An Eighteenth-Century Theory of Musical Cognition?

John Holden's Essay towards a Rational System of Music *(1770)*

Carmel Raz

Abstract Among the most exciting aspects of John Holden's *Essay towards a Rational System of Music* (1770) is its explicit ambition to explain musical practice by means of a limited set of psychological first principles. Relying primarily on introspection, it nonetheless describes phenomena that we today understand as grouping, chunking, and subjective rhythmicization. In the absence of anything resembling a modern theory of cognition, Holden's account of how we can perceive music chiefly relies on the actions of posited mental faculties, including attention, memory, imagination, and expectation. These concepts allow him to develop detailed speculations about a range of conscious and unconscious dispositions of perception. This study explicates the *Essay*'s speculative theories and contextualizes them both within eighteenth-century music theory and in light of contemporary psychology.

Keywords John Holden, history of music theory, musical cognition, attention, perception

IN 1863 the Englishman Charles Isaac Stevens (1835–1917) submitted a doctoral dissertation in absentia to the Georg-August-Universität Göttingen. Titled "An Essay on the Theory of Music," the work derived a speculative theory of music perception from the hypothesis that the mind has an innate tendency to group all stimuli into small units of equal size—an essentially cognitivist approach foreshadowing experimental work on rhythm percep-

I thank David E. Cohen for his extraordinary guidance and mentorship throughout the writing of this article. Many ideas in this piece first arose during our conversations, and I am profoundly grateful for his rigor and intellectual generosity. Many thanks are also due to the Columbia Society of Fellows, as well as to Jamie C. Kassler, Patrick McCreless, Elaine Sisman, Ellie Hisama, Christopher L. Brown, Giuseppe Gerbino, and Nori Jacoby for their incisive comments on earlier drafts, and to Nathan Martin and Eric Bianchi for their advice and support. Particular thanks go to Richard Cohn, who encouraged my interest in Holden from the very beginning and provided valuable feedback, and to Roger M. Grant and an anonymous reader for this journal for their very helpful suggestions. Finally, I would like to dedicate this article to Jamie C. Kassler, whose work continues to inspire me.

DOI 10.1215/00222909-7127670

tion from the late nineteenth century, as well as key findings in Gestalt psychology. However, this dissertation was not written by Stevens. The text is lifted wholesale and without attribution from a work published in Glasgow nearly a century earlier: John Holden's *Essay towards a Rational System of Music* (1770). As far as I can tell, Stevens's plagiarism was never detected. He lived out the rest of his life in London, eventually assuming leadership positions in the Free Protestant Episcopal Church of England. Living well before the era of searchable text, Stevens had little reason to fear that his deception would ever come to light.

For Stevens's examiners, the dissertation's exploration of music perception may have seemed attuned to contemporaneous interest in the physiology and psychology of audition, evident in works such as Hermann von Helmholtz's *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* (1863) or indeed Hugo Riemann's dissertation "Über das musikalische Hören," which was submitted to the same institution in 1874.[1] Even if we grant that these readers might have simply rubber-stamped the English-language dissertation, the very fact that Stevens considered a century-old treatise worth pirating attests to the novel status of its ideas within European musical discourse. Holden's comprehensive exploration of cognitive preferences, manifested in a system of harmony as well as various aspects of rhythm and pitch perception, is unlike any other eighteenth-century theory of music. Stevens's act of plagiarism discloses the extent to which the *Essay* anticipated subsequent developments in philosophical, psychological, and musical thought.

Still, my subject here is neither Stevens nor the state of nineteenth-century music theory. Rather, this article focuses on Holden's extraordinary treatise by explicating its speculative theories and contextualizing the work within eighteenth-century music theory and in light of contemporary psychology. Perhaps the most exciting aspect of the *Essay* lies in its explicit ambition to explain musical practice by means of a limited set of psychological first principles. Relying primarily on introspection, it nonetheless describes phenomena that we today understand as grouping, chunking, and subjective rhythmicization. The *Essay* also gestures toward discrete forms of memory and attention, as well as the psychophysical phenomenon of just noticeable difference. Although the appeal to the mind and senses has some precedent, the scale and originality of Holden's theory are exceptional within the Enlightenment context and beyond.

The remarkable nature of Holden's ideas had a considerable impact in the late eighteenth century. Within the first forty years of its publication, the treatise enjoyed two additional editions, one printed in Calcutta in 1799

1 Riemann's dissertation was published as *Musikalische Logik: Hauptzüge der physiologischen und psychologischen Begründung unseres Musiksystems* (1874). The philosopher Hermann Lotze evaluated the dissertations of both Stevens and Riemann, as well as that of his own advisee Carl Stumpf, although Stumpf's dissertation does not deal with music (Woodward 2015, 468–72).

and the other in Edinburgh in 1807. The work was commended by Johann Nikolaus Forkel (1792, 418a) as "one of the best of its kind" and praised by Johann Georg Sulzer (1793, 457) a year later as "understandable and concise."[2] François-Joseph Fétis (1839, 188) deemed it "the best musical treatise published in England for over a century . . . characterized by a philosophical spirit that even today, renders it worth of attention."

In spite of its importance in the eighteenth century, Holden's music theory has attracted limited scholarly consideration.[3] We can hazard a number of reasons for this neglect, including the decline of the Scottish Enlightenment, the relatively limited interest in speculative music theory in Britain during the nineteenth century, and Holden's death within a few months of publishing his work.[4] Another reason lies in the fact that the two-part treatise, which adheres to the traditional theory/practice divide, begins with the latter ("The Rudiments of Practical Music") rather than the former ("The Theory of Music"), so the speculative portion is not readily apparent to the casual reader.[5]

Additional grounds for the relative neglect of the *Essay* surely lie in the work's unexpectedly challenging technical demands. The treatise contains few musical examples and does not engage with composers or their works.[6] Its eighteenth-century English is deceptively simple, masking a conceptual rigor that is easily overlooked. Furthermore, many of Holden's innovations are described in words, rather than formulas, and require an awareness of

2 All translations are mine unless otherwise indicated. Sulzer's description of Holden's treatise as "verständlich und bündig" first appears in this (second) edition.

3 Louis Chenette (1967, 346–66) provides a brief overview of Holden's conception of cadences, harmony, rhythm, and tuning. Jamie C. Kassler (1971, 83–87) discusses Holden's theory in relationship to Scottish musical thought. In her magisterial book, *The Science of Music in Britain, 1714–1830*, she elaborates on Holden's ideas in relation to French music theory and Scottish common sense philosophy while providing new details about Holden's life, influences, and the first edition of the *Essay* (Kassler 1979, 524–30). George Houle (1987, 78–79) mentions Holden's theory of pairwise accentuation and cites his example of accents imposed on watch ticks. Leslie Brown (2001, 122–32) examined the relationship between the first part of Holden's treatise and common sense philosophy, and Justin London (2004, 146) described Holden's account of metrical accents as "remarkably prescient of [Bruno] Repp's work." More recently, David Damschroder (2008, 149–50, 166–68) explicated aspects of Holden's harmonic system.

4 Much of what we know today about Holden's life comes from a 1772 letter by the philosopher Thomas Reid (2002, 74), who describes him as follows: "A Teacher of Musick, Writing, & Mathematicks in Glasgow. [He] was really an ingenious literary Man as well as of an excellent Character. He wrote a Treatise on the theory of Musick which does him much honour in the Opinion of good judges. He commonly had some Students of better Rank at the University who boarded in his house. He wrote our Records and Diploma's [*sic*], and Directed the Church Musick in the College Chappel, so that he was much connected with the College & much respected by the Masters. He had been some time a teacher as I have heard in an English Academy & both he and his wife were English People." This letter is also reproduced in Brown 2001, 131–32. I will discuss new archival findings regarding Holden's biographical circumstances, as well as his influence on other Scottish music theorists, in a subsequent article.

5 Stevens plagiarized only the first three chapters of the second part of Holden's *Essay* and removed all references and citations that could have dated the manuscript.

6 The only exceptions are as follows: in the first part of the treatise Holden refers to the opening of an aria by Thomas Arne as exemplifying the programmatic use of a deceptive cadence (part 1, §255, 292). He also discusses some Scottish psalm settings in the first part of the treatise (part 1, §35–36, 34–35) and comments on the use of accidentals rather than key signatures in editions by Corelli and Rameau (part 1, §262, 239). In the second part of the treatise he notes that the Scottish composer Thomas Erskine, Earl of Kelly, employed quintuplets (part 2, §13, 290).

eighteenth-century mathematical convention. A final reason may lie in the very nature of the deeper project that Holden sets himself, one that his contemporaries may have found difficult even to recognize.

Holden's investigations focus on the operations of the mind to elucidate the nature of our innate mental tendencies and their role in the phenomenological experience of hearing music. This orientation is reflected in his methodology, which prioritizes human behavior over the laws of counterpoint or physics. He criticizes the approach of mathematicians who "have bestowed their labour on wrong materials, for want of sufficient practice in music," as well as musicians who, "though better qualified in this respect, have yet adopted erroneous principles, without submitting them to accurate trial" (Holden 1807, part 2, §1, 284). Rather, he declares, his treatise examines "other properties which have hitherto passed almost unobserved, and which seem to constitute a very essential part of the theory of music" (part 2, §2, 284). This aim may explain the reversed order of the first and second parts of the treatise: rather than presenting abstract theories to be subsequently applied to the acquisition of practical skill, the *Essay* first expounds the rudiments of musical practice, often with experiential description, and then distills them into psychological principles. It therefore seems clear that Holden understood his focus on the perception of music as diverging from both the speculative and the practical traditions of his day.

In the absence of anything resembling a modern theory of cognition, Holden's explanation of how we can perceive music chiefly relies on the actions of posited mental faculties including attention, memory, imagination, and expectation. Although he uses these terms in a way that diverges at times from their modern colloquial meaning, these concepts allow him to develop detailed speculations about a range of conscious and unconscious dispositions of perception. The key difference between Holden's theory and any other eighteenth-century music-theoretical treatise is his comprehensive grounding of musical experience in a hierarchy of explicitly mental actions that are largely independent of any affective or mimetic response. His unambiguous focus on the mind and its psychological first principles can therefore with justice be understood as a precocious attempt at what we today would recognize as a cognitive theory of our mental processing of music. Seen in this light, Holden's approach is best characterized as synthesizing the rationalist tradition of music-theoretical engagement with cognition that is frequently ascribed to Jean-Philippe Rameau with Scottish philosophy of perception, in a strikingly bold and even visionary effort to explain precisely how the human mind perceives, grasps, and enjoys music.[7]

7 For more on Rameau's engagement with cognition, see Cohen 2001 and Moreno 2004. The relationship between Holden's writings and those of the Scottish philosophers Alexander Gerard, Thomas Reid, and Adam Smith will be the subject of a follow-up article.

The present study is therefore motivated both by the remarkable originality of Holden's thinking and by the desire to understand how and why his theories appear so modern to us today. I approach his treatment of scale and key from the dual perspective of the theoretical framework introduced in the second part of the *Essay* and the phenomenological observations in the first. Through this lens I explore how Holden derives both the foundational elements and the psychological experience of the tonal system from the claim that an innate preference for isochronous grouping determines our perception of music. I conclude by briefly positioning Holden's thought in relation to contemporary psychology.

Holden's speculative theory of music

1. Isochronous parcels and the analogy to rhythm

In line with the ideals of eighteenth-century rationalism, Holden begins the theoretical portion of his *Essay* with a new principle that he regards as primary and foundational. He asserts that our auditory perception is structured by the innate strategy of grouping sounds into small equal units, which he terms *isochronous parcels*.[8] Our preference for certain rhythms and pitch relationships within music is a consequence of a single far-reaching imperative: "The new principle we have here to propose, as being that whereby the various choices of a musical ear are best accounted for, is that of our distributing the vibrations of musical sounds by *isochronous,* or *equal timed parcels,* something very similar to the distributions we find naturally to be made among quavers, or other short notes, in the timing of music" (part 2, §10, 288).

According to Holden, it is the role of the ear, as distinguished from the mind, to determine the isochronism of successive parcels and to conclude whether the sound has a recognizable pitch and is hence suitable for music. The distinction between ear and mind seems at least partly intended as a way of evading the difficulty of claiming that we are able actually to count, and thereby determine the frequency of, sound vibrations.[9] This seems to be the problem he is trying to address when he adds, "It is one thing for the sense to be pleased with a number, and another thing to count it" (part 2, §10, 288).

8 Holden's very attempt at deducing a theory of harmony from psychological first principles explicitly responds to Rameau's bid to derive much of his theory quasi-deductively from the empirical fact of the *corps sonore*. In *Génération harmonique*, Holden charges, Rameau tried to ascertain "the choice of the ear [by] examining the properties of those sounds which we receive into our music" (Holden 1807, part 2, §9, 288). He condemns this argument as circular, demonstrating only that "the ear chuses such sounds because it chuses them" (part 2, §9, 288). The new first principles he proposes thus avoid Rameau's bias toward musical pitches by relying on the cross-modal phenomenon of grouping.

9 As David E. Cohen (2001, 68–74) has shown, Rameau's theory assumes the "musical ear" extracts the consequences of the *corps sonore*—including the diatonic succession, the fundamental bass, and the perfect (authentic) cadence—from the music itself. However, Rameau does not attempt to provide an account of the mechanism of unconscious processing involved in music perception.

While individual vibrations may be too fast for us to consciously grasp, the ear is simply charged with assessing incoming sounds for the equality of consecutive vibrations. Thus Holden compares the experience of sensing a given pitch to that of hearing a drum roll, where "the quick succession renders it impossible to count the single pulses; all we can here do is to judge of, and I may say—feel—the isochronism of each" (part 2, §15, 293). This process transpires by means of an unconscious aesthetic response: "We may perceive the equality, or isochronism, not only of certain compound parcels, but also of each single pulse, when a drummer beats much quicker than can be counted; and an inequality among the single pulses, or false and irregular time among the larger compound parcels would disgust us as much in this case, nay more than if the succession of pulses were slower" (part 2, §15, 292–93). That is, Holden claims that our senses evaluate isochrony at different time scales and that this determination has an affective component.

Whereas the ear gauges the regularity of sounds, Holden entrusts the mind with grouping isochronous vibrations of sound into equal units. He completes his new principle by adding that "among the isochronous single vibrations of musical sounds the mind naturally seeks to constitute isochronous compound parcels" (part 2, §15, §292). These isochronous compound parcels are cumulative or nested groupings (groups of groups), and Holden argues that our musical preferences result from this cognitive strategy of equal grouping and that this approach is shared across a range of perceptual domains: "Where equal and equidistant objects affect our senses . . . there is a certain propensity in our mind to be subdividing the larger numbers into smaller equal parcels; or as it may be justly called, compounding the larger numbers of several small factors, and conceiving the whole by means of its parts" (part 2, §11, 288–89).

