

# Prospects for timbre physicalism

Alistair M. C. Isaac<sup>1</sup>

Published online: 13 February 2017

© The Author(s) 2017. This article is published with open access at Springerlink.com

**Abstract** *Timbre* is that property of a sound that distinguishes it other than pitch and loudness, for instance the distinctive sound quality of a violin or flute. While the term is obscure, the concept has played an important, implicit role in recent philosophy of sound. Philosophers have debated whether to identify sounds with properties of waves, events, or objects. Many of the intuitive considerations in this debate apply most clearly to timbre qualities. Two prominent forms of timbre physicalism have emerged: one identifying timbre with the spectral composition of proximal waves; the second identifying timbre with the mechanical vibrations at a sound source. I demonstrate that the first possibility is conceptually unsatisfying, while the second fails to meet the standards of rigor established by the color physicalism literature. One response to these worries might be to adopt a more modest, non-reductive realism about timbre, such as the ecological view of J. J. Gibson.

**Keywords** Timbre · Sound · Realism · Physicalism · Similarity · Constancy

## 1 Introduction

Do qualities as we experience them in perception exist independently of us, as objective features of the world? This question of perceptual realism has been pursued most extensively in the case of color, and positions developed in the philosophy of color have shaped and informed the burgeoning philosophy of other sense modalities. Color physicalism, the view that color may be identified with or

---

✉ Alistair M. C. Isaac  
tokyodrifter@gmail.com; a.m.c.isaac@ed.ac.uk

<sup>1</sup> School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building, 3 Charles Street, Edinburgh EH8 9AD, UK

reduced to a well-defined physical kind, such as surface spectral reflectance, has been especially influential—in part because it is a position that engages closely with the science of color and its external correlates. Here I examine the auditory quality of *timbre* through the lens of the science of sound and its correlates, asking whether it may be identified with or reduced to some objective physical feature(s) of the world in a manner analogous to that of color physicalism. Many arguments about the ontology of sound turn most directly on intuitions about timbre, and so a rigorous analysis of timbre physicalism should shed light on more general questions about what sounds are and where we find them in nature.

If the term “timbre” is familiar at all, it is typically from a musical context, where we use it to describe the characteristic sound quality of a musical instrument: a violin has a different timbre than a clarinet, for instance. More generally, timbre is just the quality of a sound that distinguishes it other than its pitch, loudness, or duration. This negative, or “dust bin” definition, identifying timbre as whatever is left over after subtracting out the best understood auditory qualities, is deeply dissatisfying, but it is a symptom of the fact that, unlike the case of color, we do not have a rich specialized vocabulary for describing our timbre experiences.<sup>1</sup> Nevertheless, we are all just as familiar with timbre qualities as with colors, and we do manage to refer to them, either onomatopoeically—*what is that strange buzzing?*—or by reference to a (potential) source—*it sounds like a giant bee!*

The science of timbre is not nearly as mature as the science of color. Nevertheless, it provides results analogous to those that have become important in debates about color realism on the key phenomena of correlation, constancy, and similarity. Color sensations are *correlated* with features in the world, for instance surface spectral reflectance; our attributions of color exhibit *constancy* with some of these correlates across changes in others, as when a sheet of paper looks white in both dim fluorescent and broad daylight; and colors stand in determinate *similarity relations*, e.g. orange is more similar to red than to blue. The correlates of color serve as candidates for a realist analysis, and the constancy with which we attribute colors to a correlate counts as evidence in favor of this analysis. Similarities are also evidence in the debate, and if the correlates of color do not stand in the appropriate similarity relations—if the surface reflectance for orange is not more similar to that of red than to that of blue—this fact must be explained, or explained away by the successful physicalist.<sup>2</sup> We find analogous aspects of the science of timbre in the debate about timbre physicalism: correlates of timbre are candidates for a

<sup>1</sup> Some date this negative definition to Helmholtz (1885, 19f); it is the one adopted by the Acoustical Society of America, and is discussed in classic texts such as Bregman (1990, 92f). Alternative definitions of timbre either appeal to the impoverished terminology of ordinary talk about sounds (that quality of a sound we describe as “bright” or “dull,” “harsh” or “mellow,”...), or directly state the Helmholtz theory articulated in Sect. 2.3.

