AND DISCUSSION

Descriptive statistics for Experiment 1 indexed by experimental condition are displayed in Table 1. Emotional expression ratings were subjected to a 2 (Cue Combination: major/fast vs. minor/slow) X 2 (Ritard: absent vs. present) X 2 (Order) mixed-model ANOVA. As expected, this analysis revealed a main effect of Cue Combination, $F(1, 36) = 186.96, p < .0001, \eta^2 = .84$; confirming that major key/faster tempo hymns were perceived as happier than minor key/slower tempo hymns. It also revealed a main effect of Ritard, $F(1, 36) = 8.20, p < .008, \eta^2 = .18$; suggesting that hymns ending in final ritardandi were perceived as sadder than those without them. Critically, these effects were qualified by a Cue Combination X Ritard interaction, $F(1, 36) = 6.41, p < .02, \eta^2 = .13$. To decompose this interaction, a series of Bonferroni-adjusted post hoc comparisons were computed on the condition means. These showed that major key/faster tempo (i.e., expressively happy) hymns ending in final ritardandi were perceived as no happier than those without ritardandi, as part of a separate study constructively replicating earlier work by Kleinsmith et al. (2016, 2017), participants were then asked to relisten to the same set of hymns in the same order as before, yet this time, to rate their liking for each hymn stimulus. Findings using the measure, which are not pertinent to the hypotheses assessed at present, will be presented in a subsequent report.

### Table 1. Experiment 1: Descriptive Statistics for Ratings of Sadness/Happiness of Musical Stimuli Indexed by Experimental Condition

<table>
<thead>
<tr>
<th>Ritard</th>
<th>Major Key/Fast Tempo</th>
<th>Minor Key/Slow Tempo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>$M = 6.53, SD = 0.83$</td>
<td>$M = 4.04, SD = 1.27$</td>
</tr>
<tr>
<td>Present</td>
<td>$M = 6.49, SD = 0.76$</td>
<td>$M = 3.66, SD = 1.11$</td>
</tr>
</tbody>
</table>
METHOD

Overview. Given this evidence confirming that final ritardandi enhance the musical expression of sadness, Experiment 2 was designed to conceptually replicate the effect and additionally assess whether final ritardandi also contribute to the expression of tenderness. Prior research has shown that the combination of cues used to express tenderness heavily overlaps with that used to express sadness (Juslin, 2001). Expression of both emotions is associated with a relatively slow tempo; however, given that it is a positively valenced feeling state, tenderness is more likely to be conveyed using a major rather than a minor key (Eerola et al., 2013). Therefore, in Experiment 2, the expressively happy (major key/faster tempo) hymns from Experiment 1 were slowed in tempo and either implemented with or without ritardandi. Participants in a newly appended experimental group were asked to rate the tenderness, as opposed to the happiness/sadness expressed by all hymns. It was predicted that the inclusion of final ritardandi would lead participants in this group to selectively perceive the major key/slower tempo hymns as expressing more tenderness.

Participants. A minimum sample size was estimated by computing a power analysis using G*Power (Faul et al., 2007). This indicated that at least 16 participants would be required to achieve a statistical power of .80 to detect an effect of the same size observed for the Cue Congruency X Ritard interaction in Experiment 1 ($d = .77$). This minimum sample size was then doubled to 32 to account for the addition of a between-participants factor (Rating: sadness vs. tenderness). Ultimately, 51 undergraduate students (36 female, 14 male, 1 other) at the University at Albany were recruited for this study from a psychology research methods course. Approximately 68% of participants reported having had some formal musical training and 68% reported having played an instrument for at least one year ($M = 2.06$ years, $SD = 0.95$). (There were no effects of music training on emotion expression ratings. Therefore, this variable will not be discussed further). Participants who completed the experiment were compensated for their time by receiving extra exam credit in the course. Participants were recruited in class for a study on “responses to melodies” to take approximately 20-30 min and were run in groups of up to 7 at computer workstations in a small classroom.

Materials and procedure. The procedure of Experiment 2 was identical to that of Experiment 1 with the following exceptions: First, the major key hymns from Experiment 1 were reduced in tempo from 120 bpm to 80 bpm to enable the musical expression of tender feelings. Second, two between-participants Rating conditions were created, one in which participants listened to and rated the happiness/sadness of the hymns (as in Experiment 1), and another in which they listened to and rated the tenderness expressed by the hymns. Second, as there were no effects of order in Experiment 1, hymns were presented in only a single fixed-random order. Finally, to ensure that participants understood the term “tenderness,” for participants who rated the tenderness of the hymns, instructions were provided on screen defining the term as “refer[ring] to feelings of affection, care, or warmth.” Moreover, to reinforce this conceptualization, two images were presented on screen, a photo of a mother kissing her baby and a line drawing of a woman embracing a young child. As no Self-Assessment Manikin scale exists for tenderness, participants were simply asked to rate tenderness (“How much tenderness was expressed by the music you just heard?”) on a 9-point Likert scale anchored at “1” (minimum tenderness) and “9” (maximum tenderness).

RESULTS AND DISCUSSION

Descriptive statistics for Experiment 2 indexed by experimental condition are displayed in Table 2. Emotional expression ratings were analyzed separately for participants who rated the sadness versus the tenderness of the hymns.

| Table 2. Experiment 2: Descriptive Statistics for Ratings of Musical Stimuli Indexed by Experimental Condition |
|-------------------------------------------------|---------------------------------|---------------------------------|
|                                                | Major Key/                      | Minor Key/                      |
|                                                | Slow Tempo                      | Slow Tempo                      |
| Ritard                                          | $M$ | $SD$ | $M$ | $SD$ |
| Absent                                          | 5.61 | 1.10 | 4.02 | 1.12 |
| Present                                         | 5.66 | 1.28 | 3.61 | 1.18 |
| Sadness/Happiness Rating (n = 23)               |                      |                               |
| Tenderness Rating                               |                      |                               |
| Absent                                          | 5.79 | 0.89 | 4.50 | 1.52 |
| Present                                         | 6.37 | 1.01 | 4.60 | 1.69 |
tenderness of the hymns. Beginning with the former, as in Experiment 1, SAM-based happiness/sadness ratings were subjected to a 2 (Cue Combination: major/slow vs. minor/slow) X 2 (Ritard: absent vs. present) mixed-model ANOVA. This analysis revealed a main effect of Cue Combination, F(1, 22) = 55.20, p < .0001, \( \eta^2_p = .71 \), suggesting that minor key/slower tempo hymns were perceived as sadder than major key/slower tempo hymns. This main effect was qualified by a Cue Combination X Ritard interaction, F(1, 22) = 10.14, p < .005, \( \eta^2_p = .32 \). To decompose this interaction, a series of Bonferroni-adjusted post hoc comparisons were computed on the condition means. These showed that major key/slower tempo hymns ending in final ritardandi were perceived as expressing approximately the same amount of happiness as those without ritardandi, \( p > .68 \), whereas minor key/slower tempo hymns ending in final ritardandi were perceived as significantly sadder than those without ritardandi, \( p < .003 \). These results replicate those of Experiment 1, confirming that final ritardandi amplify the sadness conveyed by music in which a sad expressive quality has already been instilled by means of more fundamental cues, here, the use of a minor key and relatively slow overall tempo. As final ritardandi did not heighten the perception of happiness when the same music was transposed into a major key, the findings also accord with those of Experiment 1 in suggesting that these interpretative variations shape the perceived quality of emotion expressed, not merely its quantity.

As with happiness/sadness ratings, tenderness ratings were also subjected to a 2 (Cue Combination: major/slow vs. minor/slow) X 2 (Ritard: absent vs. present) mixed-model ANOVA. This analysis revealed a main effect of Cue Combination, F(1, 27) = 27.69, p < .0001, \( \eta^2_p = .49 \), again suggesting that major key/slower tempo hymns were perceived as more tender than minor key/slower tempo hymns. There was also a main effect of Ritard, F(1, 27) = 11.87, p < .003, \( \eta^2_p = .30 \). These effects were qualified by a marginally significant Cue Combination X Ritard interaction, F(1, 27) = 3.56, p = .070, \( \eta^2_p = .12 \). To decompose this interaction, a series of Bonferroni-adjusted post hoc comparisons were computed on the condition means. These showed that minor key/slower tempo hymns ending in final ritardandi were perceived as expressing the same amount of tenderness as those without ritardandi, \( p > .57 \). However, major key/slower tempo hymns ending in final ritardandi were perceived as expressing significantly more tenderness than those without ritardandi, \( p < .002 \). These results support the prediction that final ritardandi amplify the musical expression of tenderness.

General Discussion

According to Juslin (2001), final ritardandi constitute part of the acoustic code employed by musicians to communicate two distinct emotional states: sadness and tenderness. To test this proposition, in two experiments, participants were exposed to a set of hymns that were modified in mode and/or tempo to primarily express either happiness, sadness, or tenderness. In addition, the inclusion of final ritardandi was experimentally manipulated such that these timing variations were either present or absent in the hymn stimuli. Participants were then asked to rate the emotions expressed by each variant of the hymns. In line with Juslin (2001), results revealed that expressively sad music was perceived as conveying more sadness, whereas expressively tender music was perceived as conveying more tenderness, when final ritardandi were included. Inclusion of ritardandi did not heighten the expression of happiness in music that was in a major key nor promote the expression of tenderness in music that was in a minor key. This suggests that final ritardandi do not generally heighten emotional expressivity and only amplify the emotional message already established by other cues, particularly those based on mode and overall tempo.

Although the present findings support Juslin’s (2001) descriptive model of the musical code used to communicate sadness and tenderness, they leave an essential question unaddressed: What is origin of this code? As discussed earlier, Juslin (2001) proposes that the acoustic cues utilized by musicians to express emotion are largely based on emulation of corresponding expressive cues in speech. However, he does not specifically discuss whether sad and tender speech are characterized by a relatively pronounced reduction in tempo at phrase boundaries. Indeed, in their comprehensive meta-analysis of the correlation between musical and speech-based emotional cues, Juslin and Laukka (2003) make no mention of the use of final ritardandi. Rather, they emphasize the use of a slower overall tempo, rather than the reduction of tempo at phrase boundaries, as a primary cue for the expression of both sadness and tenderness. As such, it remains unclear why final ritardandi might contribute to the musical code for these particular emotions and whether this contribution may find its origin in the vocal expression of emotion.

A likely possibility may be found in the link between a reduced rate of speech and diminished physical energy. Sadness is often characterized by a feeling of deenergization, which is linked to slower gross motor...
movement (Buyukdura, McClintock, & Croarkin, 2011) as well as slower action in the articulator muscles, leading to a slower speech rate (Ellgring & Scherer, 1996; Mundt, Snyder, Cannizzaro, Chappie, & Gerals, 2007). A reduction in vocal energy has also been associated with a drop in subglottal air pressure and reduced tension in the laryngeal muscles toward the end of an utterance, which can lead final speech segments to be more drawn out in time (Herman, 2000). Therefore, assuming that music communicates emotion by emulating how it is expressed in the voice, by diminishing tempo as sad music comes to a close, final ritardandi may be imitatively expressing the waning energy characteristic of feelings of despair or resignation.

How might this “vocal energy” hypothesis account for the use of final ritardandi to convey feelings of tenderness? Unlike in the case of sadness, there is little reason to assume that tenderness is associated with dwindling energy. A possible answer may be extrapolated from the work of Juslin (2001), who has posited that aspects of the musical code used to communicate emotion may arise not from hard-wired emotion expression “programs”, but from a learning process based on the way that the voice is utilized in social interaction. Here, Juslin (2001) notes that “[w]hen mothers talk to their infants . . . if they want to calm their infant, they reduce the speed and intensity of their speech and talk with slowly falling pitch contours” (p. 323). Presumably, such efforts to soothe others are often accompanied, if not impelled, by feelings of tenderness, opening the possibility that tender feelings may generally become associated with vocal deintensification. If so, the use of final ritardandi in the musical expression of tenderness may be at least partially rooted in its emulation of this feature of tender speech.

In sum, the results of the present study constitute the first experimental evidence that the final ritardando plays a supportive role in expressing at least two specific emotional states: sadness and tenderness. In doing so, it supports the cue utilization framework proposed by Juslin (2001) and advances understanding of a ubiquitous timing variation, the psychological function of which has received interest among music scholars. Further research is clearly needed to elucidate the boundary conditions and ultimate origins of these effects. However, the current results attest to the value of “using synthesized sound sequences that are systematically manipulated” (Juslin & Laukka, 2003; p. 804) to pinpoint the expressive role of a given musical cue. As such, they pave the way for analogous studies investigating how other cues, both temporal (e.g., average tempo variability, “sharpness” of duration contrasts) and non-temporal (e.g., average pitch height, sound level variability, timbre) distinctly, yet interactively, contribute to the rich language of musical emotion.

Author Note

I am grateful to Marielle Erdheim, Evelyn Karpel, and Nicholas Lathouris for their assistance in running these experiments.

Correspondence concerning this article should be addressed to Ronald S. Friedman, Department of Psychology, University at Albany, State University of New York, 1400 Washington Avenue, Albany, NY 12222. E-mail: rfriedman@albany.edu

References


### Appendix

#### List of Hymns used in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Hymn Name</th>
<th>Musical Key</th>
<th>Time Signature</th>
<th>Average Length (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faith of Our Fathers</td>
<td>A major/minor</td>
<td>3/4</td>
<td>31.59</td>
</tr>
<tr>
<td>For the Beauty of The Earth</td>
<td>A major/minor</td>
<td>4/4</td>
<td>32.38</td>
</tr>
<tr>
<td>Just Over In The Glory Land</td>
<td>A major/minor</td>
<td>4/4</td>
<td>25.13</td>
</tr>
<tr>
<td>More Holiness Give Me</td>
<td>D major/minor</td>
<td>12/8</td>
<td>31.84</td>
</tr>
<tr>
<td>When He Comes In Glory By and By</td>
<td>A major/minor</td>
<td>4/4</td>
<td>22.80</td>
</tr>
</tbody>
</table>