Our tendency to understand large numbers as compounded of smaller factors can therefore be regarded as an innate cognitive strategy: on sensing that a number of similar objects are equal in size, our minds immediately group them into sets comprising smaller equal portions. The sensing of the ear is thereby implicitly distinguished from the operations of the mind, which is charged with conceiving the whole by means of synthesizing its parts. Holden clarifies this point by appealing to a familiar visual experience:

> When we cast our eyes on nine equidistant windows in a row, they are no sooner seen than subdivided into three times three: eight appears at first to be two fours, and each of these fours, two twos; seven we conceive as two threes disjointed, and one in the middle; six most naturally divides itself into two threes; but if seen along with nine, or immediately after it, we then trisect it, in conformity with nine, and it appears three twos: five becomes two twos disjoined and one in the middle; four becomes two twos, and single three or two need no subdivision. (part 2, §11, 289)

An illustration of this passage is shown in Figure 1. Holden asserts that our minds instantly arrange larger sets of similar objects into smaller equal por-

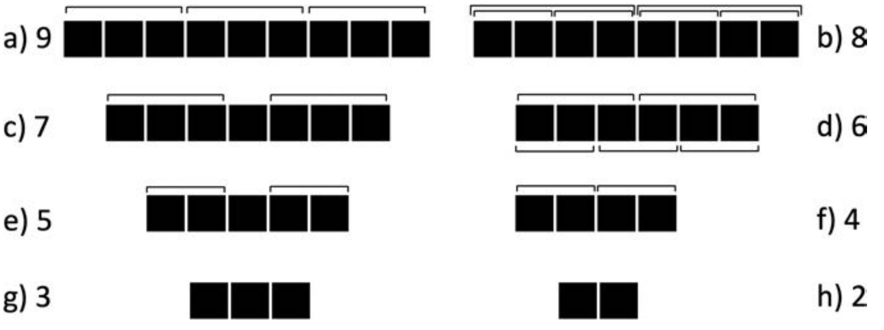


Figure 1. Illustration of Holden's theory of visual grouping.

tions factorized by the small primes 2 and 3. This grouping is internally hierarchical, with two groups of four subdividing into two groups of two (Figure 1b). He further asserts that the primes 5 and 7 divide into equal groups plus one remainder. Finally, in ambiguous cases, such as the grouping of six (Figure 1d), Holden maintains that groups are determined not only by the current state but also by the context of the grouping of the previous state. These claims prove essential to his theory of harmony.

Describing visual experiences, Holden observes that we obtain gratification from encountering "a certain symmetry which strikes us immediately with delight, in the prospect of a regular and well-designed piece of architecture" (part 2, §15, 293).[10] Regular proportions between different sounds in music, like proportions between different parts of such a building, "are the very things whence our pleasure is derived" (part 2, §15, 293). However, whereas in vision we can access an entire quantity at once, a different approach is required for sound, which occurs in time. Having now arrived at his real topic, he first takes up rhythm, where "we have to do not only with the choice of the mind, but also with the memory" (part 2, §12, 289).[11] Holden claims: "The mind insists that all our notes be made up, as it were, into isochronous parcels, which we call bars, or measures, and that the number of equal short notes which constitutes each measure, be a number some way compounded of the small factors two and three multiplied together, and rarely admits any larger factor than these" (part 2, §12, 289).

For us to successfully perceive music, therefore, our mind constantly groups incoming sounds into smaller equal units of twos and threes and further organizes these into groups of groups, that is, hierarchical structures.

10 This statement is reminiscent of contemporary eighteenth-century analogies between music and architecture. See Briseux 1753.

11 Holden uses the term *rhythm* in a broad sense to denote what we today more narrowly define as both meter and rhythm. Notably, the word *meter* does not appear a single time in his treatise, although he uses the term *mood* to indicate the various types of duple or triple time signatures.

This mental activity, in which groups of twos and threes are the principal factors according to which we comprehend larger wholes, is for Holden the principle of rhythm. Duple and triple divisions govern the mood, or metric character of a measure, as well as the surface rhythm, which subdivides further into twos or threes at the beat level. Holden notes that subdivisions of five and seven are also occasionally encountered at the beat level but emphasizes a difference between them: whereas the effort it takes to perceive groups of five can be attributed to lack of familiarity, groups of seven stand at the very limits of our ear's capacity and challenge our mental faculties (part 2, §14, 290).

Given an undifferentiated series of beats, Holden asserts, the listener will "naturally parcel them by 4 and 4 together, by giving a greater regard to every fourth beat; by which regard these beats would to him acquire a kind of emphasis or accent" (part 2, §14, 294). This accent, which arises from involuntary acts of attention, is explicitly imaginary; that is, it exists in our minds, rather than in the acoustic reality, and "would be the very same with a hearer if an inanimate machine were made to beat the drum" (part 2, §14, 294). This subject is also treated in the first part of the *Essay*, where Holden suggests that the aspiring student acquire a sense of rhythm by counting off duple and triple groupings of watch clicks. In this case, "we imagine the pulses which we count, to be really stronger than the intermediate ones, which we pass over. The superior regard which we bestow on the counted pulses is, here, the sole cause of these imaginary accents" (part 1, §95, 83). Predating Johann Philipp Kirnberger's account of accentual grouping in *Die Kunst des reinen Satzes* (1776), Holden's description of our involuntary mental impulse to metricize duple or triple groups of beats as arising from the allocation of attention may be the earliest articulation of the phenomenon known in modern psychological literature as subjective rhythmicization.[12]

Holden observes that changes in the domain of rhythm can at times require our minds to shift between comparable hierarchical levels of isochronous compound parcels. Indeed, he notes, any piece that includes triplets will feature repeated changes between triple and duple subdivisions. In such cases, he argues, we generally prefer to retain a shared beat level (we can think of this as a common denominator), which allows us to move smoothly between different groupings. Therefore we easily "pass from the division of a strain into 64 or 128 equal parts [i.e., consistent duple division: $64 = 2^6$ or $128 = 2^7$], to another kind of division into 48 or 96 parts [i.e., duple mixed with triple division, e.g., $48 = 2^4 \times 3$ or $96 = 2^5 \times 3$], and resume again the original division at pleasure, and from thence derive a very agreeable variety" (part 2, §17, 295–96).

12 For more on Kirnberger's theory of rhythm, see Grant 2014, 93–125.

Next, scaling up from beats and measures, Holden proposes that division and grouping by twos and threes also determine the lengths of musical phrases, subject to the constraints of our memory:

> The mind extends its view, and as far as the memory can be supposed distinctly to retain, goes on to constitute some number of measures into isochronous phrases, or strains of a tune; and these strains may contain a greater number of measures in quick time than in slow, because of the inability of the memory; but here, as before, the number of measures in a strain must always be either two or three, or some product of these numbers: for here five bars in one strain is not used, and seven proves much more intolerable. (part 2, §14, 290)

Thus the action of creating isochronous groups or chunks (i.e., groups of groups) does not arbitrarily stop at the measure but extends to phrase lengths in terms of measure groupings, that is, numbers of measures, which, like rhythm, generally comprise multiples of twos and threes.[13] This comparison is not naive, as Holden is intent on establishing grouping as a universal feature of our perception. Therefore he argues that, in addition to influencing the domain of sight, the imperative of grouping into the powers of twos and threes affects our hearing, specifically the perception of rhythm and form along a temporal continuum, subject to the constraints of our memory.

Holden now makes a bold analogy between our perception of rhythm and form and that of pitch and tonal relationships:

> The faculty of remembering the key note [i.e., the tonic] and the constant expectation of returning to it at the conclusion, which is so remarkably perceived by the musician, resolves immediately into that of retaining the idea of a small portion of time, divided and subdivided in some eligible manner, by the vibrations of the same key, or its octaves; and agrees exactly with the remembrance of the length of one bar or strain, and of the mood of time, in the timing of music. (part 2, §20, 298–99)

Holden here asserts that similar cognitive abilities enable the relational perception of pitch, scale degrees, and rhythm. The opening terms of his analogy—the retention of the keynote and the expectation associated with a sense of key—take up a theme that he treats at length in the first part of the *Essay*, namely, the roles of memory and syntactical expectation. Logically, to fix a keynote in our minds, we must be able to distinguish it from other pitches, and in the first chapter of the *Essay* Holden takes the unusual step of characterizing the particular quality of each scale degree: "The key note is remarkably bold and commanding; the third and seventh have something supplicative in them. . . . The sixth is a kind of plaintive sound; the fourth, as observed

13 Holden's reference to the constraints of memory in regard to the creation of nested hierarchies of groups is highly reminiscent of the psychological concept of chunking. See Miller 1956.

before, is grave and solemn; the fifth partakes of the nature of the key, and the second is not unlike to the sixth" (part 1, §21, 14–15). Holden emphasizes that these properties depend on the tonal context in which the notes occur, thereby identifying them as relational rather than inherent in the notes themselves or in their physical means of production. That is, "if we were only to consider musical sounds singly, without any regard to their relations to, or dependencies upon each other; no such properties as these could be attributed to any one sound more than another" (part 1, §21, 15).[14]

Given that the particular effect of each scale degree arises solely out of its relationship to the key, Holden argues, we must "keep the key note constantly in view during the whole course of a tune; and to consider all the other notes of the scale, chiefly with regard to their several relations to the key" (part 1, §28, 23). The fact that all pieces of music end in the home key is further proof that we must be preserving the key in our minds throughout, as "if ever [the listener] were supposed to have totally forgot it, what then could hinder him from being as fully satisfied with a final close, upon some other note?" (part 1, §29, 25). Consequently, the act of sustaining the keynote in our awareness gives rise to the familiar phenomenon of musical expectation.

To illustrate this point, Holden asks his reader to imagine the experience of listening to someone singing or playing a well-known psalm tune. Were it to break off before the end, he maintains, this would be "quite unsatisfactory, and a plain disappointment to the hearer. This natural expectation, which never ceases till some proper succession of notes occur, leading to, and terminating upon one certain sound, is not peculiar to a few tones only, but is common to all" (part 1, §29, 23–24). Indeed, he argues, perceiving pitches in relation to a key is "both necessary for the purposes of music and natural to the human mind" (part 1, §32, 28).

Holden summarizes his claim as follows:

> In the practice of music, the key note is constantly kept in mind; and all other notes which are admitted, are some way compared with the key. These comparisons, and the consequent perceptions, are, indeed, the very essence of music. It is impossible for us to hear two different sounds, either together, or in succession, without attempting to make some comparison, either between one and the other, or between each of them, and some third sound, with which our mind may previously be possessed, and which we regard as a key note. (part 1, §32, 28)

According to this theory, retaining the keynote in our minds provides a basis for the comparisons of sounds, which are essential to the perception of music, thereby giving rise to a sense of key and, consequently, tonal expectation. Yet to retain and evaluate—that is, to "attend" to—musical sounds occurring in

14 Holden writes, "We might indeed ascribe different properties to the sounds of different instruments; thus, the trumpet is bold, the violin is chearful, the bassoon is solemn, etc" (part 1, §21, 15).

time, we need to have some kind of a frame of reference, some way of measuring and comparing incoming pitches against the keynote as a ground. To enable this, and to support his analogy between the perception of time and pitch, Holden introduces the concept of the module.

2. The module

Holden defines the module as "the small portion of time, by which we suppose the vibrations not only of the key [note], but also of every other sound which we admit into our music, while we retain that key, to be measured and distributed into isochronous parcels" (part 2, §21, 299).[15] The module, that is, is both a short span of time that defines a certain frequency and a point of reference akin to a measure in rhythm, constituting an imaginary quantity held in the mind against which other sounds are perceived. (It is thus, in effect, a way of transforming frequency relations from physical events into unconscious mental acts.) The module has a similar status to a musical bar: it is a durational window in relation to which shorter durations stand as fractional divisions and are thereby measured (as a centimeter is measured as a defined fraction of a meter), and conversely, longer durations are measured as multiples (as a kilometer is measured as a definite multiple of a meter). In the latter way, the module participates in larger-scale groupings. It therefore functions as the basis of the mental process of parceling, in that it is continually hierarchically subdivided into smaller equal durations.

The analogy between a module and a measure is illustrated in Figure 2. Holden does not specify the number of vibratory cycles or pulses that should be used to constitute the module but simply notes that, for "whatever small portion of time be taken for a module, there may always be a pitch or tone of sound, whose vibrations shall divide the same module in any assignable manner whatever" (part 2, §22, 299). There will therefore always be a given vibration that corresponds to the size of the module itself, which is by definition some octave of the keynote, as depicted by the single wave in the bottom left corner in Figure 2.

Holden asserts that, as with the metrical bar, the primary division of the module must be into two or three isochronous parts.[16] Of course, he notes, for any module there is an infinite number of pitches whose vibrations divide that module in an infinite number of ways. Therefore, in the context of a musical work, we include only the pitches whose various frequencies "divide the module in certain simple and intelligible manners, and such we say have

15 Holden frequently uses *key* to indicate the keynote—the tonic itself—rather than the key in the sense of the tonality.

16 A single pulse cannot determine a module, Holden notes, as "three pulses only determine two of the small intervening particles of time; and till we have perceived two or more of these particles, we can form no judgment concerning their equality; that is, we cannot estimate the tone of the sound" (part 2, §25, 302).

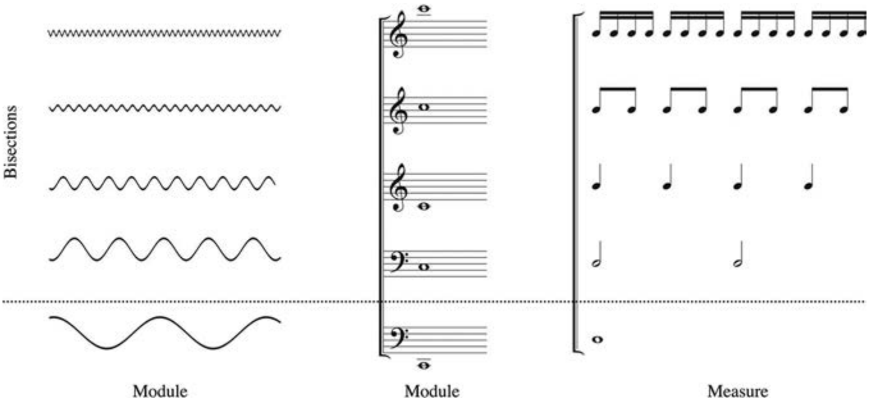


Figure 2. An illustration of Holden's analogy of the module as a measure and as a pitch. Left and center: the harmonics of a fundamental sound equivalent to the module represented as vibrations and as pitch; right: the same represented as a measure and equivalent rhythmic subdivisions.

eligible parcels" (part 2, §22, 299). "Eligible parcels" are those meeting the cognitive imperative, in Holden's view, that the number of parts into which any module is divided be a product of powers of 2 and/or 3. Divisions involving fives or sevens are also possible at times, he notes, but only after an initial division into two or three has been attained. Such units are "admitted, but with more difficulty, and generally only as dependents, or *harmonics* . . . [while] those whose parcels involve any of the higher prime, or uncompounded numbers 11, 13, 17, etc. are totally rejected" (part 2, §23, 300).[17] Holden uses these constraints to conceptualize the proportional pitch relations constituting the diatonic scale as divisions of the module.

Holden suggests that we normally establish the module for a piece of music on the basis of the piece's first pitch or chord:

> Supposing our ear to be entirely unbiassed, and not retaining the least impression of any former heard sound, when a musical sound is first proposed: in this case, we shall most naturally regard this first heard sound as a principal key note, and parcel its vibrations by continual reduplication, or by the powers of 2, rather than in any other practicable manner; and thus we shall constitute a module divided and subdivided by continual bisection, like the measure in common time. (part 2, §24, 300)

In the absence of other contextual cues, therefore, our minds extract and constitute a module from the vibrations of the first sound we hear, which is generally the keynote.

17 Holden uses the term *harmonics* here (in the sense of harmonic partials) to indicate a further quintuple or septuple division of a primary duple or triple division of the original frequency.

Holden further uses the module to speculate on the nature of our hearing. The entire range in which we can perceive pitch, he observes, comprises approximately seven octaves from F1 to F8 (part 2, §25, 302).[18] He therefore proposes the following hypothesis: since F1, being the lowest audibly identifiable pitch, has the longest module, and F8, being the highest, divides that module into the maximum possible number of parts, the latter represents "the smallest mental subdivisions which we can make in a [i.e., any] given module, [which] may be the cause why no sound still acuter can be admitted into our music" (part 2, §25, 302). Here Holden, in a remarkably bold stroke, effectively portrays the physiological limits of human hearing as a consequence of the cognitive restrictions entailed by the act of mental grouping. Our music, he concludes, does not feature pitches that cannot be distributed into groups, such as pulses whose rapidity exceeds our mental capacity to perceive their durational relation.[19] The size of the parcels therefore represents the cognitive limits of perceivable proportions. Importantly, even though Holden chose to use relative rather than absolute values to construct the module, he was familiar with the concept of frequency.[20] His decision to rely on a perceptual measure, rather than a unit anchored to absolute time, arises from the nature of the module, which functions as a cognitive framework against which other pitches can be hierarchically organized and perceived.

It is important to note that, as an idealized span of time, the module exists in our minds rather than in the physical world. It thereby joins a host of other musical phenomena that exist only in our imagination, such as the aforementioned case of subjectively rhythmicized metric accents, as well as the fundamental bass or implied dissonances, which are absent from a given sonority but are deducible from the empirical reality. Unlike these constructs, however, Holden proposes that the module is constantly retained over long spans of time—or at least until the mind can no longer parse sounds according to a given module, at which point it selects another. Although he does not state this explicitly, the module appears to be that instrument of the mind, so

18 Holden's estimation of the physiological limitations of our hearing range resembles a claim made by Leonhard Euler (1739, 8) in *Tentamen novae theoriae musicae*, namely, that the ear can perceive sounds within the range of eight octaves corresponding to 30 and 7,552 vibrations per second. However, Euler and Holden use different methods and provide distinct results. Euler bases his conclusion on the sounding frequency, that is, the number of vibrations per second, whereas Holden examines idealized numbers. Additionally, Holden's hearing range is more limited, comprising "more than six, but less than seven successive octaves" (part 2, §25, 302). While the *Essay* includes references to the music-theoretical writings of Rameau, d'Alembert, Serre, Tartini, Lampe, Pasquali, and Robert Smith, as well as Brossard's musical dictionary, Rousseau's entries on *systeme* and *battemens* from the *Dictionnaire de Musique* (1768), and d'Alembert's entry on *fondamental* from *Encylopédie* (1757), Holden does not cite Euler or mention his ideas.

19 Holden uses the terms *vibrations* and *pulses* nearly interchangeably while observing the distinction that two vibrations entail three pulses.

20 Holden mentions experiments reported by William Emerson in *The Doctrine of Fluxions* (1748) and by Robert Smith in *Harmonics; or, The Philosophy of Musical Sounds* (1749) and compares their measurement of frequency to refute the possible argument that the fastest rhythmical subdivisions used in music might influence our conception of the module (part 2, §47, 323–24).

to speak, which affords the mental acts of comparison required to hear pitches as participating in distinct relationships to a key.[21] To clarify this matter, I turn to his account of the major scale.

3. The ascending and descending major scale

Holden's first principle of music—the cognitive strategy of grouping all incoming equal stimuli into nested groupings comprising small primes—finds its ultimate expression in the major scale. He regards the proportional relationships between scale degrees as revealing our innate disposition toward perceiving pitches by means of grouping into small prime factors in comparison to a module. Indeed, in the first part of the treatise, he maintains that the scale illustrates "one of those laws which the great Author of Nature prescribed to himself, in the formation of the human mind, that such certain degrees of sound should constitute music" (part 1, §38, 37). Hearing pitches as belonging to a certain tonal context thus entails accessing idealized mental representations of divisions of the module.

According to Holden's theory, movement between scale degrees involves shifting among eligible groupings of the module, which means that each degree must stand in proportion to an initial duple or triple division of the module by means of some combination of the factors of 2, 3, 4, 5, and 7.[22] He justifies the inclusion of the latter three factors by appealing again to the case of musical rhythm. In performing a quintuplet or septuplet, he observes, we usually understand them as slightly faster versions of more familiar divisions into four and six, just as when we see a row of five or seven equally spaced windows, "we readily conceive 5 by its affinity to 4, and 7 by its affinity to 6" (part 2, §29, 305).

Holden next analogizes between the experience of visual grouping and that of hearing scale degrees in relation to a key. He proposes that in hearing, just as in vision, we can directly grasp the simple factors of 2, 3, and 4 but that the higher primes of 5 and 7 can be perceived only in reference to these simple factors. Thus we perceive 5 as a sort of variation of 4 by the addition of 1, and 7 likewise with respect to 6. Scale degree numbers whose factors include the higher primes of 5 or 7 therefore must be mediated by these smaller and simpler factors, which in the domain of tonal music appear as lower pitches that stand to those scale degrees with factors of 5 or 7 as fun-

21 In the first part of the *Essay* Holden depicts the experience of hearing a melody that remains in a single key (his term is *connected*) as follows: "A whole tune is often in reality no more than a kind of division or breaking upon the key note as fundamental. The key might be held on in the bass from first to last, as in the musette and bagpipe music; and although it should not actually be so held on, yet it is undoubtedly always kept in mind in all connected pieces" (part 1, §321, 278). The equivocation in this passage, whereby we could retain the keynote as a drone, although we do not actually do so, bolsters the interpretation of the module as the representation of the key that is sustained in our minds throughout a single diatonic piece.

22 Of course, as the square of 2, 4 ought not to be properly included in a list of prime factors, but Holden includes it for the sake of the proportion 5:4 (the ratio of the major third).

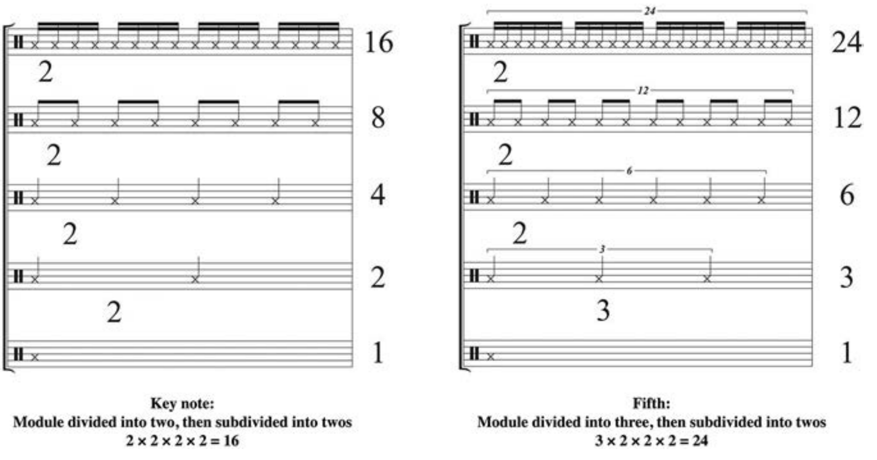


Figure 3. Key and fifth divisions of the module.

damental basses representing solely primary (i.e., duple and triple) divisions of the module. In addition to the imperative of grouping, Holden's scale derivation relies on another property of our perception. Once heard, both the pitch and its fundamental bass influence the way in which we perceive subsequent sounds (and generate their fundamentals) in the present. As a result of this directional feature, Holden assigns scale degree IV different divisions of the module according to whether the scale ascends or descends.

To illustrate Holden's approach to the scale, I begin by considering scale degrees I and V, each of which has as its fundamental a note lying one or more octaves below it with the same letter name as itself, so that they belong to what I henceforth call the same *note class* (part 2, §27, 304).[23] Moving from I to V entails only a single change from a duple to a triple division, as both can be immediately perceived vis-à-vis the module. This is illustrated in Figure 3, which depicts the module as the duration of a measure of 4/4. The initial duple or triple division of the module determines the note values that follow: the note values from the left column stand to the right column in the proportion of 2:3. As the module is equivalent to a far shorter amount of time corresponding to some number of vibrations, we can regard the note values in this illustration simply as idealized frequencies, and their proportions as intervals.

To clarify the divisions of the module shown at each hierarchical level in Figure 3, Holden uses whole numbers, rather than fractions, to discuss the divisions of the module and hence proportions of the scale in relation to

23 Holden regards octave equivalence as fundamental to his analogy between rhythm and pitch, claiming that just as "the division of a crotchet or other short note, into 2, or 4, or 8, makes no alteration in the nature of the time in music, so the doubling and redoubling the number of vibrations in a parcel, makes no alteration in the effect of such parcel on our sense" (part 2, §19, 298). I thank David E. Cohen for suggesting the concept of note class.

Table 1. The descending scale represented as divisions of the module

Note	*Scale degree*	*Division of module*
C	Key: 32	$2 \times 4 \times 4$
B	Seventh: 30	$3 \times 2 \times 5$
A	Acute sixth: 27	$3 \times 3 \times 3$
G	Fifth: 24	$3 \times 2 \times 4$
F	Grave fourth: 21	3×7
E	Fundamental great third: 20	$2 \times 2 \times 5$
D	Second: 18	$3 \times 2 \times 3$
C	Key: 16	$2 \times 2 \times 4$

some pitch.[24] He describes a module that in its simplest form is simply a fixed period of time in which pitches are compared. This creates a cognitive reference point to which higher pitches, whose frequencies relate to the module by integer ratios, can be proportionally related. Holden notates these tones by numbers, for instance, 24. This pitch relates to the module in a 1:24 ratio, which is to say that, assuming a fixed module, the numbers can be regarded as relative frequencies. For example, if the module is F1, then the number $24 = 3 \times 2 \times 2 \times 2$ describes C4.

In line with his understanding of the major scale as innate, Holden regards the proportions of the scale degrees—that is, divisions of the module—as fixed. The proportions of the descending scale are listed in Table 1, along with each scale degree's corresponding division of the module into small factors comprising 2, 3, 4, 5, and 7. Note that scale degree IV (21, or 3×7) is a *grave fourth*, which is to say, a note class equivalent to the natural seventh partial of the dominant scale degree, and is therefore lower than the perfect fourth (4:3), which appears in scale degree IV of the ascending scale, by a comma of 64:63. Holden describes the resulting discrepancy between the two kinds of fourths as his own invention, remarking that as this "small interval, not having formerly been supposed to exist among musical sounds, has no established name; we shall therefore call it a bearing; not chusing to borrow from any other language the name of an interval which a Briton first introduces" (part 2, §44, 317).[25]

24 Holden may have adopted this strategy from Euler, who likewise represents pitches using whole numbers in his *Tentamen*, although, as mentioned earlier, Holden does not cite this work or otherwise refer to its ideas. Jean-Adam Serre (1753, 133–34), whom Holden cites on other subjects, transmits this information in his *Essais sur les principes de l'harmonie* in a section titled "Observations sur les formules de M. Euler, qui représentent les sons du mode majeur, & ceux du mode mineur," which critiques Euler's algebraic formula for the minor scale.

25 The ratio 64:63 is precisely the amount by which the ratio of the major whole tone, 9:8, is exceeded by the larger whole tone with ratio 8:7 that lies between V and the grave fourth. Thus 72:64 = 9:8, while 72:63 = 8:7. These relationships are further detailed in Holden's chart, reproduced in Figure 10. In the "Conjecture sur la raison de quelques dissonances généralement reçues dans la musique," Euler (1766, 172) likewise makes use of 64:63, claiming that the proportions of the dominant seventh chord are 36:45:54:64 but that they are heard as 36:45:54:63 and thus reducible to 4:5:6:7. I thank Roger Grant for this observation. It seems unlikely, however, that Holden was influenced by the "Conjecture," as he issued *A Collection of Church-Music*, a compilation of twenty-four new hymn settings advertising his forthcoming *Essay* the same year. In the preface Holden (1766, 1–12) claims that his only predecessor in

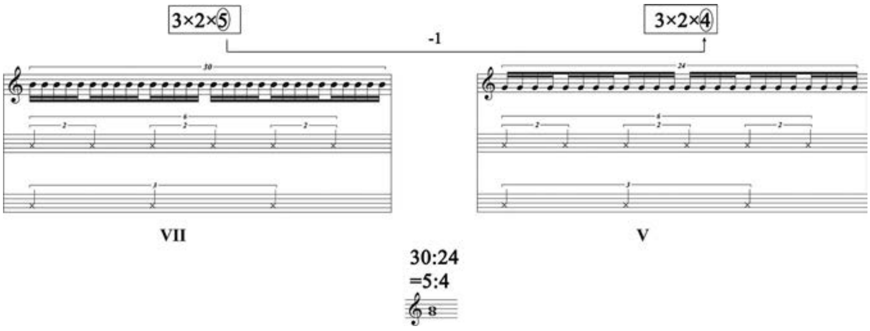


Figure 4. Perceiving VII within the context of a scale means hearing it in relation to V.

One example is the transition from scale degree VIII (32), a purely duple division of the module, to scale degree VII (30), which decomposes into the factors of 2, 3, and 5. In Holden's words, "We take the seventh instead of the fifth; that is, we take 30 vibrations instead of 24; and this presents a module divided, like the fifth, into 3; but each of these subdivided into 5, instead of 4; or 10, instead of 8" (part 2, §33, 308). In this rather terse formulation, Holden proposes that we regard scale degree VII ($3 \times 2 \times 5$) in relationship to scale degree V ($3 \times 2 \times 4$), as these degrees share two of three factors. Therefore, "as we conceive of 5, by its affinity to 4, so, while we sound the seventh of the scale, the fifth is essentially implied, and is our fundamental" (part 2, §33, 308). Holden here understands the fundamental bass as a consequence of our mind's attempt to grasp the incoming sounds in terms of various duple and triple primary divisions of the module. According to this approach, every scale degree will be accompanied by a fundamental bass note internally generated in our imagination via a primary duple or triple division of the module. When we move from VIII to VII, he argues, our imagination instantly projects their corresponding fundamental basses, namely, I and V, respectively, an act that essentially reflects how our mind perceives tonal relationships.[26] We can compare this to the experience of sight: recall Holden's claim that when we view five windows, our minds involuntarily sort them into two groups of twos and a remainder of one by perceiving them as two twos with a surplus, as illustrated in Figure 1e. The relationship between scale degree VII and its fundamental, V, is illustrated in Figure 4 by an analogy with rhythm and in Figure 5 by the analogy with vision.

using 7 as a musical ratio was William Jackson (in *A Scheme Demonstrating the Perfection and Harmony of Sounds* [1726]).

26 Holden's presentation of the scale as composed of scale degrees accompanied by fundamentals subject to constraints on their possible motion responds, of course, to Rameau, who repeatedly addressed this subject, starting with the *Traité de l'harmonie réduite à ses principes naturels* (1722), which Holden cites in the *Essay*, though not on this point.

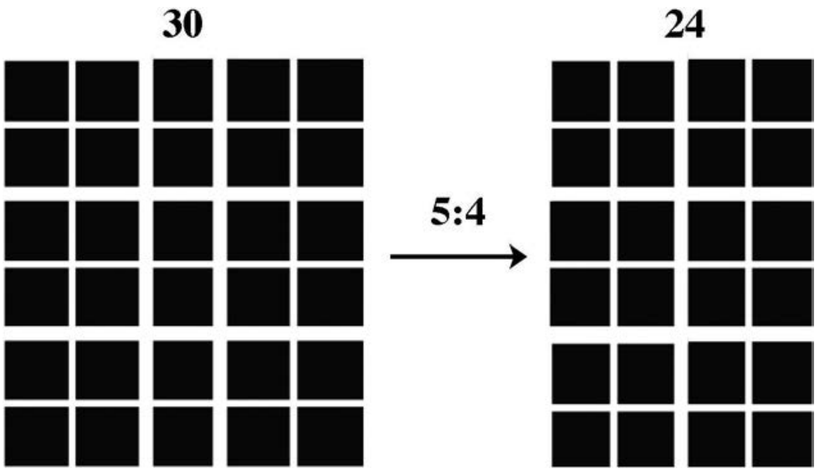


Figure 5. We can compare hearing VII:V to the cognitive act of grouping that transpires when we view windows. If we see 30 windows, we view the five rows as four groups (of six windows), with an extra group included in the middle, analogous to Figure 1e.

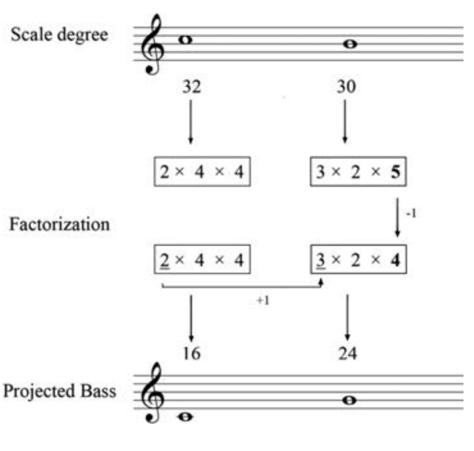


Figure 6. Movement between scale degrees implies movement between fundamentals.

To summarize, in the context of a scale, moving from scale degree VIII ($2 \times 4 \times 4$) to VII ($3 \times 2 \times 5$) involves simultaneously moving from fundamental bass VIII or I ($2 \times 4 \times 4$ / $2 \times 2 \times 4$) to V ($3 \times 2 \times 4$), as shown in Figure 6. (Again, I treat VIII [$2 \times 4 \times 4$] and I [$2 \times 2 \times 4$] interchangeably, as they are members of the same note class.)

Holden thus posits that the listener's mind automatically and unconsciously supplies an appropriate accompanying fundamental bass note below each melodic scale degree heard. Therefore diatonic melodic motions among such scale degrees, for example, from scale degree I to VII, entail coordinated motions among those unconsciously provided fundamentals. In this regard his theory is reminiscent of Rameau's, but with significant differences. For Rameau, as David E. Cohen (2001, 71–72) has shown, the agency within the listener responsible for this activity is not the mind but a faculty that he calls "the ear" and later "instinct." However, the detailed operations of this faculty and the specific means by which it achieves its results remain occult. In contrast, Holden assigns the generation of fundamental basses to the mind rather than the ear and provides a coherent explanation for how we accomplish this task by proposing that the fundamental bass is an artifact of our overarching cognitive tendency toward grouping certain kinds of (auditory and visual) stimuli into nested hierarchies of small primes. Thus, while both Rameau and Holden assume that our projection of fundamental basses is unconscious, Holden's theory provides a detailed and mechanistic account for how exactly the precise pitches are generated within the mind.

A number of consequences follow from Holden's conception of the relationship between the scale degrees and the module as mediated by projected fundamentals. Given that we are dealing with factors of 2, 3, 4, 5, and 7, the factorization of any scale degree will differ from that of its projected fundamental by a value of either 0 or 1.[27] To think of this in more music-theoretical terminology, this means that the relationship between fundamentals and scale degrees will be one of either note class identity or a superparticular ratio.

In the *Essay* Holden walks us through his derivation of the scale, invoking rules on an ad hoc basis. However, his language suggests that he is thinking algorithmically, in that he discusses the influence of past and future moves on determining the current step. Here I first present my distillation of the constraints that are implicit in his scale and then interpret the scale in the form of a directed graph.

Holden's implicit "rules" can be summarized as follows:

1. "The real fundamentals themselves include only the numbers 2 and 3 in the composition of their parcels" (part 2, §39, 311).[28]
2. The keynote, fifth of the descending scale, and fourth of the ascending scale take a member of their own respective note classes as their fundamental.[29]
3. Any scale degree with a 7 or a 5 in its factoring can only be conceived as being in a superparticular ratio with its fundamental.[30]

27 Thus, for example, scale degree V (24, or 3 × 2 × 4) is its own fundamental, whereas scale degree VI (27, or 3 × 3 × 3) accepts a fundamental bass of II (18, or 3 × 3 × 2), or 3:2. See Figure 7 for the complete inventory of scale degrees and fundamentals.

28 Again, I treat 2 and 4 interchangeably for the reasons discussed above.

29 Holden writes, "The next note is the fifth itself, which we naturally take as a fundamental" (part 2, §36, 309).

30 According to Holden, "In the descending scale . . . the sounds which include the number 5 in the composition of their parcels naturally become great thirds [to their fundamentals], and the sound which includes the number 7, will be either a less third, or an added seventh [to its fundamental]" (part 2, §39, 310–11).

4. Movement between fundamentals can only occur by fifth; that is, only a single bisection (group of three shifting to a group of two) or a trisection (group of two shifting to a group of three) is permitted at a time.[31]

Figure 7a collates the transitions between scale degrees (and their corresponding fundamentals) that are compatible with rules 1–4. The rows and columns on the right show the compatibility or incompatibility of each scale degree with each of the available fundamentals, listed in the leftmost column. VIII, V, and I are compatible only with themselves (i.e., pitches one or more octaves below them), and VI and II are compatible only with II and V, respectively. III is compatible only with I, and VII with V, as in both cases the factor 5 must be understood in terms of 4. A more challenging case arises in regard to scale degree IV, which appears compatible with the fundamentals II and V, both of which would have been eligible according to rule 3. At this point Holden argues that there should be a local accordance between fundamentals and their respective scale degrees, as well as a consistency of transitions between divisions of the module that are retained in memory. Recall that, according to his aforementioned grouping principles, we perceive seven against six (i.e., 7:6), which means that IV is comprehended in proportion to II as its fundamental. Indeed, he writes, this would have been the case

> if the succession [i.e., the fundamental-bass progression] would have admitted it; but perceiving that the third of the scale is next to follow, which having 5 vibrations for 4 of the key, will inevitably require the key for its fundamental. . . . If we here take the second as fundamental, we cannot take the K*f* [i.e., the keynote, or tonic] immediately after it: we therefore refer the grave fourth, and its implied second to the fifth fundamental which existed in the preceding note, and to which the fourth becomes an added 7th, as the second is its 5th. (part 2, §37, 310)[32]

As Figure 7a shows, pairing IV with fundamental II would leave us too distant from fundamental I and cause us to break the fourth rule in the future (fundamentals II and I are not a perfect fifth apart). Therefore, instead of taking II as our fundamental, we conceive of IV as an added seventh to the preceding scale degree (V) and retain its fundamental. The remaining movement from III to II to I is straightforward. Note the elegance of this solution: the descending scale is fully compatible with the four rules above, as shown in Figure 7a, and no other possible solution fulfills all these conditions. The

31 Holden writes that, "after the module has been heard divided simply into 3 parts, we may, at the next step, proceed to trisect each of its third parts; and, by this means, introduce a division into 9, which corresponds with the second of the scale: but we cannot easily reconcile the making of two such trisections at one step, that is, we can substitute 3 vibrations instead of 4, and contrariwise, 4 vibrations instead of 3, but we cannot immediately substitute 9 vibrations instead of 8 nor 8 instead of 9, which is exactly the case in the timing of music also" (part 2, §35, 308).

32 K*f* here refers to the fundamental of the key, or I.

Possible Fundamentals	Scale Degree and Factorization							
	C (VIII) 2×4×4	B (VII) 2×3×5	A (VI) 3×3×3	G (V) 3×2×4	F (IV) 3×7	E (III) 2×2×5	D (II) 2×3×3	C (I) 2×4×2
C (I / VIII) 2×4×2/2×4×4	✓	✗	✗	✗	✗	✓	✗	✓
G (V) 3×2×4	✗	✓	✗	✓	✓	✗	✓	✗
D (II) 2×3×3	✗	✗	✓	✗	(✓)	✗	✗	✗
A (VI) 3×3×3	✗	✗	✗	✗	✗	✗	✗	✗

Figure 7a. Factorization and directed graph of the descending scale. The flawed option of taking II as the fundamental to IV is shown in brackets.

Figure 7b. The resulting descending scale. Top: proportions of scale degrees among themselves; middle: proportions of scale degrees to projected fundamentals; bottom: proportions between projected fundamentals.

proportions between adjacent scale degrees, their successive fundamentals, and each scale degree and its fundamental are shown in Figure 7b.

Holden next introduces the ascending scale by remarking that we know from experience that there are other ways to derive scale degrees, namely, by continual trisection, which generates a scale degree III of 81 and a VII of 243.[33] He suggests that a trained musician can opt to perform III:I and VII:V as 5:4 "in order to confine his fundamental progression more closely to the original key," or he can relate III:VI and VII:III as 3:2, which is "no more than what nature suggests to the most uncultivated singer" (part 2, §42, 314).[34] He therefore includes alternate variants of certain scale degrees as possible options within his system, depending on context.

33 Assuming octave equivalence, the difference between the third scale degree generated according to Holden's given ratio of 5:4, which renders a major third represented by the number 80 (= 5 × 2^4), and a major third generated by four stacked fifths resulting from four trisections of the module, represented by 81 (3^4), is the syntonic comma, 81:80. The same comma occurs between his seventh scale degree (15:8) and a seventh generated by continual trisection (3^5), or 243, which, compared to 240 (the 15 of 15:8 multiplied by 2^4), reduces to 81:80.

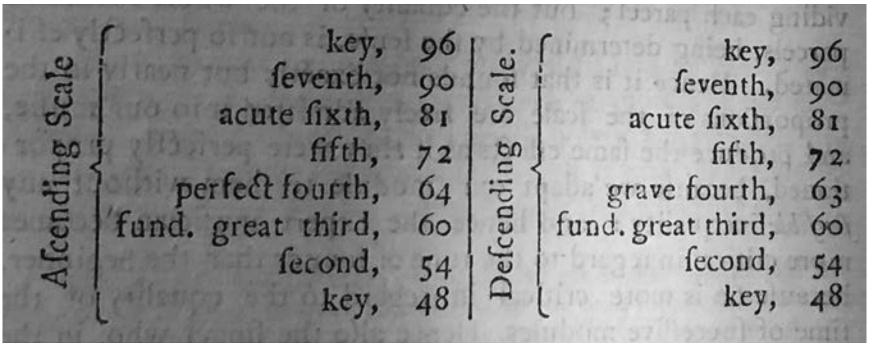
34 These proportions, Holden observes, are found in Scottish folk tunes, which are notoriously difficult to harmonize (part 2, §42, 314).

Referring to musical practice, Holden emphasizes that we frequently replace a given sound with another a comma away without noticing the discrepancy. Moreover, he observes that we can substitute not only sounds but also projected groupings: "These different ways of conceiving the third and seventh of the scale may effectually take place, although no alteration be made in the real pitch of the sounds" (part 2, §42, 314–15). This is possible because our senses generally cannot detect the difference when we vary the length of our module between 80 and 81. Holden compares this to the case of musical timing, where "one bar may be, and always will be some very small matter longer than another, perhaps much more than one eightieth part of the whole bar, and yet we are sensible of no impropriety" (part 2, §42, 315).[35] He thereby argues that a module derived from the keynote is actively retained in the mind and used to evaluate and constitute successive pitches such that they reflect the specific proportions of module divisions that are characteristic of the diatonic scale. This mental strategy holds even in cases where, in acoustical reality, there is a discrepancy between the proportions of actual pitches and idealized divisions of the module (e.g., as is the case in fixed keyboard instruments). The mental action of extracting a module from the keynote and sustaining it in our minds thus allows us to perceive the various scale degrees as the different divisions of the module, which is precisely the feature that lends them their distinct tonal qualities.[36]

Holden appeals again to his earlier distinction between the senses and the mind. Recall that the former conveys the scale degree to the mind by assessing the equality of successive pulses, while the mind, once it has received a sound, groups the pulses into small units and determines their proportional relationship to the module by means of a projected fundamental. He argues that "the mind allows not the least deviation from the proper method of dividing and subdividing each parcel; but the equality of the whole successive parcels, being determined by the sense, is not so perfectly estimated" (part 2, §42, 315). As the senses are responsible for simply determining the isochrony of the pulses that constitute a given sound, pitches that are close enough to the proportions of the scale "produce the same effects as if they were perfectly proportioned, because we adapt our module to them without any *sensible* inequality" (part 2, §42, 315). This also explains why an expert

35 Holden appears here to intuit the phenomenon of just noticeable difference, subsequently formalized by Gustav Fechner as Weber's law in 1860 based on findings by Ernst H. Weber in *De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae* (1834). Descriptions of our mind's tendency to approximate pitch discrepancies so as to conform to integer ratios can be found in a letter by Leibniz (1988) to the mathematician Christian Goldbach in April 1712 and in Euler's "Conjecture"; however, Holden's generalization of this principle to encompass both pitch and timing appears to be unique. His insights anticipate research on microtiming in the twentieth century (for a detailed summary, see Goldberg 2017, 8–33).

36 This notion anticipates Riemann's (1992, 90) assertion that our "tonal imagination" (*Tonvorstellungen*) leads us to perceive various pitches and chords in ways that sometimes significantly deviate from hypothetical tuning differences.


Aſcending Scale
key, 96
ſeventh, 90
acute ſixth, 81
fifth, 72
perfect fourth, 64
fund. great third, 60
ſecond, 54
key, 48

Deſcending Scale.
key, 96
ſeventh, 90
acute ſixth, 81
fifth, 72
grave fourth, 63
fund. great third, 60
ſecond, 54
key, 48

Figure 8. Holden's ascending and descending scales.

Possible Fundamentals	Scale Degree and Factorization								
	C (I) 2×4×3×2 48	D (II) 3×3×2×3 54	E (III) 2×5×2×3 60	F (IV) 2×4×4×2 64	G (V) 3×4×3×2 72	A (VI) 2×4×2×5 80	A (VI) 3×3×3×3 81	B (VII) 3×5×3×2 90	C (VIII) 4×4×3×2 96
F (IV) 2×4×2×2	✗	✗	✗	✓	✗	✓	✗	✗	✗
C (I / VIII) 4×4×3×2 / 2×4×3×2	✓	✗	✓	✗	✓	✗	✗	✗	✓
G (V) 3×4×3×2	✗	✓	✗	✗	✗	✗	✗	✓	✗
D (II) 3×3×3×2	✗	✗	✗	✗	✗	✗	✓	✗	✗
A (VI) 3×3×3×3	✗	✗	✗	✗	✗	✗	✗	✗	✗

Figure 9. Factorization and directed graph of the ascending scale (the *double emploi* at scale degree VI is indicated in dashed lines).

musician is more sensitive to intonation than a beginner: because "he is more critical in regard to the equality of the time of successive modules" (part 2, §42, 315).

Figure 8 displays Holden's ascending and descending scales side by side. These are identical with a single exception: the former features a perfect fourth (64) and the latter a grave fourth (63).[37] Although Holden does not make this point explicitly in the second part of the treatise, he discusses *double emploi*, which he terms "double meaning," in the first part of his treatise.[38] As shown in Figure 9, coming from V, scale degree VI can be interpreted as 80:64, or 5:4 of IV, while in relation to VII it is heard as 81:64, or 3:2, of II.

37 The descending scale in Table 1 was fixed between 16 and 32; however, here, to obtain a full comparison with the ascending scale, Holden multiplies these ratios by 3. The justification for this move is discussed in section 4.

38 See part 1, §212, 190–91, for Holden's discussion of "double usage"; for more on Rameau's concept of *double emploi*, see Christensen 1993, 193–95.

Holden's harmonization of the scale refers to both harmony and melody and is identical to Rameau's (1737) *règle de l'octave* harmonization from *Génération harmonique*, which Holden cites elsewhere in the *Essay*.[39]

To summarize, retaining a module in our minds and using it to interpret successive pitches provide us with a constant sense of key. This aspect of our mental processing allows us to retain and attend to the keynote in our minds throughout a piece. Because, moreover, hearing different pitches as scale degrees requires us to mentally carry out distinct divisions of the module, the sense of key is maintained even when the actual pitch is not phenomenally present. Significantly, the various scale degrees can be conceived—and consequently heard—differently according to context. These variations, which are reflected in the distinct module divisions and fundamental bass progression involved in each case, help inform us about their tonal function.

4. Modulation

Holden maintains that we retain the key throughout a tune unless we encounter sounds that require a change in the size of our module. To account for this experience, he introduces the mental operation of attention as the faculty that enables us to recognize when sounds are part of, or foreign to, a given diatonic context: "When a person has fixed his attention on one sound as a key note, if other sounds accidentally intrude upon his ears, which belong to a different key, they have often a most disagreeable effect. . . . Before he can hear such incompatible sounds, with satisfaction, he must quit his own former key, and turn his attention altogether on that to which they properly belong" (part 1, §31, 27).

The experience of hearing pitches in relation to a key, Holden argues, requires us to have already fixed our attention on a given keynote, and he regards this act as already entailing the determination of a module. It should be emphasized that attending here does not necessarily entail a conscious awareness of the mental act. Holden appears to implicitly distinguish between—and invoke—both voluntary and involuntary forms of attention throughout the treatise.[40] The former is typically described using active verbs, whereby the listener fixes, turns, or places his attention on a certain sound as a keynote, as in the passage above, whereas in the case of the latter, the listener's attention is attracted, distracted, carried, or diverted by a sound, line, or other aspect.[41]

39 Compare Holden's Plate 6 Example XLII to Rameau's (1737, 129) Plate 15, *Génération harmonique*. Holden does not provide the generation of the ascending scale in the second part of the *Essay*.

40 A bifurcated approach to attention as composed of both voluntary and spontaneous aspects has been consistently theorized since the eighteenth century. See Hagner 2003, 672; and Riley 2004, 29–30.

41 For example, Holden observes that "our attention is carried either to the lowest or the highest sound which exists in a chord, rather than any of those which lie, as it were, concealed in the middle" (part 1, §144, 127).

Throughout the first part of the *Essay* Holden repeatedly depicts the mental effort involved in sustaining our attention as informing us about the tonal function of the pitches we encounter. For example, a beginner can determine the mode of a piece by attempting to see whether its keynote (tonic) accepts a major or minor third scale degree. This exercise succeeds, he writes, "because it is very difficult to turn our attention from one sort of scale to the other, while we retain the same key note" (part 1, §35, 32). The transition between major and minor thus requires a "change of attention," in which previous sounds "must be entirely disregarded, and quite new ideas substituted in their stead" (part 1, §35, 33).

There is, however, one exception: when moving between the most closely related keys, we can rely on our recollection of the original key to a large extent. This is because "there is such an affinity between the key and its fifth or fourth that we can for a little while turn our attention upon either of these as an occasional new key, without entirely relinquishing our principal key" (part 1, §31, 27). To make sense of truly extraneous sounds, therefore, we must reconfigure our attention around a new key center by deriving its concomitant module. The experience of hearing nondiatonic pitches thus gives rise to a change in attention, and this serves as an indication that we are in the process of modulating.

In accord with this theory, Holden understands divided attention as occurring under three circumstances: (1) when the module cannot provide a sufficient explanation or division for a given pitch, (2) when there are multiple potential explanations, or (3) when we have changed keys. The state of attending thus corresponds to our innate capacity for "retaining the impressions of musical sounds, for some considerable time after they cease to be heard, [which] is purely natural, and requires no improved abilities at all" (part 1, §31, 28). In contrast, the ability "to turn our attention readily from one key note, and the sounds dependent upon it, to another different key, which introduces with it a different sett [*sic*] of sounds, is an attainment due to practice alone, and a kind of force upon nature" (part 1, §31, 28). Consequently, the capacity to recognize and adapt to new modules requires us to develop our natural mental inclinations.

The implications of regarding the faculty of attention as participating in a broader system of music perception are fully revealed in Holden's conceptualization of the full chromatic system, which he obtains by juxtaposing the diatonic scale of the key with scales built in the identical fashion on IV and V, plus the (raised) leading tone of each of these three keys' relative minors. That is, a tonic key (with its relative minor) shares one module with the key of its adjunct fifth (and its relative minor) and another with the key of its adjunct fourth (and its relative minor), and these two modules stand in a 3:2 (or 4:3) relationship to the tonic. In other words, each module is shared by four keys, of which two—the tonic key and the relative minor key—are duplicates, except for the raised leading tone of the latter, which lies a major third in 5:4 ratio above scale degree III. Taken together, these two modules supply both the perfect (4:3) and grave (21:16) versions of scale degree IV, as well as both the just (5:3) and Pythagorean (27:16) versions of scale degree VI, allowing us to actively select which one to perform or mentally project.

Holden asserts, "In every passage of music our attention is partly divided, either between the key and its adjunct fifth, or between the key and its adjunct fourth, as fundamentals" (part 2, §45, 317). When we listen to a passage of music, therefore, our minds constantly retain and project—that is, attend to—two possible modules, each divided into its corresponding scale degrees. Both modules are necessary, as comprehending all twelve chromatic pitches via a single module would require its division into 3,072 (48 × 64) parts, that being the least common denominator of all the ratios present in Holden's complete chromatic scale. To clarify: by taking the same C to equal 48 units of one module and 64 of another, we can obtain whole-number module values for all members of the full chromatic gamut.[42] Shifting between these modules and "conceiving the principal key occasionally in one or the other of these ways" thus afford us the entire range of chromatic pitches (part 2, §45, 317). In a sense, we can think of moving between these two pairs of scales as representing a cognitive grouping strategy that enables us to comprehend highly complex proportions.

Holden encapsulates this theory in a chart titled "Scheme of the System of Modulation of C," shown in Figure 10. (*Modulation* here is used in the eighteenth-century sense of *Ausweichung*, a brief tonicization of one of the diatonic scale degrees of the key rather than a definitive change of tonal center.) The descending column on the left labels each row with an arabic numeral purely for ease of reference, while the roman numerals across the top and at lines 1, 5, and 9 represent scale degrees (diatonic or altered) with reference to C as keynote. Starting at the center and working outward, and ignoring the italicized letters between scale degrees for now (they are explained in the caption), line 5 simply contains the roman-numeral representation of the scale degrees of C major, as well as the label "scale of the principal Key." Line 6, labeled "Final," contains the numerical proportions for each scale degree taking C as 64,[43] while line 4, termed "Medial," likewise contains the numerical proportions for each scale degree but starts by taking C as 48.[44] The

42 The numbers 48 and 64 represent the smallest division of two modules standing in 3:4 proportion that afford the construction of the chromatic scale. Holden maintains that we cannot grasp this gamut directly, as he places an upper perceptual limit on possible divisions of the module at just under seven octaves, which means that the most detailed divisions of the module that we could still comprehend equal 3×2^7, or 384 (part 2, §25, 302). Thus "the highest power of 3, which we can admit in the parcelling of musical vibrations, is its fifth power 243, for the sixth power, which is 729, is unquestionably beyond our limits" (part 2, §41, 312).

43 That is, taking as reference pitch a C with a module divided into 64 equal parts.

44 Holden terms the relationship between the key and its adjunct fourth *medial* and that of the key and its adjunct fifth *final*. He uses these terms to refer to a key's degree of "perfection," that is, conclusiveness (see part 1, §146, 130). He argues that we cannot conclude on a medial chord or key, and at times he even uses the terms *medial* and *inconclusive* interchangeably (part 1, §197, 173). In contrast, he regards the key of the adjunct fifth as more per-

SCHEME OF THE SYSTEM OF MODULATION OF C.

N°	
1	IV. *t. g.* V. *t. l.* VI *ſ.d.* VII♭. *t. r.* K. *t. g.* II. *t. l.* III. *ſ.p.* IV. &c. Scale of the adjunct fourth.
2	32. 36. 40. 42. 45. 48. 50. 54. 60. 64. 72. 80. 84. 90. 96. 100. 108.
3	F. G. A. b♭. B. C. c※. D. E. F. G. A. b♭. B. C. c※. D.
4	Medial 48. 54. 60. 63. 72. 81. 90. 96.
5	Scale of the prin. Key K. *t. g.* II. *t. l.* III. *ſ.d.* IV. *t. r.* V. *t. g.* VI. *t. l.* VII. *ſ.p.* K. &c.
6	Final 64. 72. 80. 84. 96. 108. 120. 128.
7	G. g※ A. B. C. D. d※. E. F. f※. G. g※. A. B. C. D.
8	48. 50. 54. 60. 63. 72. 75. 81. 90. 96. 100. 108. 120. 126. 144.
9	V. *t. g.* VI. *t. l.* VII. *ſ.d.* K. *t. r.* II. *t. g.* III. *t. l.* ※IV. *ſ.p.* V. &c. Scale of the adjunct fifth.

Figure 10. Holden's chart of the relationship of C to F and G major. The module of lines 1–4 is divided into 48; the module of lines 6–9 is divided into 64. Note that the size of the interval between each scale degree is classified according to one of five options: the whole tone can be greater (*t. g.*), 9:8; lesser (*t. l.*), 10:9; or redundant (*t. r.*), 8:7. The semitone can be proper (*s. p.*), 16:15; or deficient (*s. d.*), 21:20. These intervals arise from the derivation of scale degrees from the factors of 2, 3, 4, 5, and 7 and affect the size of larger intervals. It bears repeating that these proportions represent cognitive divisions of the module rather than any kind of prescriptive system of tuning. In practice, Holden notes that the dominant seventh chords on the fifth scale degree of all three minor scales should be tempered to conform to the proportion of 4:5:6:7.

proportions taken here are those of the descending scale, which is to say that the fourth in both cases is grave (21:16). That is, we have here two identical representations of the descending scale of C that stand to each other in the proportion of 64:48, or 3:2.

Moving on to lines 1–3, line 1 contains the scale degrees in relation to C major as generated by the descending scale of its lower fifth, F (hence including B♭), labeled as the "Scale of the Adjunct Fourth," that is, the scale of F major but with its degrees numbered with reference to C as keynote; line 2 contains the proportions of this scale according to the same module we encountered in line 4; and line 3 contains their letter names. However, here the descending scale has been built on F, rather than C, and hence the grave fourth is located between F and B♭, at VII♭, and the fourth itself is a "bearing" higher than the fourth we encountered in the scale built on C, shown in lines 4 and 6.[45] This can be seen if we compare the numerical value of F in row 4 (63) and row 2 (64). Likewise, the A that is scale degree VI in C major (81) is a syntonic comma higher than the A that is scale degree III in F major (80). Lines 7–9 represent an analogous scenario in mirrored order: line 9 contains the scale degrees of C major and the label "Scale of the Adjunct Fifth"; line 7 contains

---

fect than that of the adjunct fourth, not only because every piece must conclude with a movement forming a perfect authentic cadence from V to I (part 1, §191, 167) but also because he maintains that it is more closely related to the key (part 1, §242, 221).

45 The interval of the bearing (64:63) was discussed in section 2; see also part 2, §44, 317.

the letter names; and line 8 contains the proportions of the descending scale on G, built according to the module we encountered in line 6. In this case, the grave fourth is located between G and C; hence the C in line 8 is a bearing lower than the C in line 6, and the syntonic comma is located between scale degree III in C major and the E in G major, which would have been scale degree VI. Taking all these elements together, we have C major, flanked by G major below and F major above. Furthermore, each of these three diatonic pitch collections is augmented with ♯V, the raised leading tone of its relative minor, and thus we have in effect six keys represented, and hence the full chromatic gamut.

Figure 11 offers an alternative way of representing Holden's "Scheme of the System of Modulation of C," and its derivation of the chromatic gamut by means of the scales of F, C, and G major and their harmonic minor relatives. Unlike Holden's chart, which depicts the relationship in terms of the scale degrees of C major, Figure 11 displays the degrees of each scale with reference to its own tonic. As shown in Figure 10 (and as necessarily occurs somewhere in any justly tuned diatonic scale), the minor triad built on scale degree VI, that is, the tonic of the relative minor, contains intervals that reduce to a Pythagorean minor third of 27:32 and a grave fifth of 27:40, both intervals being smaller than pure by the syntonic comma 81:80. Hence, to secure a justly tuned triad on VI to serve as the tonic of the relative minor, Holden raises scale degree v of A minor (E), D minor (A), and E minor (B) by the syntonic comma (shown in Figure 11 in rows 2, 7, and 12), allowing him to obtain A-, D-, and E-minor triads that reduce to the ratios 54:64:81, indicated in circles in Figure 11. However, like the two options for conceiving IV in major (63 and 64), discussed in section 3, there are situations in which v is conceived of differently. Thus, for example, a dominant seventh chord built on scale degree v of the relative minor calls on the grave v rather than the raised v. Furthermore, as shown in rows 2, 7, and 12, scale degree iv of the relative minor must be tempered by the septimal comma 36:35 to conform to the proportion of 7:6 vis-à-vis ii, the top interval of the justly tuned dominant seventh chord represented by the proportional series 4:5:6:7, and indicated in dashed squares in Figure 11.

Let us further examine the case of the tempered dominant seventh chord of D minor, the members of which are indicated in dashed squares in lines 2 and 5 of the chart. (In Figure 11 this corresponds to 35, 40, 50, 60; the seventh [G] should be moved up an octave, resulting in a root position seventh chord of 4:5:6:7.) Here, the seventh (G) has been (mentally) adjusted by the septimal comma from 36 to 35. Holden concedes that conceptualizing the G as a product of the factors of 7 and 5 can help "reconcile this chord to the imagination" (part 2, §86, 368).[46] Yet given the undeniable mental strain

46 He further posits that "the difficulty of conceiving such a parcel is the only reason why beginners in singing find it much harder to join a seventh to the leading chord of a flat [minor], than to that of a sharp [major] series" (part 2, §86, 368).

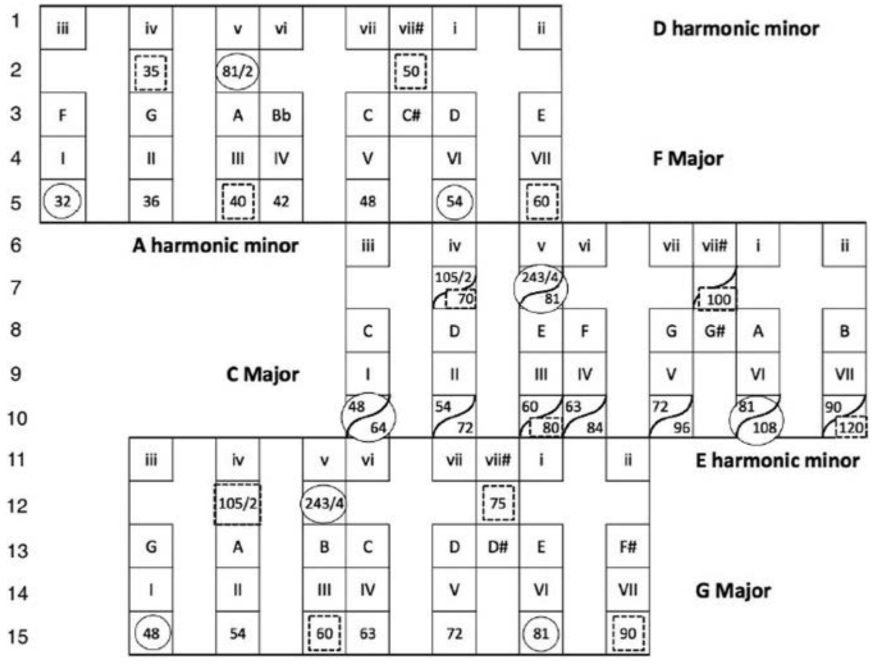


Figure 11. An illustration of the six different scales involved in the full chromatic gamut, with the adjustments required for minor. For ease of visual reference, roman numerals pertaining to minor scales are indicated by lower case letters. F major, D harmonic minor, and C major stand in proportion to a module of C taken as 48, while C major, A harmonic minor, G major, and E harmonic minor stand in proportion to a module of C taken as 64 (the raised leading tone G# of A harmonic minor is shared with this module alone). Adjusted scale degrees in minor keys are indicated in rows 2, 7, and 12; all other pitches are shared with their relative major keys, and identical to the numbers in Figure 10. The fractions in these rows are an artifact of the low module values of this graphic representation (which range between 32 and 90 for the first module, and 48 and 120 for the second module) and reflect the aforementioned tempering of the minor triad and the dominant seventh chord built on v of minor. Members of the tonic triad of the minor keys are indicated in circles, and members of the dominant seventh chord of the minor keys are indicated in dashed squares.

involved in comprehending a chord comprising only the higher prime factors 5 and 7, he observes that we must "conclude that the key must effectually be changed, and the IV or the V, must be received as a new key, before either of their substituted sixths [relative minors] can be introduced as such" (part 2, §87, 368–69). That is to say, according to this theory we cannot modulate to a minor key without first conceptualizing its relative major as a tonic. Most significantly, Holden here provides us with what is essentially a cognitive measure for knowing when we have modulated to a new key: if the mental effort of tempering the seventh of a leading chord is excessive, we should consider ourselves in a new key and conceptualize the module accordingly.

Holden is notably emphatic that alternative tunings for scale degrees conceived according to different proportions or modules should never "be depressed or elevated in an unnatural or forced manner" (part 2, §84, 366).[47] The designation of grave or acute describes the process of hearing a sound in relationship to a fundamental and a key. As such, it is intended only "to distinguish the different sounds of notes, which are apparently the same, on different occasions, according as one or another of the three scales takes place at the same time" (part 2, §84, 366).

5. Double fundamentals and implied sounds

As I describe in section 4, Holden regards attention as the faculty that determines the association of pitches within a given tonal context. In the first part of the *Essay* he proposes that our attention has a second function as well: it is the faculty by which we distinguish between consonance and dissonance, which he defines by writing, "When the several sounds mix and unite, in a manner agreeable to the hearer, it is called a *consonance*, or *consonant chord*; when they do not unite, but separately distract the attention of the ear, it is called a *dissonance*, or *dissonant chord*" (part 1, §125, 113).

This conception of consonance and dissonance is tightly linked with another distinction that Holden makes between concord and discord, namely, "two sounds are said to be *concord between themselves*, when both of them can be referred to one and the same fundamental perfect chord; and two sounds are called *discord*, when they cannot both be referred to one perfect chord" (part 1, §153, 138). A "perfect chord" for Holden is any major or minor triad, in any voicing, in root position or inversion.[48] Concord, then, for Holden is the common property of all of these triadic intervals, while discord is the property of all other nontriadic intervals.[49]

In characterizing concord, Holden notably foregrounds the criterion of triadic membership rather than any sonic character or aesthetic effect. Only subsequently does he invoke the traditional attribute of unified sound for concord, and this as a consequence of triadic membership by way of the mental phenomenon of attention: "Allowing that the mind *naturally chuses* to conceive every sound in music as belonging to some perfect chord, it is plain, that two sounds will seem to unite, when both of them are included in the idea of one perfect chord, but separately *distract our attention*, when this cannot be done, or when they must necessarily be referred to two different fundamentals" (part 1, §153, 138). Since discord is the subjective impression received

47 In *Collection of Church Music* Holden (1766, 5) also emphasizes this point, noting that "an accent placed over any note does not intimate that the note is to be *distorted*, or put out of its natural place, but rather it shows what is its most natural place, in that occurrence."

48 This chord, Holden claims, is ideally suited to be the final chord of a piece, as it does not generate expectations in its listeners for resolution or further motion. See part 1, §143, 127, and part 2, §76, 354. He further notes that in earlier times it frequently served as the final chord of music in a minor key by means of the raised Picardy third. See part 1, §269, 243, and part 2, §88, 372.

49 According to James Tenney (1988, 65–85), the view that consonance is a result of triadic membership was an innovation of Rameau's that had been widely adopted by the eighteenth century.

when a lack of shared triadic membership causes a combination of notes to "distract our attention," it is plausible to infer that, conversely, in concord "the two sounds will seem to unite" because, as members of a single triad, they are perceived as more unified. Unlike *consonance* and *dissonance*, therefore, which describe the impressions such sonorities make on the ear in terms of qualities such as unity, distinctness, agreeability, and their opposites, Holden's definitions of *concord* and *discord* speak to the cause of these impressions, namely, triadic membership, or lack thereof. Here consonance is implicitly reconfigured as a consequence of concord, as dissonance is of discord.[50] Moreover, Holden emphasizes, it is important to distinguish between "what only *divides our attention* and what *displeases*. The former of these is properly a *discord* in music, and the latter a *false, or inharmonical relation*" (part 1, §153, 139).[51]

Holden thus appears to consider consonance and dissonance as characterizing the opposing ends of a continuum between unity and distinctness. This is also evident in his unique appeal to acoustics, specifically to the overtone series.[52] Rather than taking as paradigmatic the familiar example of the vibrating string, Holden invokes the experience of listening to bells:

> Although the harmonics, belonging to the perfect chord of any sound, be generally, if not universally, predominant among its *natural consonances* when examined in this manner [i.e., the ringing of a bell], yet there are not wanting other sounds to coincide with, and assist our imagination, in case we have occasion to regard the total sound, not as a fundamental, but as one of the harmonics of some other fundamental. . . . Whatever particular tone we have previously estimated, and set our attention upon, we are almost sure to find something like it among the consonances of a bell. (part 1, §137, 121–22)

Holden's example of bells allows him to remark on the alignment of the perfect chord with real-world acoustic phenomena such as the overtone series

50 Thus the triad for Holden is represented in our minds as a sort of categorial construct that determines our perception of intervals as concordant or discordant on the basis of membership or nonmembership. This reverses the traditional understanding—still evident in the writings of Holden's contemporaries, such as Kirnberger (1771, 23–26)—in which consonance is a primitive property of certain dyads and the triad, as the class of sonorities composed of such dyads, is consonant owing to the consonance of its components. I thank David E. Cohen for this idea.

51 Dissonances therefore differ from inharmonical relations, which characterize pitches that do not stand in sufficiently close approximation to simple integer ratio proportions and hence sound mistuned. In some cases, however, inharmonical relations *are* permitted to stand in relation to a module by analogy. For example, Holden notes that music performed on an organ, whose pitches cannot be adjusted, is perceived much as it is when performed by violins (part 2, §94, 379). Furthermore, in the case of brass instruments such as the natural horn, "the interval between the II and IV is exactly 9 to 11: and yet this interval in compositions for a first and second horn, is very frequently used where the ear must inevitably conceive the sounds by the proportion of 6 to 7" (part 2, §94, 380). That is, although the second and fourth scale degrees function as the top two members of the dominant seventh chord built on the V of the scale, when they are performed by a natural horn whose fundamental is tuned to the tonic, Holden argues that we hear partials 9 and 11 of the tonic, yet understand them as if they were the partials 6 and 7 of the dominant.

52 See part 1, §137, 122 and 126; and part 2, §7, 287, and §79, 357. Holden limits the ideal voicing of the perfect chord to partials 1–5, because "6, being the octave of 3, viz. the fifth, cannot be admitted without rendering the same fifth too remarkable, and thereby taking off its dependance on the key," whereas "7, being a difficult factor in the parcelling of vibrations, produces a sound difficult to be sung, and unsatisfactory, except when it can be properly followed by another more easy sound, in the way of *resolution*" (part 2, §78, 356–57).

while positioning his theory as cognitivist: our imagination and attention can ultimately extract what they wish from any bell sound, regardless of its actual acoustic presence. This account supports the main claim of his theory, namely, that our minds can construct a potential interpretation of nearly any given pitch as a harmonic to the module of some keynote (and thus a different fundamental), which is retained in the memory. The choice of bells thereby expresses Holden's agenda of subordinating acoustics to cognition while retaining the privileged status of the relationship between the perfect chord and the overtone series.

Holden therefore regards consonance as a continuum bounded on one end by the five-note voicing of the major triad that exactly mimics the harmonic series of the first five partials, in the proportion 1:2:3:4:5, which constitutes the most unified chord possible (part 2, §76, 354). Any alterations in voicing, as well as doublings or omissions, detract from the perfect chord's perceptual unity, as "the further we depart from the disposition of the most perfect of this or any other kind of chord, the greater imperfection is thereby introduced" (part 1, §135, 120). When we hear a triad voiced differently or with omitted sounds, he asserts, our minds nevertheless understand it as representing the proportions 1:2:3:4:5 in varying degrees of attenuation. Moreover, "not only one or two of the notes of this chord may be wanting, but they may all be wanting except one, and yet still the same idea conceived by the hearer, as if they were all joined together" (part 1, §136, 120–21).[53] Therefore, Holden argues, the most perfect configuration of the minor chord will omit its third entirely, corresponding to 1:2:3:4 (part 1, §140, 124–25).

Of course, in a real piece of music, many of the sounds that we hear will not be perfect chords in root position. Holden therefore makes some general recommendations for the sake of enhancing the unity of a sonority: the fundamental should be doubled where possible, as "the unity of the whole is destroyed, when the third or fifth attracts too much of our attention" (part 1, §142, 125). Chord voicing is also important because "our attention is carried either to the lowest or the highest sound which exists in a chord, rather than any of those which lie, as it were, concealed in the middle" (part 1, §144, 127). In the case of chord inversions, he proposes, it is best to place an octave of the fundamental in the highest position, "that the hearer, being disappointed of it below, may yet find it above" (part 1, §144, 128).

Along similar lines, Holden's approach to dissonance develops from the assumption that such chords are perceptually less integrated. To explore this, he takes up an idea proposed by the Swiss music theorist Jean-Adam Serre some two decades earlier in *Essais sur les principes de l'harmonie* (1753), namely,

53 Holden here and elsewhere posits something like Riemann's (1882, 184–85) concept of *Klangvertretung*, in which a single note can represent a major or minor triad. For example, Holden writes, "It is necessary to conceive, not only the compound sounds of harmony, but also every single sound in melody, as belonging to some perfect chord, in order to account for their various effects, and the preference due to some successions of sounds rather than others" (part 1, §146, 131–32). I thank David E. Cohen for this observation.

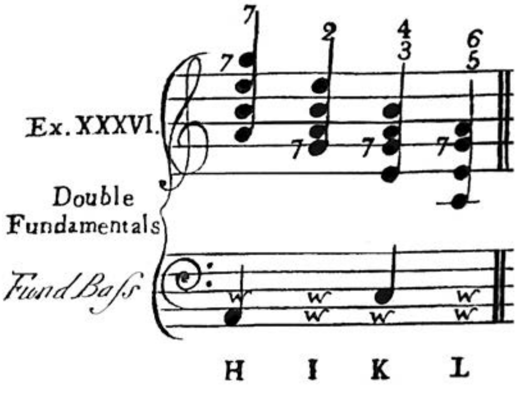


Figure 12. Seventh chord with double fundamental.

that dissonant chords have two fundamentals.[54] Holden asserts: "No one sound can properly be called the *sole* fundamental of a dissonant chord, because of the divided attention which the discord creates, but notwithstanding we must be supposed to have two fundamentals partly in view, yet one of them may, for various reasons, claim the greater share of our regard; and therefore may be properly called *governing*" (part 1, §160, 144). The fundamental of any well-formed chord is thus determined by its possession of a perfect fifth, which reinforces our perception of that fifth's lower note as the fundamental.[55] Therefore, a chord containing several perfect fifths, such as a minor seventh chord, will have competing fundamentals.[56]

Holden observes: "The chord of the seventh, when the fundamental bears a less third, may be considered as a mixture of two perfect chords whose fundamentals are the two terms of the same less third" (part 1, §178, 156). That is, we can conceive of this either as a minor triad with an added seventh or as a major triad with an added sixth whose root is the first chord's minor third. To ascertain the governing fundamental, he proposes we rely on the voicing of the chord itself. Figure 12 reproduces Holden's interpretation of a minor seventh chord of A–C–E–G as two overlapping perfect triads A–C–E and C–E–G, which thus contains a double fundamental.

54 According to Thomas Christensen (1993, 167), Serre's theory remained speculative, as it lacked "criteria by which one may prioritize the various roots."

55 Holden writes, "Every note which has its own harmonics, and especially its fifth existing along with it, will in some degree attract our attention as a fundamental" (part 1, §176, 155–56).

56 Two other intervals also possess the power to invoke fundamentals, namely, the diminished fifth, which has "a very peculiar property of referring the hearer to a fundamental note, at the distance of a greater third below its lower term," whether or not that note appears in the chord, and the tritone, which "refers to a fundamental, at the distance of a less third below its lower term" (especially when the perfect fifth above that fundamental is present), or it refers to the major third below the tritone's upper note (part 1, §179–80, 158).

In providing guidelines for determining the fundamental of a dissonant chord, Holden's analysis in Figure 12 relies on his earlier claim that our attention tends to be drawn to the outer voices of a chord, that is, the bass and soprano (i.e., melodic) lines (part 1, §144, 127).[57] In root position (letter *H*), the A commands our attention through its presence in the bass, while the C is concealed within the inner voices; this is indicated by the custodes in place of the C in the bass clef. A similar situation occurs in the second inversion (letter *K*), where the C commands our attention, as it is the highest part of the chord, and the custodes indicates the conflicting fundamental on the A in the bass clef. In the case of first and third inversion, however, the attention is more or less equally divided between potential fundamentals, as shown by the custodes on A and C. Here, Holden observes, the attention must be assisted through doubling or rendering the intended fundamental note conspicuous by other means:

> At L, the note A is uppermost, and C lowest; which causes a divided attention, or *double fundamental*, and the preference due to one or the other is to be determined from other considerations. The case is much the same at I, where both C and A are in the middle. In these ambiguous cases, if one of the notes which thus stand in competition be doubled, or be made more remarkable by falling in with the expectation of the hearer, the choice will be determined by either of these circumstances. (part 1, §178, 157)

By doubling the notes, as Holden here suggests, we fortify one of the perfect fifths, thus both strengthening its lower term's claim as fundamental and helping us identify the chord as either a minor seventh chord or an added sixth chord. But he also proposes that the harmonic and melodic context in which the chord occurs on a given occasion influences our determination of its fundamental. Much as in Rameau's *double emploi*, we can also determine the chord's fundamental according to its fulfillment of melodic or harmonic expectation.[58]

In the second, speculative part of the treatise, Holden formalizes these perceptual claims by introducing a new principle to explain various aspects of intervals, chords, and inversions, which he terms "implied sounds" (part 2, §76, 354). As I show below, this principle, an abstraction of the phenomenon of difference tones, allows him to provide a cognitive, rather than acoustic,

57 While Holden's belief that we can better perceive a chord's fundamental if it is placed in the soprano line seems counterintuitive, it is likely that he arrived at this conclusion from observing that bass and soprano parts tend to attract most of our attention; see part 1, §199, 176, and §144, 128, respectively.

58 The criterion of expectation also applies to perfect chords with substituted sounds, or suspensions. To ascertain the fundamental of suspensions, Holden suggests that we search for notes that possess their own harmonics in the chord, particularly their fifth, as this "will in some degree attract our attention, [particularly if] such a note be rendered more distinguishable by its situation above or below the rest" (part 1, §176, 155–56). However, we must also factor in the syntactical context of the chords and "take into account the expectation, whether founded on nature or custom, of hearing such particular fundamentals in succession," as this may sometimes cause us to favor a different fundamental (part 1, §176, 155–56).

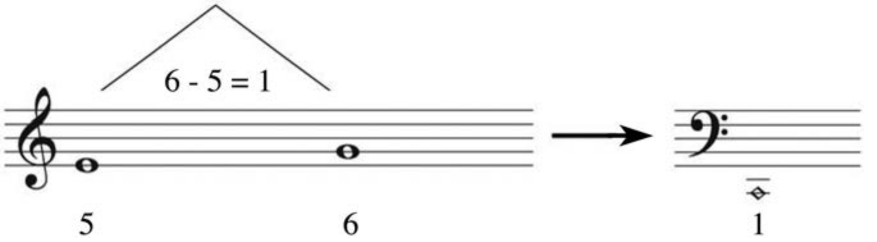


Figure 13. Implied difference tone generated by a minor third.

account for the perceptual unity of consonance. Holden asserts that, on hearing two different sounds, our imagination projects an implied sound, that is, a third sound equal to their difference. Say, for example, that we are in the key of C major, and hence measuring all sounds by a corresponding module of C (designated as f). If we hear G and E with the frequency $6f$ and $5f$, Holden predicts that our mind recognizes the difference of $6f - 5f = 1f$ and internally generates an implied sound of $1f$. Therefore, although the C is not present in the actual sonority (E, G), our imagination supplies the sound, as shown in Figure 13.

In the eighteenth century it was unclear whether the origins of difference tones were acoustical sounds or a psychophysical illusion.[59] Holden sidesteps this debate altogether by arguing that the implied tones arise from the process of perception. He makes another analogy to visual experience:

> It may be inferred, that as the velocity with which one moving body approaches towards or recedes from another (which may be called its relative motion, and is equal to the difference of the two absolute motions) is a circumstance which always attracts part of our regard: so the difference of the vibrations, or the relative velocity of the pulses of one sound in comparison to those of the other, may some way be perceived, abstract from all considerations either of real or *sensible* coincident pulses. (part 2, §75, 353–54)

Holden applies his principle of implied sounds to explain a range of musical features. For example, he ascribes the power of "the most complete and natural arrangement" of the perfect chord (1:2:3:4:5) to the fact that the difference between any two pitches in this chord will always be equal to one of the members of the chord itself (part 2, §76, 354). He remarks, "The implied sounds, in this perfect chord, produce no other effect than that of fortifying, or doubling, all the real sounds of the chord, except the highest" (part 2, §76, 355). As shown in Figure 14, the root of the chord is thus always most strongly reinforced, as the difference of 2-1 appears four times in the sonority, while the fifth is reinforced only a single time, and the third is not fortified by any implied tones.

59 Difference, or Tartini, tones are now explained as being physical but intra-aural, arising from the nonlinearity of the cochlear membrane (Langner 2015, 32).

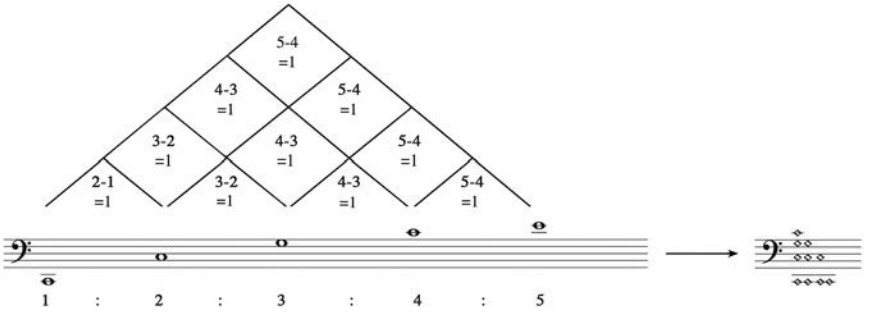


Figure 14. Finding the implied sounds (shown in diamond note heads) of a perfect chord.

Holden maintains that implied sounds influence musical practice in a number of ways. For one, they support his characterization of the octave as the most perfect consonance, followed by the double octave.[60] Therefore the familiar convention of enhancing an accompaniment with octave doublings fortifies the harmony by "essentially [implying] the coexistence of all its octaves, both above and below, not exceeding the limits of audible sound" (part 2, §77, 356). Implied sounds can also explain why we tend to prefer certain chord voicings. For example, the close-position sixth chord, or 5:6:8, comprises the implied sounds 6 – 5 = 1, 8 – 6 = 2, and 8 – 5 = 3, to which we project the implied sonority of C2, G2, and C3, all of which support our perception of the chord's root as C. However, if we change the voicing by lowering the E4 by an octave, the resulting implied sounds are 12 – 5 = 7, 16 – 12 = 4, and 16 – 5 = 11. Holden notes that even though these sounds stand in complex relations to the root, and are therefore barely perceivable, "still they contribute something toward the imperfection of the chord" (part 2, §81, 360). Finally, implied sounds can help us understand why intervals tend to sound less consonant in very low registers: if a chord's implied sounds exceed our hearing range, it is perceived as less consonant, as shown in Figure 15 (part 2, §91, 376–77).

To summarize, Holden understands consonance and dissonance as qualities occupying a continuous spectrum defined by greater and lesser degrees of perceptual unity and distinctness.[61] The degree to which a sonority will be classified as one or the other depends on two factors: (1) the conformance of

60 Holden writes, "When the two terms of the single octave are sounded, completely true, the upper term unites so perfectly with the lower, that the ear is sensible of nothing more than the sound of the lower term made louder and *fuller*, by the concurrence of the upper term. The case is different with the double octave. . . . Here the upper term will be heard distinctly, as well as the lower, and the ear is abundantly sensible of two sounds, but, at the same time, they are so perfectly like to each other, that we can attend to both without any difficulty or distraction" (part 1, §126, 113–14).

61 Holden's approach here resonates with Euler's project in the *Tentamen* of classifying intervals according to their "degrees of agreeability" (*gradus suavitatis*). See Grant 2013.

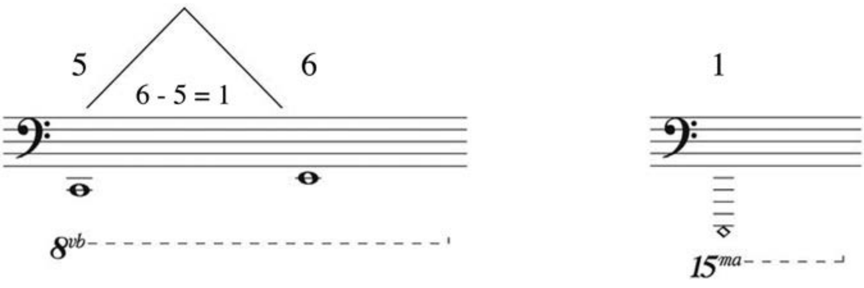


Figure 15. Inaudible implied tones do not contribute to the consonance of dyads at the lower boundary of our hearing range.

the specific voicing in question to the first five partials of the overtone series as embodied by the fundamental perfect chord 1:2:3:4:5, which also exhibits the most desirable configuration of implied tones; and (2) the degree to which implied tones are cognitively perceivable (thus interacting with physiological limits of our hearing at registral boundaries). Regardless of their ambiguous status as acoustic or perceptual objects, the primary function of implied tones is the heuristic one of supplying the listener with additional information about the quality (and potentially the fundamental) of a chord. Thus their greatest value lies in providing a coherent explanation for how differences in chord voicing and doubling can either reinforce or conflict with the function of a chord, not only with respect to a given module but also in a broader tonal context.

Conclusion

As a cognitive theory *avant la lettre*, the *Essay* challenges our assumptions about the understandings of musical audition available in the eighteenth century. Holden explicitly states what he regards as the conditions and the fundamental mechanism of musical perception and then derives his entire theory as an orderly series of deductions and consequences therefrom. By approaching perception as composed of hierarchies of mental processes, Holden's methodology enables him to articulate a series of prescient intuitions essential to the modern fields of music theory and music cognition.

One remarkable aspect of Holden's theory lies in the fact that it is modeled on the phenomenon of rhythm rather than pitch. The *Essay* thus presents an early articulation of the notion that our perception of a range of musical parameters alike occurs over a temporal continuum, an idea that was naively developed by the eighteenth-century Scottish music theorist Thomas Robertson (1784, 229–33), who notated the "rhythms" of different pitch intervals a few years later. Related approaches would independently resurface over the nineteenth and twentieth centuries, most notably in the writings of Hector Berlioz and Moritz Hauptmann,[62] as well as in the work of Friedrich Opelt, Henry Cowell, and Karlheinz Stockhausen, as Alexander Rehding (2016, 273–75) has shown. More recently, the music psychologist Guy Madison (2014, 15) has proposed that our "perception of rhythm is built on similar principles as the perception of pitch, since both seem to utilise small integer subdivision of intervals. The metrical structure is in this sense analogous to the partials (overtones) that characterise complex auditory tones."[63]

Beyond music-theoretical analogies between rhythm and pitch, however, Holden's treatise invites us to reconsider the history of a number of key aspects of music psychology and cognition. The music psychologist Henkjan Honing (2011, 23) defines musical cognition as entailing "discussions of the role of memory, attention, perception and expectation in listening to music." As we saw, Holden repeatedly deploys similar aspects of our mental capacities in order to link musical pleasure to anticipation,[64] as well as to explain cross-modal cognitive limits on memory,[65] perception (just noticeable difference),[66] and mental processing,[67] in addition to offering what appears to be the earliest account of the phenomenon of auditory grouping known as subjective rhythmicization.[68] While some of these findings were reported by other Enlightenment writers, the fact that the *Essay* independently describes *all* of these phenomena is exceptional in the context of eighteenth-century musical culture and attests to the essentially cognitivist nature of Holden's project.

Holden's *Essay* can be productively compared to a number of subsequent psychological approaches. Burdette Green and David Butler (2002, 246) distinguish between "outside to inside" theories of music cognition, exemplified by Helmholtz's development of psychophysics, and "inside to outside" mentalist theories of music, which they link to Carl Stumpf and the Gestalt psychologists. Holden's theory clearly falls along the lines of the latter and indeed has

62 Consider Berlioz's (1837, 1) claim that "there are rhythmic dissonances, there are rhythmic consonances, there are rhythmic modulations; nothing could be more obvious" (Il y a des dissonnances rhythmiques, il y a des consonnances rhythmiques, il y a des modulations rhythmiques; rien de plus evident), or Hauptmann's (1853) dialectical approach to both harmony and meter.

63 I thank Richard Cohn for bringing this passage to my attention.

64 Holden (1766, 4) observes, "Our enjoyment of music depends, in a great measure, on a faculty of retaining the ideas of former sounds, and anticipating those which are to follow." Compare, for example, David Huron's (2006, 138) claim that "if a listener predicted the occurrence of G4, then the tone itself is likely to be experienced as pleasant. . . . It is not frequency of occurrence per se that accounts for the experience of pleasure, but sure and accurate prediction."

65 Compare part 2, §14, 290–91, with Bob Snyder's (2000, 56) ascription of rhythmic and melodic grouping to short-term memory: "Sequences of events at the melodic and rhythmic grouping level that fall within the limitations of short-term memory [between 1/16 of a second and 8 seconds long] are perceived as being in the present and as forming various kinds of groupings and phrases that can be apprehended in their entirety."

66 See part 2, §42, 315, discussed at n. 35.

67 See Holden's discussion of the limits of our hearing range in part 2, §25, 302, discussed in n. 18. Today we know that rate limits on pitch are predominantly determined by low-level transduction mechanisms (Cheveigné 2010, 88). However, the notion that rate limits on processing determine important aspects of cognition has been widely accepted in modern psychology (Salthouse 1996).

68 Compare part 2, §14, 294, discussed in section 1, with Kirnberger 1776, 114–15, and Bolton 1894, 185.

significant similarities with Gestalt psychology. A number of principles stated in the *Essay* are almost identical to preference rules articulated by the Gestaltists. For example, Holden's formulation of our mental defaults in grouping is highly reminiscent of the principle of *Prägnanz* (conciseness), according to which "stimuli are perceived and remembered as coherent wholes (often as more coherent than they are), and are organized in perception in as stable, symmetrical, simple, and meaningful a way as possible" (Sutherland 1991, 232).

Another way in which Holden's approach resonates with modern psychology lies in his use of musical syntax as evidence for the structure of our cognition.[69] In this regard, his theory seems unexpectedly close to the ideas of the so-called cognitive revolution of the 1950s associated with George Miller and Noam Chomsky. Comparable to Fred Lerdahl and Ray Jackendoff's *Generative Theory of Tonal Music* (1983), and other theories derived from Chomsky's approach, Holden assumes that when we hear a sound or a string of sounds, our mind generates a parallel mental representation in the form of a path through divisions of the module. Holden thereby breaks completely with explanations grounded in acoustics, numerology, or experiment to conceptualize perception as an active process in which the mind extracts features of the musical input and organizes them in accordance with innate psychological preferences.

Yet in twenty-first-century hindsight, perhaps the most remarkable feature of Holden's model is its reliance on what is termed in modern psychology "hierarchical processing," specifically a mixture of "bottom-up" and "top-down processing." Bottom-up processing begins with sensory input, which is then "transformed and combined until we have formed a perception. The information is transmitted upwards from the bottom level (the sensory input) to higher, more cognitive levels" (Rookes and Willson 2000, 13). In contrast, top-down processing argues that we require "our stored knowledge about the world in order to make sense of [sensory] input" (13). Holden's approach repeatedly invokes both bottom-up and top-down processing at various stages of perception.

Figure 16 maps Holden's theory onto the tripartite hierarchy of mental processing widely accepted in modern psychology (see Adelson 1999, 3–9). The low level comprises the level of the senses, and in the case of hearing, this entails the physiological transductions of the acoustic signal by the ear. As an example of low-level processing, for Holden it is the ear, not the mind, that evaluates the isochrony of successive pulses or vibrations (as discussed in section 1).

69 Holden (1766, 4) himself uses the phrase *syntax of music* in a description of his theory: "To shew how the several chords which are admitted in harmony, are connected one with another, in every strain which gives us pleasure, may not improperly be called the *Syntax of music*; and is a subject of which no competent idea can be conveyed within the limits of a preface. . . . [The author] hopes shortly to finish, a familiar treatise on this subject. . . . The latter part contains a short sketch of the theory of music, upon principles in a great measure new, which seem to be more simple and satisfactory, than any which have formerly been pursued."

With a few exceptions, the second, speculative part of Holden's *Essay* explores phenomena that we today ascribe to mid-level processing. According to the neuroscientist Daniel Pressnitzer (2009, i), mid-level audition pertains to "processes that sit between an acoustical description of sound and the use of auditory information to guide behavior," in which a series of operations extract and bind complex features, thereby distilling complex auditory stimuli into more hierarchically organized mental representations. As an instance of mid-level processing, consider Holden's claim that the fundamental bass arises as a consequence of the mind's acts of grouping, discussed in section 3. This example demonstrates how the signals conveyed by the ear are refined into more complex mental representations.

In contrast, the first, practical part of the *Essay* generally describes processes that we would today categorize as higher level, the stage at which auditory objects are perceived as coherent wholes participating in syntactical contexts. As an example of high-level processing, consider Holden's account of the experience of hearing a minor seventh chord with ambiguous fundamentals, which divide our attention, discussed in section 5. These processes include the role of attention in selectively modulating our interpretation of sounds perceived, as well as the influence of stylistic conventions such as function, voicing, and register.

As Figure 16 illustrates, at the lowest level, the ear judges whether a sensed sound is regular. If not, it is discarded as noise; otherwise the ear sends the sound up to the mid-level. At the mid-level, the mind performs a series of analyses: it decomposes the sound into small prime factors and compares them to previous divisions of the module, attempting to determine how the sound fits into the key. If the relationship is straightforward, the imagination projects a corresponding fundamental; if not, the sound is excluded as mistuned. If the sound is composed of more than one pitch, the mind ascertains its implied sounds and amends its fundamental as necessary. At successively higher levels, considerations of voicing and register, and subsequently function and chunking, affect the organizing of extracted features into coherent wholes, modulated by concerns of syntax and the faculty of attention.

Figure 16 demonstrates how Holden incorporates both bottom-up and top-down processing in his theory. Bottom-up processing is evident in the successive refinement of the internal representation of sounds, starting from the ear's low-level estimate of pitch and continuing through the acts of feature extraction and binding associated with mid-level processing and the resultant representations that are then further interpreted by high-level processes (Snyder 2000, 4). However, Holden's model also takes top-down processes into account. In tandem with bottom-up processing at the mid-level (such as grouping via a comparison with the module and associated factoring, as well as the assigning of implied tones and projected fundamentals), an internal representation held in implicit memory—the previous module, factoring, and fundamental—directly affects the way in which these features are extracted and perceived at a given moment.[70] The influence of the immediate past thereby exerts a top-down influence on our interpretation of the current state.

At higher levels, these mid-level features are further organized by established musical conventions pertaining to register, voicing, and comparable qualities, in order to chunk individual elements into larger units such as phrases or functions. Higher-level processing can also call on the attention to determine the function of the sounds perceived, particularly around ambiguous situations, such as the case we encountered in section 5, where a chord performed by natural horns can be functionally heard in terms that are considerably different from its acoustic expression.[71] This, too, calls on top-down processing, as our interpretation relies on previously acquired syntactical knowledge to shape our experience of a current sound or event.

Perhaps the clearest example of Holden's reliance on both top-down and bottom-up processes is seen in his bifurcated approach to attention as both active and passive, discussed in sections 4 and 5. On the one hand, we can place our attention on a certain sound as a keynote, exerting a top-down influence on the maintenance of the module and associated key relationships. On the other, our attention may be attracted by competing fundamentals or implied tones and influenced from the bottom up. At the same time, considerations of syntax can affect both voluntary and involuntary forms of attention, as indicated with bidirectional arrows in Figure 16. Musical features that activate and divide our attention, located at the very top of the diagram, sometimes breach the focus of our conscious awareness, either by causing the sensation of distraction or by calling on our minds to actively intervene and interpret incoming stimuli (such as a modulation, a dissonance, or other rare or significant sonic objects).[72]

Holden's treatise reveals a number of unexpected ways in which a rigorous engagement with music theory can give rise to sophisticated insights into cognition. Viewed within the context of eighteenth-century musical culture, the *Essay* is exceptional by any measure. However, much of the astonishing detail of Holden's theory arises from the fact that he was attempting to understand perception through the music theory of his day. He deduces the workings of the mental faculties of attention, memory, and expectation via introspection by closely observing the behavioral and syntactical characteristics of a cultural practice. In a sense, his theory offers a cognitive interpretation of a familiar phenomenon that was already thoroughly characterized in terms of hierarchical structures. Holden's greatest innovation in this respect lies in theorizing, in detail, our preference for simple integer-ratio rhythmic and

70 See Holden's derivation of the descending scale, and in particular his account of the decision to provide scale degree IV with a fundamental of V rather than II, discussed in section 3.

71 See n. 51.

72 My interpretation of Holden's account of divided attention as indicating a state of near or full awareness supports Natalie Phillips's (2016, 4) theorization of late eighteenth-century understandings of distraction as a productive cognitive state.

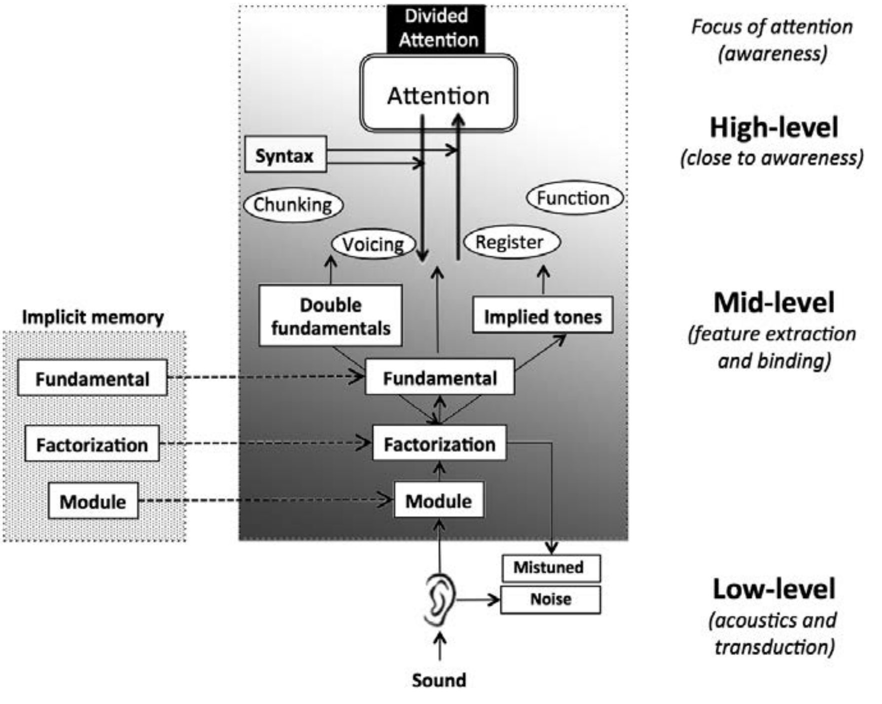


Figure 16. **A flowchart of Holden's theory, mapped onto the modern tripartite division of the cognitive faculties involved in audition, shown at the right. The gray shading illustrates the gradual boundaries between low-, mid-, and high-level processing.**

pitch relations as emerging from the innate structure of our minds. This allows him consequently to regard complex configurations of musical objects—starting from the scale and ascending through musical syntax—as reflecting mental rather than acoustic or numerological constraints. More than any other factor, this insight allowed him to theorize the experience of the music of his time as arising from a set of universal cognitive principles.

Works Cited

Adelson, Edward H. 1999. "Layered Representations for Vision and Video." In *Proceedings of the IEEE Workshop on Representation of Visual Scenes*, 3–9. Cambridge, MA: IEEE.

Berlioz, Hector. 1837. "Strauss: Son orchestre, ses walses.—De l'avenir du rhythme." *Journal des débates*, November 10, 1.

Bolton, Thaddeus L. 1894. "Rhythm." *American Journal of Psychology* 6/2: 145–238.

Briseux, Charles-Étienne. 1853. *Traité du beau essentiel dans les arts*. Paris.

Brown, Leslie E. 2001. "The Common Sense School and the Science of Music in Eighteenth-Century Scotland." In *Essays in Honor of John F. Ohl: A Compendium of American Musicology*, edited by Enrique Alberto Arias, 122–32. Evanston, IL: Northwestern University Press.

Chenette, Louis F. 1967. "Music Theory in the British Isles during the Enlightenment." Ph.D. diss., Ohio State University.

Cheveigné, Alain de. 2010. "Pitch Perception." In *The Oxford Handbook of Auditory Science: Hearing*, edited by Christopher Plack, 71–104. Oxford: Oxford University Press.

Christensen, Thomas. 1993. *Rameau and Musical Thought in the Enlightenment*. Cambridge: Cambridge University Press.

Cohen, David E. 2001. "The 'Gift of Nature': Musical 'Instinct' and Musical Cognition in Rameau." In *Music Theory and Natural Order from the Renaissance to the Early Twentieth Century*, edited by Suzannah Clark and Alexander Rehding, 68–92. Cambridge: Cambridge University Press.

Damschroder, David. 2008. *Thinking about Harmony: Historical Perspectives on Analysis*. Cambridge: Cambridge University Press.

Euler, Leonhard. 1739. *Tentamen novae theoriae musicae ex certissimus harmoniae principiis dilucide expositae*. St. Petersburg. Facs. ed., New York: Broude, 1968.

Euler, Leonhard. 1766. "Conjecture sur la raison de quelques dissonances généralement reçues dans la musique." In *Histoire de l'Académie royale des sciences et des belles-lettres de Berlin année 1764*, 165–73. Berlin.

Fétis, François-Joseph. 1839. *Biographie universelle des musiciens*. Vol. 5. Brussels.

Forkel, Johann Nikolaus. 1792. *Allgemeine Litteratur der Musik, oder Anleitung zur Kenntniss musikalischer Bücher*. Leipzig.

Goldberg, Daniel L. 2017. "Bulgarian Meter in Performance." Ph.D. diss., Yale University.

Grant, Roger M. 2013. "Leonhard Euler's Unfinished Theory of Rhythm." *Journal of Music Theory* 57/2: 245–86.

Grant, Roger M. 2014. *Beating Time and Measuring Music in the Early Modern Era*. Oxford: Oxford University Press.

Green, Burdette, and David Butler. 2002. "From Acoustics to *Tonpsychologie*." In *The Cambridge History of Western Music Theory*, edited by Thomas Christensen, 246–71. Cambridge: Cambridge University Press.

Hagner, Michael. 2003. "Toward a History of Attention in Culture and Science." *MLN* 118/3: 670–87.

Hauptmann, Moritz. 1853. *Die Natur der Harmonik und Metrik*. Leipzig.

Holden, John. 1766. *A Collection of Church-Music Consisting of New Setts of the Common Psalm-Tunes . . . for the Use of the University of Glasgow*. Glasgow.

Holden, John. 1807. *An Essay towards a Rational System of Music*. 3rd ed. Edinburgh.

Honing, Henkjan. 2011. *Musical Cognition: A Science of Listening*, translated by Sherry Marx and Susan van der Werff-Woolhouse. New Brunswick, NJ: Transaction.

Houle, George. 1987. *Meter in Music, 1600–1800: Performance, Perception, and Notation*. Bloomington: Indiana University Press.

Huron, David B. 2006. *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

Kassler, Jamie C. 1971. "British Writings on Music, 1760–1830: A Systematic Essay toward a Philosophy of Selected Theoretical Writings." Ph.D. diss., Columbia University.

Kassler, Jamie C. 1979. *The Science of Music in Britain, 1714–1830: A Catalogue of Writings, Lectures, and Inventions*. Vol. 1. New York: Garland.

Kirnberger, Johann Philipp. 1771. *Die Kunst des reinen Satzes in der Musik*. Vol. 1. Berlin.

Kirnberger, Johann Philipp. 1776. *Die Kunst des reinen Satzes in der Musik*. Vol. 2. Berlin.

Langner, Gerald D. 2015. *The Neural Code of Pitch and Harmony*. Cambridge: Cambridge University Press.

Leibniz, Gottfried Wilhelm. 1988. "Letter no. 6 to Christian Goldbach (1712)." In "La correspondance de Leibniz avec Goldbach," edited by Adolph-Andrei P. Juschkewitsch and Juri C. Kopelewitsch. *Studia Leibnitiana* 20: 175–89.

London, Justin. 2004. *Hearing in Time: Psychological Aspects of Musical Meter.* Oxford: Oxford University Press.

Madison, Guy. 2014. "Sensori-motor Synchronisation Variability Decreases as the Number of Metrical Levels in the Stimulus Signal Increases." *Acta Psychologica* 147: 10–16.

Miller, George A. 1956. "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information." *Psychological Review* 63/2: 81–97.

Moreno, Jairo. 2004. *Musical Representations, Subjects, and Objects: The Construction of Musical Thought in Zarlino, Descartes, Rameau, and Weber.* Bloomington: Indiana University Press.

Phillips, Natalie M. 2016. *Distraction: Problems of Attention in Eighteenth-Century Literature.* Baltimore, MD: Johns Hopkins University Press.

Pressnitzer, Daniel. 2009. "Mid-Level Audition." Habilitation, Université Paris Descartes et Département d'Études Cognitives, École Normale Supérieure.

Rameau, Jean-Philippe. 1737. *Génération harmonique.* Paris.

Rehding, Alexander. 2016. "Three Music-Theory Lessons." *Journal of the Royal Musical Association* 141/2: 251–82.

Reid, Thomas. 2002. *The Correspondence of Thomas Reid*, edited by Mark Wood. Vol. 4. University Park: Pennsylvania State University Press.

Riemann, Hugo. 1882. "Die Natur der Harmonik." In vol. 4 of *Sammlung musikalischer Vorträge*, edited by Paul Graf Waldersee, 157–90. Leipzig.

Riemann, Hugo. 1992. "Ideen zu einer 'Lehre von den Tonvorstellungen,'" translated by Robert W. Wason and Elizabeth West Marvin. *Journal of Music Theory* 36/1: 69–79.

Riley, Matthew. 2004. *Musical Listening in the German Enlightenment: Attention, Wonder, and Astonishment.* Burlington, VT: Ashgate.

Robertson, Thomas. 1784. *An Inquiry into the Fine Arts.* London.

Rookes, Paul, and Jane Willson. 2000. *Perception: Theory, Development, and Organisation.* London: Routledge.

Salthouse, Timothy A. 1996. "The Processing-Speed Theory of Adult Age Differences in Cognition." *Psychological Review* 103/3: 403–28.

Serre, Jean-Adam. 1753. *Essais sur les principes de l'harmonie.* Paris.

Snyder, Bob. 2000. *Music and Memory: An Introduction.* Cambridge, MA: MIT Press.

Stevens, Isaac. 1863. "An Essay on the Theory of Music." Ph.D. diss., Georg-August-Universität Göttingen.

Sulzer, Johann Georg. 1793. *Allgemeine Theorie der Schönen Künste.* 2nd ed. Vol. 3. Leipzig.

Sutherland, Stuart. 1991. "Law of Prägnanz." In *The Macmillan Dictionary of Psychology*, edited by Stuart Sutherland, 232. London: Macmillan.

Tenney, James. 1988. *A History of "Consonance" and "Dissonance."* New York: Excelsior.

Woodward, William R. 2015. *Hermann Lotze: An Intellectual Biography.* Cambridge: Cambridge University Press.

Carmel Raz is a research group leader at the Max Planck Institute for Empirical Aesthetics in Frankfurt, where she directs the group "Histories of Music, Mind, and Body." Prior to taking up the position in Frankfurt, she was a postdoctoral research fellow at the Columbia Society of Fellows.