<sup>2</sup> For instance, Hardin (1988) argues that, because no physical correlates of color exhibit the similarities that obtain between colors, we should be color eliminativists. In contrast, Churchland (2007) attempts to derive color similarities from surface spectral reflectance, arguing that success would constitute an “answer to [the] color realist’s prayer” (133).

physicalist reduction; constancy in timbre attribution counts as evidence in favor of some correlates over others; similarities between timbres need to be explained or explained away.

Despite the analogous structure in physicalist debates on color and timbre, I argue that the most popular candidate for a physicalist reduction of timbre, the component vibrations of distal resonant events, is disanalogous with surface spectral reflectance. In particular, a robust basis for the physicalist reduction of some set of qualities should satisfy two constraints: first, it should be exhaustive, reducing all the qualities in the set; second, it should form a natural category by the lights of our best science. Though strict, these constraints appear to be met for the reduction of color to surface spectral reflectance: the space of all surface reflectances is a scientifically well-defined natural category, and it provides a reduction basis for all possible surface colors. In contrast, I demonstrate that these two constraints cannot be simultaneously satisfied for a reduction of timbre to the component vibrations of a resonant event. The well-defined natural category of mechanical vibrations is not adequate to reduce all timbres, but an enlarged reduction basis is no longer well-defined or scientifically legitimate as a kind. An alternative candidate for timbre physicalism, the spectral composition of a sound wave, fares better as a basis for reduction, but fails to satisfy the demand that the similarities between timbres be adequately explained.

I begin with an introduction to the basic science of timbre, before surveying the major philosophical theories of timbre. The key question for the would-be timbre physicalist is: what is the bearer of timbre? It is the properties of this bearer that suggest themselves as candidates for identifying or reducing timbre. This leads to a detailed analysis of the prospects for reducing timbre to the resonant interactions that compose an audible event, the most prominent current candidate for the bearer of timbre. Either these interactions are limited to the well-defined class of mechanical vibrations, in which case they do not subvene all possible timbres, or they are grouped with other resonant features of the event that together might exhaustively subvene timbre, but no longer form a scientifically well-defined kind. The general moral of this discussion is that prospective timbre physicalists have yet to provide an adequate reduction basis for timbre. I conclude with some reflections on the prospects for a reductive approach to timbre, suggesting that a non-reductive, ecological realism may be more promising.

## 2 Introducing timbre

This section introduces the features of timbre science relevant to the timbre physicalism debate. After a discussion of the timbre stimulus, I examine one method for studying the experience of timbre, and how it might produce a “timbre solid” analogous to the color solid. Finally, I introduce the two main scientific theories of timbre, which identify different external correlates as the targets for the scientific study of timbre.

## 2.1 The timbre stimulus

Timbre is typically studied as a property of complex sound stimuli. The simplest sound stimulus is a sine wave, which can be described by three parameters: frequency, amplitude, and duration. Complex sound stimuli are just complex waves. Fourier's Theorem states that any complex wave homogeneous in time may be exhaustively described as a combination of sine waves at different relative frequencies, amplitudes, and phases; the pattern of simple frequencies in a complex wave is its spectral composition, and may be derived by means of a Fourier transformation. One of these sine waves, typically the lowest, provides the fundamental frequency, responsible for the pitch at which the sound is perceived. The other sine waves are the overtones—in a musical context these typically stand in whole number ratios to the fundamental frequency and are called harmonics. Finally, realistic stimuli will change in spectral composition over time, so a complete description requires not just a Fourier transformation of a time slice of the stimulus into component sine waves, but an assignment of a dynamic envelope, or description of changing relative amplitude as a function of time, to each wave.

The mathematics of Fourier analysis ensures that sound stimuli in any experimental context may be rigorously described. Even when the stimuli are derived from recordings of natural sources, for instance musical instruments or the human voice, spectral analysis allows a complete description and precise comparison of their properties. This mathematical analysis does not, however, deliver a prediction about how the stimulus will sound. Given two complex waves, will they sound similar or different? The degree of similarity in our experience of timbres may not correspond to any similarity between the stimuli that is easily definable in mathematical or physical terms. Just as surface reflectance profiles that are quite different may be perceived as identical in color, i.e. as *metamers*, it is *prima facie* possible that waves that are physically quite different may be perceived as similar in timbre.

## 2.2 Timbre experience

Psychophysics provides methods for empirically determining the perceived similarity relations between stimuli. There are many experimental methods for obtaining data from which a measure of such similarities may be derived: the subject might be asked if they can discriminate between two stimuli, to assign a number to the degree of perceived similarity between stimuli, or to order a set of stimuli by relative similarity or difference. This data may then be arranged into a *quality space*, which plots possible sensations as points, and degrees of (dis)similarity between these as distances within the space. In the case of color, the quality space derived from studies such as these is familiar as the color solid, which represents in three dimensions the relative degrees of similarity between colors. Can the same techniques be used to derive a timbre solid?

One reason to expect that they should is that we clearly make consistent judgments of timbre (dis)similarity. In fact, timbre typically trumps all other attributes when it comes to categorizing sounds—two knocks against wood, two

buzzes from different insects, two notes played on a violin: in each case, the pair is judged similar despite differences in pitch, loudness, or duration.<sup>3</sup> Nevertheless, there is a major stumbling block in the search for a timbre solid, namely we do not have any antecedent folk theory, or specialized vocabulary, for describing timbres. Correspondingly, we have no pre-scientific theory of the basic features that organize our experience of timbre, i.e. we do not know the grounds on which timbre comparisons are made. This has motivated the use of multidimensional scaling in the study of timbre similarities. Multidimensional scaling is an algorithmic procedure for extracting a low-dimensional model from high-dimensional data. A set of similarity judgments between  $n$  stimuli may always be represented (up to monotonicity) in an  $(n - 1)$ -dimensional space, as this means that each point may be arranged at any distance whatsoever from the remaining  $n - 1$ . From this high-dimensional representation, multidimensional scaling extracts a low-dimensional model, the dimensions of which correspond to the underlying factors that account for judgments of similarity. To see how this works, let's look at an example.

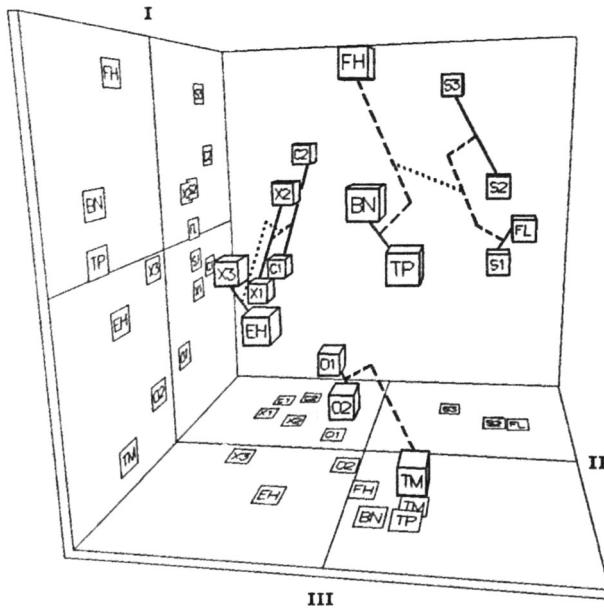
An early, influential analysis of musical timbre similarities into a low-dimensional timbre space was performed by Grey (1977).<sup>4</sup> Grey took recordings of common musical instruments and normalized them for loudness, duration, and pitch. He played these recordings in pairs to 20 subjects, asking them to rate the degree of (dis)similarity between each on a 30 point scale. Since there were 16 stimuli, these judgments could be represented in a 15-dimensional space for each subject. Grey then used multidimensional scaling to generate a low-dimensional geometric representation of the relative similarity distances, on the assumption that the same underlying features of timbre were responsible for all subjects' similarity judgments (Fig. 1).<sup>5</sup> Blocks in this space stand for perceived timbre qualities and are labeled by names of the instruments from which the corresponding stimuli were generated—BN for bassoon, FH for French horn, TM for trombone, TP for trumpet, FL for flute, EH for English horn, C's for clarinets, O's for oboes, S's for strings, and X's for saxophones. Lines connecting the blocks are the result of a separate analysis of the data using a hierarchical clustering algorithm.

How is Grey's timbre space to be interpreted? If it is to serve as a quality space, we'd like an understanding of its axes in terms of *perceptual* attributes of timbre—in the case of the color solid, for instance, axes correspond to the perceptual qualities of hue, saturation, and lightness, or to red–green, blue–yellow, and white–

<sup>3</sup> In fact, Wolpert (1990) demonstrates that non-musicians judge different melodies played on the same instrument to be more similar than the same melody played on different instruments (in contrast to trained musicians).

<sup>4</sup> More recent surveys in this research program confirm Grey's basic results, for instance McAdams et al. (1995), or Howard and Angus (2009), Section 5.3.2. For a general defense and discussion of multidimensional scaling as a method for studying timbre, see Plomp (1976, Chap. 6).

<sup>5</sup> In particular, he used the INDSCAL algorithm of Carroll and Chang (1970), which treats the individual differences of subjects as different weights on the axes of a common "psychological space." One feature of this algorithm (as opposed to other methods for the dimension reduction of similarity data) is that the space cannot be arbitrarily rotated (say, to search for more intuitively "meaningful" axes). This motivates the discussion in the text of the significance of these particular axes—axes that are largely confirmed by subsequent studies of musical timbre perception.



**Fig. 1** Timbre space derived from subject similarity judgments by means of multidimensional scaling (reproduced from Grey 1977, with the permission of the Acoustical Society of America)

black axes of perceived color opposition.<sup>6</sup> Without any antecedent expectations about the relevant attributes of timbre experience, however, Grey must begin his analysis by looking for qualities of the stimulus captured by each dimension. This at least is possible as Grey's stimuli may be precisely described in terms of their spectral compositions. We might then attempt to reinterpret these physical descriptions in perceptual terms. For instance, axis I correlates with both the width of the sound's spectral energy distribution and its center of mass. Narrow bandwidth and low frequency are at the top (the low, muted sounds of the French horn and the double bass), while wide bandwidth is at the bottom (the trombone's tendency to produce a very broad spectrum of overtones, both low and quite high, at the same time). Does this stimuli-defined attribute correspond to a perceptual one? A likely candidate is "brightness," one of the most consistently identified perceptual

<sup>6</sup> More generally, since a quality space is a model of perceptual experience, it is organized by perceptual attributes; in some representations these may not correspond directly to the axes of the space, but should be recoverable through rotation or some simple transformation. In the case of color, only some models exhibit a direct correspondence between axes and intuitive perceptual qualities, for instance those devised to capture psychological features of color such as the Swedish Natural Color System. Color solids derived directly from psychophysical data, such as CIELAB, or motivated by technological concerns, such as the CMYK space, may be defined for convenience by axes that do not correspond directly to psychologically salient qualities, but they exhibit the same gross structure as the NCS in virtue of preserving the same similarity relations. (See Kuehni and Schwarz 2008, for a survey of color solids and their properties.) Isaac (2014) argues at length for the importance of characterizing perceptual attributes in psychological terms, distinct from the physical attributes of the stimulus.

attributes of timbre across a wide array of different studies (e.g. Alluri and Toiviainen 2010; Pressnitzer et al. 2015).

The other two axes are much more mysterious. Grey identifies axis II with the synchronicity of the dynamics across the high end of the spectral distribution—harmonics of the woodwinds enter and exit the sound together, while in the strings and flute, harmonics fade in and out over the course of the sound at very different rates. Axis III captures something about the inharmonicity in the sound's attack (i.e. the onset of its volume envelope)—at one extreme, the trombone and some saxophones exhibit high-end “noise” during attack, while at the other, the double bass, cello, and clarinets exhibit low-end inharmonic grumbling during attack. These attempts at describing the regularities in the spectral composition of stimuli that track axes II and III are quite rough; moreover, there are no obvious perceptual attributes to which they correspond. Subsequent work has improved the precision with which the regularities in the stimulus that motivate timbre judgments may be described, but the identification of these with intuitive timbre qualities continues to be tentative (e.g. