Facilitative Agency in Performance

Roger C. Graybill

NOTE: The examples for the (text-only) PDF version of this item are available online at:

<http://www.mtosmt.org/issues/mto.18.24.3/mto.18.24.3.graybill.php>

KEYWORDS: audiation, agency, analysis and performance, phenomenology, Beethoven, Edwin Gordon, David Lewin

ABSTRACT: This paper explores how a performer generates a cognitive infrastructure in support of a performance, and advocates for regarding such cognitive processes as a kind of agency. This notion of “facilitative agency,” which draws on Edwin Gordon’s work on audiation as well as David Lewin’s p-model in his “Music Theory, Phenomenology, and Modes of Perception,” is illustrated in an analysis of the opening theme from Beethoven’s Piano Sonata in G major, op. 14, no. 2, second movement. The paper closes by considering the implications of the above for the training of student musicians.

Received April 2018

Volume 24, Number 3, September 2018
Copyright © 2018 Society for Music Theory

[1] The last decade has seen a burgeoning of research on the embodied and gestural aspects of musical performance. Most of these studies address the observable aspects of performance, typically gleaning data from video recordings of performers and analyzing how the performer’s gestures correlate to events in the music.⁽¹⁾ Such observable kinesthetic aspects of musical performance seem particularly well suited to the present collection of papers on performer’s agency. Other studies—for instance, Doğantan-Dack 2011, Hatten 2004 and 2006, Mead 1999, Pierce 2007, and Tzotzkova 2012—have adopted a different angle, describing the subjective experience of performing. This article falls into the second category, with a focus on the performer’s internal processes rather than on what might be observed by another person. But I will be examining not the subjective physical experience of performing so much as the internal mental representations upon which the performer draws while performing. I will propose that this hidden realm is itself the site for a particular kind of agency. Through *facilitative agency*, as I’ll call it, the performer generates a cognitive infrastructure as a basis for her physical act of performance, with the latter exemplifying *primary agency*. Distinguishing between these two types of agency can be either a simple or vexingly complicated task, depending on one’s vantage point. The difference is clear if we consider the two agencies from the perspective of an observer, since only primary agency can be *seen* in its workings. But if we try to distinguish the two types as subjectively experienced by the performer, the picture is much less clear; indeed, any attempt to elucidate the difference bumps up against the well-known mind/body problem that has haunted philosophy since Descartes. The conclusion of this study will reconsider the relationship between facilitative and primary agency in some detail; for the present, I will assume that differentiating them is useful and intuitively appropriate.

[2] This essay will regard facilitative agency as a category of *audiation*, a word coined by the American music educator Edwin E. Gordon.⁽²⁾ Ear training instructors generally use the term to denote auditory imagery in the absence of a sounding musical signal—for instance, hearing a tune mentally, or reading a score while hearing it in one’s head. Such an ability to generate auditory imagery is obviously useful to a musician, but Gordon’s concept is far richer than is implied by this common understanding. Indeed, Gordon considers audiation to be nothing less than “the foundation of musicianship” (2008), and he assigns it the central role within his theory of musical development. Audiation ideally informs all our musical activities: “We audiate when listening to, recalling, performing, interpreting, creating, improvising, reading, or writing music” (2012, 4).

[3] But what exactly *is* audiation? Gordon defines it as a “cognitive process by which the brain gives meaning to musical sounds” (2008). Our ability to generate such meaning in turn depends on our awareness of the *context* for any musical phenomenon to which we are attending: “When listening to music we are at any given moment organizing in audiation sounds that were recently heard. We also predict, based on our familiarity with the tonal and rhythmic conventions of the music being heard, what will come next” (2008). Note that Gordon is describing two kinds of contexts here: the events within the piece itself, and stylistic norms outside the piece.

[4] It is this contextual component that distinguishes Gordon’s conceptualization of audiation from the common understanding cited earlier. Gordon is not so concerned with our ability to conjure up sound images from scratch, but rather with our processes for “assimilating and comprehending” music (2012, 3), for which contextual awareness is indispensable. This essay will reserve the term audiation for Gordon’s contextually-based interpretation of the word, while adopting Gary Karpinski’s more specific notion of “auralizing” (2000, 49) for our generation of sound images in the absence of actual sound.

[5] Since facilitative agency is a kind of audiation—more precisely, audiation in the service of performance—any attempt to describe its operation will need to address how the performer mentally represents and accesses multiple contexts while performing. Gordon provides little assistance here, for his aim is to construct a learning theory for developing audiation skills rather than to theorize about how it works in real time. His emphasis on contextual awareness, however, calls to mind another study that may assist us—that is, David Lewin’s (1986) application of Husserlian phenomenology to the listening process in his “Music Theory, Phenomenology and Modes of Perception.” Lewin models the experience of a listener who interprets a series of events with respect to multiple temporal contexts; for any given event, a relevant context normally includes either music heard prior to the event, or music that is yet to happen. Lewin relates the former to Husserl’s notion of “retention,” which refers to the “projections of remembered past times (and past durations) into my present consciousness” (329). A future-oriented musical projection resembles Husserl’s “protensions,” which Lewin describes as the “projections of future expectations into present consciousness” (329). An implication/realization dynamic frequently gives rise to such protensions, as when a dominant seventh chord leads us to expect a tonic chord.⁽³⁾ Note how Lewin describes the directionality of a protension: rather than stating that we project ahead to the future, Lewin says that our expectations of future events project into our phenomenological present.⁽⁴⁾ Thus both past experience *and* future expectation project into our present, a two-way flow that recalls Gordon’s description of audiation.

[6] Lewin’s formal perception model (which he designates as the *p-model*) suggests several ways for refining Gordon’s sketchy account of audiation in real time. **Example 1** illustrates Lewin’s p-model analysis of a passage from Schubert’s song “Morgengruss.”⁽⁵⁾ For each perception (“p” in the left column), Lewin lists an event (EV); a context (CXT) for that event; selected “perception-relationship” pairs that show how the perception in question bears on *other* perceptions; and selected statements about the perception in a certain “language” (in the present case in the form of musical examples, not shown here). Observe that several perceptions on the list involve retention (for instance, compare the EV and CXT columns for p2 and p3b), others involve protension (p4, p5, etc.), while yet others include both (p6b, p7a). Both types of projection in Example 1 exhibit properties that carry over into my audiatonal analysis later in this essay. First, each retention has a clear initiation point (e.g., for p2, this moment is the downbeat of m. 9), and I will maintain this degree of temporal specificity when describing a performer’s audiatonal retention of past events.

Strikingly, each of Lewin's retentions flows into and includes the event that stimulates the perception (e.g., p2 includes m12 in the CXT column), blurring the distinction between the retention and the event itself. My analysis will assume such blurring as well, though I will also include retentions of earlier self-standing passages that are not contiguous with the present event.

[7] Lewin's protensions present a mirror image of his retentions insofar as they *end* at a specified moment (with the exception of p5, which has an unspecified endpoint). However, these protensions do not extend very far—only a measure, in fact. Here the p-model needs to be adapted for our purposes. Lewin assumes that the listener's protensions derive from his expectations without regard to what he may know is actually coming up—in other words, as if he is experiencing the music for the first time—and this probably accounts, at least in part, for the brevity of the protensions. But this conceit of the naïve listener does not transfer well to a performer, who needs to know where she is heading over longer spans; in fact, she needs to *intend* some kind of future outcome. For this reason, I will assume that a performer is able to pretend longer contexts than shown in Lewin's diagram.

[8] But this in turn raises a question: to the extent a performer's projections of future events result from intention, can they really be regarded as protensions? Lewin suggests that the answer is no. He states that his p-model is poorly suited for performance since it assumes a binary distinction between the perceiver and the perceived—a distinction that Lewin feels is not operative for a performer (or composer) engaging in a creative act.⁽⁶⁾ To be sure, he does not specifically mention a performer's intention vis-à-vis future events, but rather the performer's musical *present*—that is, to the way in which she attends to “what-is-being-played-right-now” (377). Nevertheless, it seems reasonable to extrapolate from Lewin's argument that, in his view, the perceiver/perceived distinction collapses in the case of intentional future projections as well. If this is correct, it appears that Lewin does not regard such projections as objects of perception.⁽⁷⁾

[9] But one can concur with this view without denying that a performer has access to the same musical intuitions as a naïve listener who is protending future events—for instance, the kind of harmonic implication/realization dynamic described earlier. For a performer, however, such intuitions are unavoidably intertwined with intentionality. Here we encounter an intriguing paradox: a performer's protensions often arise spontaneously and involuntarily, whether through an innate human cognitive capacity (as [Larson 2012](#) argues with respect to musical “forces”) or audiatonal competence in a particular musical style (e.g., while hearing/playing a ii6 chord, anticipating a V-I). Thus characterized, protensions arise prior to any intentionality on the performer's part. Yet these very protensions inform and shape the performer's intentional commitment to the ongoing life of the musical passage.⁽⁸⁾

[10] Despite the fact that a performer's ability to “intend” a future raises a question about the nature of her protensions, the validity of this study does not depend on how we come to terms with that question. A performer must generate *some* sort of mental representation of her future intended action, and that anticipated context certainly bears on her present consciousness whether or not it is a perceptual object. Throughout the rest of this essay, I will continue to use the term “protension” for any projection of a future context, whatever its status as a perceptual object.

[11] In short, Lewin's p-model shows how a listener's—and with some adaptation, a performer's—present awareness is shaped by past and anticipated events, and in specifiable ways. This is no small benefit, given that a performer's subjective audiatonal experience can seem highly fluid and resistant to description. While the analytical portion of this study will not utilize Lewin's formal p-model with the degree of rigor suggested by Example 1, I will draw upon its broad assumptions for insights on how a performer audiates in the flow of time.

[12] The remainder of this paper explores the workings of facilitative agency within a brief passage of music, mm. 1–8 of the second movement from Beethoven's Piano Sonata in G major, op 14, no. 2. The ensuing discussion of the Beethoven passage falls into two portions. First, I consider how a performer might exercise facilitative agency in a real-time performance of the passage. The paradigmatic setting for such a performance is the public concert, but I define such “real-time” performance more broadly as any intentional play-through of the passage, whatever the setting.⁽⁹⁾

In the second portion, I describe ways in which a performer harnesses facilitative agency while preparing for a stage performance, with the practice room as the paradigmatic setting. I will propose that facilitative agency in the practice room entails *audiatonal play*, which in turn has significant implications for theory pedagogy.

Facilitative agency in real-time performance

[13] **Example 2** provides the Beethoven passage along with an audio clip of my own performance. Since I do not have direct access to the internal processes of other minds, I will necessarily be reflecting on my own audiatonal experiences as I perform the passage. I do so while acknowledging the near impossibility of rendering a complete and accurate account, given that much of this process must lie below consciousness.

[14] So let me start by asserting something of which I am reasonably sure: before I play each phrase, I hold in my awareness some kind of mental representation of the forthcoming phrase as a whole. And this representation has a gestalt-like quality; I experience it as an expressive shape, not merely a span of time. A physical analogy might be helpful here. In the following video clip (**Example 3**), I demonstrate the 2+2+4 phrasing with arm gestures.⁽¹⁰⁾ The sounding music in the clip should be interpreted as a stand-in for my internal auralization of the passage. Note that each arm gesture spans an arc from my right side to my left, or vice versa, and that I time each gesture to complete its arc as the phrase comes to its end. Nevertheless, as I move my arm across space, I am not making rapid time-space calculations to ensure that I end at the right place. Rather, I “know” the trajectory ahead of time—indeed, I know it as early as my preparatory lift before the phrase.⁽¹¹⁾

[15] Something analogous occurs in the mental realm as I prepare to play each of these three phrases at the keyboard: at each of those three preparatory moments, I protend the forthcoming phrase into my present consciousness.⁽¹²⁾ I also find that this in-the-moment image vividly captures the overall gestalt and energy curve of the phrase, while conveying only a dim sense of its details. (Later I will describe lower-level projections in which the details emerge with great clarity.) Again, the directionality of my projections is critical; my claim is not that I mentally project these spans into the future, but rather that my awareness of a future context protends into my present. Here then we already see a fairly rudimentary, though powerful, form of audiation.

[16] The video clip reveals traces not only of my protensions, but my retentions as well. For instance, my arm gesture on the third phrase conveys my sense of its *slowness*, as though I am shifting into a lower gear or moving my arm through a resistant substance of some kind. This sense of slowness of course depends on my retention of the relatively fast spans of mm. 1–2 and 3–4. But paradoxically, this span also feels more energetic as my gesture responds to the ascending contour in the outer voices, as well as the increasing chromaticism, as the passage pushes up to a cadence in the dominant. My desire to express this energy in my arm motion simultaneously increases my sense of resistance described earlier, producing a great deal of tension. And again, I already “know” this qualitative difference during the breath preceding the third gesture and feel it in my preparatory gesture.

[17] The video clip also conveys a lower-level distinction between the first two gestures. Observe that my arm gesture for mm. 3–4 is somewhat more expansive than for mm. 1–2, reflecting its higher melodic range and its inflection of its climax A with the chromatic passing tone G#. This sense of expansiveness of course arises from my retention of the comparably contained motion of mm. 1–2. This in turn enriches my subsequent protension of mm. 5–8 during the juncture between mm. 4 and 5. During that moment I retain not only the relatively brief durations of mm. 1–2 and 3–4, but also their different energy profiles; this in turn leads me to protend mm. 5–8 as the culmination of a three-stage intensification that spans the entire eight measures.

[18] As already noted, my projections during this passage depend on some kind of mental representation of the retained and protended context. Thus I do not need to rapidly auralize these past and future contexts within my musical present in order for me to feel their force. It lies beyond the scope of this article (as well as my level of self-awareness) to specify the precise nature of these

representations, but recent music-cognitive research suggests that they are likely a multimodal blend of visual, kinesthetic, and sound imagery (Keller 2012, 206). This would explain their vivid and gestalt-like reality for me. Moreover, while the preceding description of my internal processes may imply that these representations suddenly spring up only at junctural moments in the music (that is, between mm. 2 and 3, and again between mm. 4 and 5), that is surely not the case; at some level I hold these representations in consciousness throughout the entire passage.

[19] So far I have focused on the 2+2+4 gestural spans of the Beethoven passage. But the relative expansiveness of the third unit invites a rich internal network of protensions and retentions. At the juncture between mm. 4 and 5, I pretend a 2+1+1 subgrouping pattern for mm. 5–8 based on rhythmic patterning.⁽¹³⁾ This awareness vivifies my audiation in two ways. First, it ironically increases the power of the forthcoming four-measure gesture in my imagination, since I now feel a proactive urge to push through the potential grouping breaks between mm. 6 and 7, and again between mm. 7 and 8, in order to convey the full sweep of the phrase. Second, my protension of the 2+1+1 subgrouping draws my attention to the correlation between the first subgroup, mm. 5–6, and an earlier context that corresponds to it, mm. 1–2. Now I anticipate the melodic F $\sharp$ –G at the end of m. 6 as especially salient, since it contrasts so markedly with F–D motion in m. 2. This enhances for me the pull of the F $\sharp$ to G, not only because such pull exists in its immediate context of mm. 5–6, but because the F $\sharp$ *replaces* the F that I retain from the corresponding spot in m. 2.

[20] My hearing of F $\sharp$ in m. 6 raises another question. As noted, I hear this pitch as marked in part because of my memory of mm. 1–2, which contained F $\natural$ in the corresponding moment. But how do I cognize this difference between these passages? Again, some kind of mental representation must come into play; I do not auralize mm. 1–2 as I play mm. 5–6 and hear a clash in the second measure of the respective units. Lewin is helpful here, for he regards apparently conflicting perceptions as occurring in different “phenomenological locations,” citing the deceptive V-to-vi motion as a prototypical example. That is, while we hear the V chord, we pretend a resolution to I;⁽¹⁴⁾ the subsequent actual appearance of the vi chord then denies the arrival of I, but does not negate our earlier projection of a I, nor does it create a clashing sound image with that earlier projection. Our two perceptions, one of the projected I, and the other of the actual vi, exist in different phenomenological locations (Lewin, 331–34). Similarly, while I am playing the F $\sharp$ in m. 6, the F $\natural$ resides in a different phenomenological location in my musical present, directly bearing on my experience of the F $\sharp$. The F $\sharp$ appears more urgent, and certainly more marked, than it would have otherwise.⁽¹⁵⁾

[21] If I expand my audiatonal attention beyond the melody of mm. 5–6 to include the entire texture in those measures, harmony enriches the sound image. As noted, Lewin’s contexts can involve a purely harmonic dynamic of implication-realization, as in a V–vi progression. Consider, for example, the V $\frac{4}{3}$ of V that I am playing in support of the melodic F $\sharp$ in m. 6. This chord would qualify for Lewin as an event that leads me to pretend a resolution to a G-major chord in either root position or first inversion. Here I’m not referring to my aural projection of what I know is coming up, but rather my audiated protension of what “wants” to come up (regardless of what I actually know will happen). This might seem like an overly fussy distinction, since both operations anticipate the same chord; yet the difference is critical if we remember Gordon’s view of audiation. When I project a first-inversion G major chord simply because I know I am about to play it—i.e., without experiencing the implication-realization dynamic set up by the V $\frac{4}{3}$ of V—I am only predicting its arrival in the most limited sense.⁽¹⁶⁾ Even if one grants that such a prediction qualifies as a kind of audiation (it is not clear that Gordon would do so), it would at best count as a particularly weak manifestation thereof.

[22] One might protest that it is impossible for anyone with ears accustomed to Western tonal music not to hear some degree of implication in the V $\frac{4}{3}$ of V in m. 6. I would tend to concur, but audiation is not an all-or-nothing affair; a performer’s audiatonal powers can fall anywhere on a broad continuum with respect to the *vividness* of his projections. For instance, imagine a performer who pretends two possible resolutions of the V $\frac{4}{3}$ of V (i.e., to either V or V 6) with a vivid awareness of the likely voice-leading details in both versions, compared to a second performer who, although able to hear that the F $\sharp$ wants to resolve up to G, possesses only a vague aural image

of what the bass and harmonic support for that resolution would be. The possibility that a performer might possess different *degrees* of audiatonal skill in turn has enormous consequences for musical training, an issue to which I will return in the final portion of this essay.

[23] In any event, the $V \frac{4}{3}$ of V does resolve to V^6 on the second half note of m. 6 (one of the two options that we projected in purely audiatonal terms, regardless of what we “knew” was coming up). Continuing on into m. 7, I experience the second beat as a veritable hot spot for audiatonal retentions and protensions. If I momentarily pause there, I realize that I protend a forthcoming six-four chord—again, not only because it *will* follow, but because it *wants* to follow. I am protending three distinct sub-events here. First, in the melodic line I hear the Bb wanting to go up to B natural, since the resulting A–Bb–B motion in that measure would parallel the ascending E–F#–G motion of m. 6, which I retain in my memory (**Example 4**). Second, I protend an upward resolution for the bass C# on beat 2 due to my audiatonal competence in early 18th-century tonality. That is, I aurally recognize the first two chords of m. 7 as a harmonic formula, ii $\frac{6}{5}$ to vii^o $\frac{7}{5}$ /V in G, which leads me to hear its bass motion as $\hat{4}-\hat{\#4}$ (in G), which clearly protends a bass continuation to $\hat{5}$ in G (**Example 5**). Third, given my protension of the outer voices B and D (**Example 6a**), I expect a six-four chord within that frame (say, rather than a iii⁶) due to my awareness of Beethoven’s early style (**Example 6b**).⁽¹⁷⁾

[24] If I switch to a rhythmic/metrical audiatonal filter for m. 7, however, I experience an odd disturbance. On the second beat of m. 7, I protend a half-note duration on beat 3 based on my retention of m. 6, which similarly presented a rhythm of two quarters and a half.⁽¹⁸⁾ But I also audiate a stylistic norm here that gives rise to a conflicting expectation. In the Classical style, cadential six-four chords are almost always metrically stronger than their resolution. That leads me to protend a V or V⁷ chord on beat 4, which would make the six-four on beat 3 relatively strong. **Examples 7a and 7b** illustrate two possible continuations (with accompanying audio clips). But Beethoven delays the V7 until the following downbeat, which places the six-four chord in a *weaker* metrical position than its resolution. Overall, I hear the effect as somehow “wrong,” even comical; and the *sf* on the six-four just adds to this effect, as if the chord is trying to overcome its awkward metrical position through a willful assertion. And I respond this way because somewhere in my consciousness I am imagining a stylistically normative behavior for the six-four chord that projects into my present awareness.⁽¹⁹⁾

[25] So far I have attempted to convey my audiatonal experience while performing the Beethoven passage in real time. But real-time performance occupies only a small slice of a performer’s professional life, which includes countless hours of work in the practice room. What role does facilitative agency, and audiation in particular, assume in that setting? To the extent that a practice session involves the playing through of musical passages (of whatever length), the kinds of in-time projections described thus far indeed continue to operate. But in addition, a practice room setting grants the performer a freedom to explore alternative interpretations in a conscious and deliberate way. The following discussion describes one such possible exploratory path through the Beethoven excerpt, a path that ultimately undercuts the 2+2+4 grouping shown with my arm gestures in Example 3.

Facilitative agency in the practice room: audiatonal play

[26] Let us go back in time and imagine that I am learning the Beethoven movement for the first time. The first eight measures seem simple enough at first, yet as I play them over several times, they start to cause me some trouble. I can’t get these measures to move naturally; either they are sound too static, or I press ahead too much, making the forward motion sound forced. So I play around with just the opening two bars, attempting to get inside the music to sense what motivates it to move towards its goal in m. 2. I come upon two ways of hearing and playing these measures. First I consciously attend to its prolongational structure, noticing how the soprano and bass move within a prolonged tonic harmony for five chords before taking us to a ii6–V cadence. My overall impression here is of melodic motion within an over-riding harmonic stasis, followed by a sudden and decisive move to the cadence. My hearing is also shaped by an extra-opus reference: the opening tonic expansion strongly recalls the Fenaroli schema (Gjerdingen 2007). Hearing the first

five chords as a Fenaroli actually increases my impression of its harmonic stasis over that span, for the pattern invites its own canonic self-perpetuation (soprano $\hat{1}-\hat{7}-\hat{1}-\hat{2}-\hat{3}-\hat{7}-\hat{1}-\hat{2}-\hat{3}$, etc., and bass $\hat{1}-\hat{2}-\hat{3}-\hat{7}-\hat{1}-\hat{2}-\hat{3}-\hat{7}-\hat{1}$, etc).⁽²⁰⁾

[27] But a second hearing also vies for my attention, lured by the evenly paced rhythm of the chordal progression, as well as its crisp and incisive articulations. Now the contours of the outer voices seem especially salient (**Example 8**). I hear the soprano line moving insistently and deliberately towards its goal, each pitch contributing equally to its sense of forward motion. This interpretation of the soprano resists even a hint of linear segmentation, unlike my Schenkerian cognitive filter that led me to hear it as a series of overlapping prolongational units. Meanwhile, I hear a more complex shape in the bass: a three-note ascent, following by a descending sequence of two ascending seconds, B–C and F–G.

[28] Thus far, my practice-room scenario illustrates how a performer might experiment with different audiatonal frames during a practice session. We had already seen a hint of multiple audiatonal framing in my earlier discussion of m. 7, which distinguished tonal projections from rhythmic projections. But in that case such projections arose spontaneously in the flow of a real-time performance, whereas the practice room setting invites the performer to slow down and explore different hearings. In so doing, he engages in *audiatonal play*. The rest of this scenario suggests that such play can have wide-ranging ramifications, leading a performer to question interpretative assumptions that had before seemed stable and even indisputable.

[29] To return, then, to my practice-room scenario: after playing around with both audiatonal options for mm. 1–2, I decide that my focus on the outer-voice contours helps me to shape mm. 1–2 in a way that feels most compelling (though I retain a shadow awareness of the prolongational structure). When I continue on to mm. 3–4 with this contrapuntal aural filter, I find that this phrase creates an even stronger sense of direction than the first, with the bass now leading the way; it progresses downward by step from G3 on the upbeat towards the G2 in m. 4. Meanwhile, the upper voice picks up on G4, a step higher than the F4 climax of the first phrase, and pushes up chromatically to A4 for the climax of mm. 1–4 as a whole. Both the soprano and bass therefore intensify the impression of motion I already experienced in mm. 1–2. Measures 5–8 continue this intensification in the outer voices, with chromaticism now featured in both soprano and bass, and the melody pushing even higher than before. In retrospect, I experience mm. 3–4 as an intermediate level of intensification between mm. 1–2 and 5–8. (It will be recalled that my arm gestures in Example 3 demonstrated such a three-fold gestural expansion.)

[30] During my next practice session I basically adhere to this broad audiatonal framework for mm. 1–8 while exploring the more quirky aspects of the passage (especially mm. 7–8, discussed earlier). I also try to refine the voicings of the chords with special attention to the soprano/bass counterpoint. I decide to isolate the outer voices, and while playing through mm. 1–4, I am struck by a peculiarity in the bass line, namely the slur on the bass octave leap in m. 2. It's the first slur in the passage, and the sheer size of the leap itself seems completely unprepared by what precedes it. How should I handle this moment gesturally? In quickly auralizing the bass line for mm. 1–2.3, I notice that it gradually winds its way downward, and that the ascending motion is exclusively by step. It sounds somehow more cautious than before. Moreover, its staccato attacks, *piano* dynamic, and constricted range all lend the line a furtive, almost comical character. Earlier I had dimly noticed this rather quirky aspect of the bass opening, but now its character is more sharply etched for me. The stimulus for this new awareness is my protension of the forthcoming octave leap, which now looms as a remarkably bold gesture.

[31] In auralizing the leap itself, I am now struck by its power. The slur forces me to inject a great deal of energy into the low G so that I can lift upward and land gracefully on the high G without hammering it unmusically. As I then finish the phrase up to the third beat of m. 4, I now notice that I am slowly filling in the large gap that was created by that upward leap (which of course I am retaining in my audiatonal present).

[32] When I then go back to play the first four bars in their entirety, the effect of that octave leap ripples outward; what I thought had been a stable 2+2 grouping suddenly tilts. Before I had heard

the octave slur as a surface connection between a cadential G2 and the G3–F3 lead-in to the next phrase. Now it feels as though the left hand is getting a jump on the right hand by starting its phrase two beats early. (The G2 in m. 2 still sounds cadential, but now the pitch takes on a second function as the initiator of a phrase as well.)

[33] My audiatonal tilt affects not only my hearing of the phrase groupings in mm. 1–4, but also how I hear the relationship between the outer voices in mm. 3–4: the bass has taken over as the leading voice, and the soprano sounds subordinate to it. But that’s not all: the textural demotion of the upper line in mm. 3–4 facilitates a new hearing for me. I can now perceive a hidden statement of the melodic line of mm. 1–2 embedded in the right hand of mm. 3–4.⁽²¹⁾ This affects my hearing of the upper voice, which now sounds like a descant or cover line above that embedded melody. This in turn leads me to play around with possible voicings at the keyboard (**Examples 9a and 9b**). Example 9a brings out the top voice, while 9b subordinates that voice to the embedded melody.⁽²²⁾

[34] My new possible hearing of the top line in mm. 3–4 as subordinate in turn greatly weakens my impression that those measures represent an intermediate stage of intensification within mm 1–8 as a whole: now mm. 3–4 sound more like a varied repetition of mm. 1–2.⁽²³⁾ This in turn suggests the possibility of hearing mm. 1–4 as a presentation within either a sentence or hybrid structure for the eight measures as a whole.⁽²⁴⁾ Now the “rippling out” effect produced by the octave leap in m. 2 permeates the entire eight-measure passage.⁽²⁵⁾

[35] In sum, the practice room provides ample opportunity for exercising several kinds of audiatonal play. First, we can explore different audiatonal frames, as seen in our discussion of mm. 1–2.⁽²⁶⁾ Second, even within the context of a single operative frame we can try out different audiatonal paths through that frame, as just seen with respect to the grouping structure in mm. 1–4. Here such play is especially fruitful when the frame supports several distinct interpretive options, say in the case of competing Schenkerian readings (Dodson 2008), the appearance of a shadow-meter along with the primary meter (Rothstein 1995; Samarotto 1999), or conflicting grouping structures (as in the present analysis). But my practice-room scenario also introduces a third kind of play that seems to be ad hoc in nature, as when the performer suddenly notices an event in the music that seems peculiar or quirky—for instance, my encounter with the bass octave leap of m. 2 of the Beethoven. We have seen that such a noticing can set off a chain of *new* noticings, and perhaps even overturn what had seemed audiatonally stable to us.

[36] While the preceding discussion has emphasized the relatively large-scale (i.e., phrase-level) ramifications of my contrapuntal audiatonal frame for mm. 1–2, it should be noted that this frame also affects more local projections within the phrase. For instance, on the fourth quarter note of m. 1, a Schenkerian framework would pretend a termination of a tonic prolongational unit on the downbeat of m. 2. In contrast, focusing on the outer-voice counterpoint leads me to pretend a continuation of both lines in accordance with some linear principle, such as Steve Larson’s (2012) “forces.” (In the latter case, I do not pretend a termination of anything on the downbeat of m. 2, but rather an event within a larger linear trajectory.)

[37] Finally, note that my practice room scenario “reads” to a large extent like a musical analysis of the Beethoven passage. Thus the boundary between a practice session and the analytical act can be remarkably permeable. To put it another way: by adopting a playful audiatonal attitude towards the music it engages, analogously to a performer in the practice room, analysis itself becomes a form of play. Play, in fact, is an essential aspect of analysis, whether we are in the practice room, silently studying a score, or simply auralizing multiple “takes” of a musical passage while taking a walk.⁽²⁷⁾

[38] Let me close by revisiting the distinction I made at the outset between facilitative agency and primary agency. How do the two relate? First, let me propose a hypothetical answer: perhaps facilitative agency, and more specifically, audiation, does the mental “mapping out” for action, and then hands over the actual execution to a different faculty, which I’m calling primary agency. But this view is seriously flawed, whatever its common-sense appeal. The Cartesian view that audiation could be a purely mental process that somehow converts into a purely physical act of

performance is not sustainable philosophically; nor, for that matter, does it accord with my own experience or of any musicians I know.⁽²⁸⁾

[39] A second possibility is to regard facilitative agency not as the prime mover of the action, but rather as the performer's mental construction of a framework for meaningful action (that is, primary agency), somewhat analogous to the schemata that a jazz musician draws on during an improvisation. The performer acts as a semi-free (primary) agent within that framework, deciding in the moment such matters as rhythmic inflection, the bringing out or downplaying of potentially salient details, whether to telegraph a striking event ahead of time, and so on. Often the performer "decides" these issues unconsciously, and depending on her experience and audiatonal skill, even the framework itself may be an unconscious construction.⁽²⁹⁾

[40] Even this second possibility, however, runs afoul of the dualist fallacy. Granted, the mental construct no longer directs the action, but this explanation still differentiates the mental and bodily components of performance. In addition, as argued by several authors cited earlier, what I am calling primary agency is experienced by performers as a synthesis of cognitive activity and physical action. And coming at the duality problem from the opposite direction, cognitive scientists have challenged the notion that our internal imagery is merely "mental" in nature: we also access auditory and kinesthetic imagery.⁽³⁰⁾

[41] Yet a third possibility is to abandon the mind-body distinction altogether and to view facilitative and primary agencies as two sides of a single integrated process. According to this understanding, facilitative agency simply would be our shorthand term for the mental aspect of the process, and primary agency would be our designation for its physical aspect. This third view, which accords well with the testimony of many performers, suggests that the two agencies identified in this paper are better regarded as sub-agencies contributing to a single and integrated performative agency. In other words, we would now ascribe agency to the performance act *in toto*, and describe that agency as having two inseparable and intertwined aspects—the mental and physical.⁽³¹⁾

[42] But does not this integrative view of performer's agency call into question the very notion of a reified facilitative agency (or sub-agency, per the preceding), since the latter now appears to be subsumed within a higher-level agency in which the distinction between mental and physical processes is slippery at best? Two answers may be offered to this question, depending on one's reason for distinguishing these two agencies in the first place. If the objective is to explain how the performer exercises agency in a real-time performance, it may well be that the notion of a facilitative agency per se will prove to be illusory or otherwise problematic. For instance, future advances in neuroscientific research and/or in our philosophical understanding of the mind-body problem may render meaningless any distinction between facilitative and primary agency during a performance. Still, such a judgment would be premature at this point, and in any case, there is considerable pragmatic (and musical) value in treating facilitative agency *as though* it is an agency in its own right, as demonstrated in this paper.

[43] If, on the other hand, our objective for distinguishing these two kinds of agency is to elucidate the skills-based foundation for a real-time performance, the distinction is indisputably of value. Specifically, these two agencies roughly correspond to two widely accepted aspects of *musical training* for performers: the development of the physical technical skills needed for performance (correlating with primary agency), and the development of audiatonal skills (correlating with facilitative agency). Framing our understanding of the two agencies in this way expands the focus from real-time performance to include the training regimen that prepares the student for such performance. Moreover, insofar as such training entails purposive and intentional activity on the student's part, the role of agency—both primary and facilitative—is similarly expanded; that is, the performer exercises agency not only during real-time performance, but also at every stage on her developmental path as a performing musician.

[44] With respect to facilitative agency in particular, this path includes a traditional musicianship-training regimen, with its emphasis on ear training and sight-singing. But it ultimately must involve much more than that, assuming we take seriously Gordon's understanding of audiation as

the very “foundation of musicianship,” a capacity that informs all kinds of musical activity, not only ear training and singing. Gordon’s learning-sequence model, in fact, essentially proposes a unified-field theory for musical development that gradually *internalizes* and *integrates* multiple modalities of musical understanding.⁽³²⁾ Not all the particulars of his developmental model are transferable to the undergraduate level, but its broad objective of training a deep multimodal understanding of music certainly is.⁽³³⁾ My ultimate objective as a theory teacher, then, is to train my students to audiate as richly and as vividly as possible—not only in their singing and ear training, but through their work in keyboard, counterpoint, part writing, analysis—and ultimately, of course, in their performing.⁽³⁴⁾

Roger C. Graybill
New England Conservatory
290 Huntington Ave.
Boston, MA 02115
roger.graybill@necmusic.edu

Works Cited

- Caillois, Roger. [1958] 1961. *Man, Play, and Games*. Translated by Meyer Barash. The Free Press of Glencoe. Originally published as *Les jeux et les hommes*. Librairie Gallimard.
- Caplin, William E. 1998. *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. Oxford University Press.
- Cook, Nicholas. 2013. *Beyond the Score: Music as Performance*. Oxford University Press.
- Covington, Kate. 2005. “The Mind’s Ear: I Hear Music and No One is Performing.” *College Music Symposium* 45: 25–41.
- Doğantan-Dack, Mine. 2011. “In the Beginning was Gesture: Piano Touch and the Phenomenology of the Performing Body.” In *New Perspectives on Music and Gesture*, ed. Anthony Gritton and Elaine King, 243–266. Ashgate.
- Dodson, Alan. 2008. “Performance, Grouping and Schenkerian Alternative Readings in Some Passages from Beethoven’s ‘Lebewohl’ Sonata.” *Music Analysis* 27 (1): 107–34.
- Gjerdingen, Robert. 2007. *Music in the Galant Style*. Oxford University Press.
- Gordon, Edwin. 2008. “Audiation.” *GIML - The Gordon Institute for Music Learning*.
<https://giml.org/mlt/audiation/>
- . 2012. *Learning Sequences in Music: A Contemporary Music Learning Theory*. GIA.
- Gritten, Anthony, and Elaine King, eds. 2006. *Music and Gesture*. Ashgate.
- Graybill, Roger. 2014. “Thinking ‘in’ and ‘about’ Music: Implications for the Theory Curriculum.” In *Engaging Students: Essays in Music Pedagogy*. Vol. 2.
<http://flipcamp.org/engagingstudents2/essays/graybill.html>
- . 2018. “Activating Aural Imagery through Keyboard Harmony.” In *Norton Guide to Teaching Music Theory*, ed. Jeffrey Swinkin and Rachel Lumsden, 182–97. W. W. Norton.
- Hasty, Christopher. 1997. *Meter as Rhythm*. Oxford University Press.
- Hatten, Robert S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Indiana University Press.
- . 2006. “A Theory of Musical Gesture and its Application to Beethoven and Schubert.” In *Music and Gesture*, ed. Anthony Gritten and Elaine King, 1–23. Ashgate.

- Held, Klaus. 2010. "Phenomenology of 'Authentic Time' in Husserl and Heidegger." In *On Time – New Contributions to the Husserlian Phenomenology of Time*, ed. Dieter Lohmar and Ichiro Yamaguchi, 91–114. Springer.
- Huron, David. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press.
- Kane, Brian. 2011. "Excavating Lewin's 'Phenomenology.'" *Music Theory Spectrum* 33 (1): 27–36.
- Karpinski, Gary S. 2000. *Aural Skills Acquisition: The Development of Listening, Reading, and Performing Skills in College-Level Musicians*. Oxford University Press.
- Keller, Peter E. 2012. "Mental Imagery in Music Performance: Underlying Mechanisms and Potential Benefits." *Annals of the New York Academy of Sciences* 1252 (1): 206–13.
- Larson, Steve. 2012. *Musical Forces: Motion, Metaphor, and Meaning in Music*. Indiana University Press.
- Lewin, David. 1986. "Music Theory, Phenomenology, and Modes of Perception." *Music Perception* 3 (4): 327–92.
- Mead, Andrew. 1999. "Bodily Hearing: Physiological Metaphors and Musical Understanding." *Journal of Music Theory* 43 (1): 1–19.
- O'Hara, William. 2012. "Music Analysis as Play." Paper presented at the annual meeting of The New England Conference of Music Theorists, New London, CT.
- Pierce, Alexandra. 2007. *Deepening Musical Performance Through Movement: The Theory and Practice of Embodied Interpretation*. Indiana University Press.
- "protend, v." OED Online. June 2015. Oxford University Press. <http://www.oed.com/>
- Rothstein, William. 1995. "Analysis and the Act of Performance." In *The Practice of Performance: Studies in Musical Interpretation*, ed. John Rink, 217–40. Cambridge University Press.
- Samarotto, Frank. 1999. "Strange Dimensions: Regularity and Irregularity in Deep Levels of Rhythmic Reductions." In *Schenker Studies* 2, ed. Carl Schachter and Heidi Siegel, 222–38. Cambridge University Press.
- Schmalfeldt, Janet. 2011. *In the Process of Becoming: Analytic and Philosophical Perspectives on Form in Early Nineteenth-Century Music*. Oxford University Press.
- Sheehy, August. 2013. "Improvisation, Analysis, and Listening Otherwise." *Music Theory Online* 19 (2).
- Taggart, Bruce. 2005. "Music Learning Theory in the College Music Theory Curriculum." In *The Development and Practical Application of Music Learning Theory*, ed. Maria Runfola and Cynthia Crump Taggart, 345–58. GIA.
- Tzotzkova, Victoria. 2012. "Theorizing Pianistic Experience: Tradition, Instrument, Performer." PhD. diss., Columbia University.
- Urista, Diane J. 2016. *The Moving Body in the Aural Skills Classroom: A Eurhythmics Based Approach*. Oxford University Press.

Footnotes

1. [Gritten and King 2006](#) contains several exemplary essays of this sort.
[Return to text](#)
2. Gordon has been a major influence on American K–12 education since the 1970s, when he published the first edition of *Learning Sequences in Music* (the most recent edition of which came out in 2012).
[Return to text](#)

3. While Lewin does not explore the possibility, a given event will often imply a *range* of realizations, each of which we unconsciously rank according to degree of probability. (For instance, see [Huron 2006](#) (158–62) regarding the relative likelihood of different continuations from a given scale degree.) One might imagine the result as a collection of multiple superimposed protensions, some of which strike us as especially vivid.

[Return to text](#)

4. Throughout this essay, I will be using the verb forms “retain” and “protend” to correspond to the nouns “retension” and “protension,” respectively. My choice of “protend” requires some elaboration. While Lewin himself uses the word, its definition in the Oxford English Dictionary differs from his apparent meaning: “In phenomenology: to extend (the consciousness or perception of a present act or event) into the future” ([OED 2015](#)). This seems to contradict Lewin’s notion of protension as a projection of anticipated events into present consciousness. Held ([2010](#)) acknowledges the difficulty in terminology, and even suggests alternatives: “If . . . we emphasize that a ‘protention’ anticipates contents of intentional fulfillment, i.e., that consciousness possesses these contents in advance, ‘holding’ them in it, the corresponding verb would be ‘to protain’ or ‘to protenuate’ (from Latin *tenere*, ‘to hold’)” (112, fn 19). Despite the appeal of these alternatives, I have decided to keep “protend,” while assuming Lewin’s interpretation of the term. (On “protention” versus “protension,” see [Eugene Montague’s footnote](#) in this issue.)

[Return to text](#)

5. This example reproduces a portion of Lewin’s Figure 7 on page 345 of his article.

[Return to text](#)

6. As [Kane 2011](#) notes, Lewin’s belief in the inadequacy of his p-model in accounting for creative acts of musical production (i.e, composing and performing) leads him to shift towards a post-Husserlian (“embodied”) phenomenological perspective in Part V of his essay.

[Return to text](#)

7. Lewin does grant that a performer “can enter into noetic-noematic exchanges, even subject/object relationships, with parts of the acoustic signal already produced” (376–77). But note the absence of any reference to future contexts, which supports the likelihood that Lewin does not regard the performer’s intentional projections of future contexts as objects of perception.

[Return to text](#)

8. To be sure, the distinction between intentional and spontaneous projections is not always clear. It is probably best to think of these two categories as idealized endpoints on a continuum.

[Return to text](#)

9. Alexandra Pierce’s notion of “reverberation” conveys this intentional attitude more precisely: “Reverberation reflects the intention that precedes an action, commitment during the action (including a desire to communicate with an audience), and fulfillment of the action” ([2007](#), 121). While the definition refers to an audience, a performer can embody the same attitude during a play-through in the practice room, *as though* an audience is present (or alternatively, with the performer herself assuming the role of “audience”).

[Return to text](#)

10. Here I use the term “phrase” informally to denote a grouping unit that comes to some kind of an arrival.

[Return to text](#)

11. The use of arm gestures to experience (and show) phrase shapes will be familiar to readers with experience in Dalcroze eurhythmics. ([Urista 2016](#) describes other ways of physically expressing phrasing from a Dalcrozian perspective.) This particular exercise bears considerable resemblance to Alexandra Pierce’s “arcing” ([2007](#), 108–12), though the latter entails extending the arm upward along a vertical plane, while my arm gestures move horizontally.

The claim that I can “know” the physical trajectory ahead of time assumes both stylistic competence on my part—that is, familiarity with classical phrasing—and experience with this

particular Dalcrozian exercise as a means for expressing such competence.

[Return to text](#)

12. My claim that I “protend” a forthcoming passage of music must be understood as a convenient shorthand for a more complex perceptual process as described by Lewin’s model. This shorthand formulation operates on two levels. First, it adopts the word “protend” in the looser intentional sense that Lewin’s model would probably not countenance, as explained earlier. Second, even if we adopt here the role of a naive listener—say, a listener who expects m. 5 to repeat the content of m. 1 (given her stylistic awareness of the parallel phrase structure)—that listener is not directly protending the forthcoming music itself. Rather, she is protending a *perception*, the object of which is the expected content of m. 5. Lewin’s perceptual model requires such descriptive rigor, but as noted earlier, this study is more interested in broader aspects of that model—especially the precision with which it defines contexts and its claims about how those multiple contexts impinge on the present moment without producing auditory chaos.

[Return to text](#)

13. It is also possible to protend a 2 + 2 grouping, with the last two measures bound together by means of a cadential progression within a tonicized G major. I wish to thank one of my anonymous reviewers for this insight.

[Return to text](#)

14. Again, this is a shorthand formulation (cf. footnote 12). Lewin is careful to explain that as we hear the V, we protend a *perception* of a forthcoming I chord. He contrasts this claim with the normal explanation that we “expect” a forthcoming I that “ ‘has not yet happened.’ ” In Lewin’s model, our perception of an anticipated tonic chord “*does actually happen*” while we are hearing the dominant (332; italicized in original).

[Return to text](#)

15. How I might choose to express this quality of the F[♯] (or not) in my actual performance is a different matter; here again we see the distinction between facilitative and primary agency.

[Return to text](#)

16. To put it another way, my intention with respect to the V[♯] in m. 6—that is, to resolve it to the V⁶ that I know is coming next—is not the same as my audial understanding of that V[♯].

[Return to text](#)

17. Note that the protensions I am describing here are essentially equivalent to those of the listener in Lewin’s p-model; they are more or less involuntary projections, independent of my will or “intention” as a performer.

[Return to text](#)

18. Such projection of a durational span on the basis of a *preceding* durational span calls to mind the projectional analyses in [Hasty 1997](#).

[Return to text](#)

19. To the extent that my audiation invokes extra-opus stylistic norms, I am engaging in neither retention or protension, strictly speaking. Rather, I am projecting a style-specific harmonic/metric schema into my present consciousness. Recall that Gordon’s understanding of audiation similarly draws on stylistic norms.

[Return to text](#)

20. Indeed, Gjerdingen (2007, 42) notes that the Fenaroli is normally repeated, even though Beethoven does not do so here.

[Return to text](#)

21. Janet Schmalfeldt suggested this possible hearing during the question-answer period following my presentation of this topic at the EuroMAC 2015 conference.

[Return to text](#)

22. To clarify, while my “playing around” in this way with mm. 3–4 was spurred by Janet Schmalfeldt’s observation (see fn 21), she herself did not address a link (or non-link) between that observation and a possible realization in performance.

[Return to text](#)

23. I find special pleasure in hearing the melody “jump” voices from D (a chordal seventh) to the E —another example of an audiatonal tilt.

[Return to text](#)

24. To be sure, both interpretations—sentence and hybrid—are problematic in some respect. We do hear a hint of a continuation function in mm. 5–8 due to the accelerating 2-1-1 measure grouping. But the acceleration is uncharacteristically late, starting only in the seventh measure. Moreover, m. 5 sounds like a “starting over” after the strong half cadence in m. 4, suggesting the initiation of a consequent function rather than a continuation.

Hearing mm 5–8 as a consequent in turn supports a “presentation + consequent” hearing for all eight measures. But again, this interpretation is not unequivocal, since the strong half cadence in m. 4 is more suggestive of an antecedent function than a presentation. Caplin (1998) also notes that the “presentation + consequent” is very rare; indeed, he does not even include this possibility as one of his four standard hybrids (63).

[Return to text](#)

25. The influence of that octave leap in fact bears on a formal juncture within the theme as a whole (m. 1–20). Observe that this same leap returns on the half cadence at the end of the B section (m. 12, beat 3), which calls into question where exactly A’ begins: the last half of m. 12, or the downbeat of m. 13? Further clouding our sense of return here is that the stepwise bass descent following the leap (recalling mm. 2–3) produces a relatively unstable I⁶ on the downbeat of m. 13, precisely where we might have expected a root-position I (since the soprano melody from m. 1 returns there).

[Return to text](#)

26. For another example, see Rothstein 1995, which investigates the tension between grouping and harmonic syntax in several works.

[Return to text](#)

27. O’Hara 2012 identifies four species of play (after Caillois [1958] 1961), each of which can be manifested through musical analysis: *agon*, *alea*, *mimesis*, and *ilynx*. Of these, *alea* corresponds to the kind of audiatonal play I am describing here. As O’Hara notes, the word *alea* connotes the idea of chance, and therefore an analytical attitude of openness to surprise: “We find it in action every time a vivid detail leaps out only after our tenth hearing, or when an intertextual resonance involuntarily emerges in response to something else that we’ve recently heard or read. . . . playful analysis searches for new ways of listening and hearing, through private, preliminary performances.” Here “preliminary” alludes to the way that such analytical explorations may culminate in a public communication of one’s analytical findings, for instance through a publication or conference talk. For more on analysis as play, also see Sheehy 2013.

[Return to text](#)

28. Nicholas Cook regards this kind of dualistic construction as a manifestation of a pervasive “page-to-stage” approach to music-theoretic analyses of performance, in which the notated score assumes higher ontological status than a performance of that score. He argues that this approach “transforms such dualism into a means of disciplining the performing body, subjecting it to a mentalist construal of the musical work” (2013, 41).

[Return to text](#)

29. Keller 2012 confirms that musical imagery “may be generated through either deliberate thought or automatic responses to endogenous and exogenous cues” (206).

[Return to text](#)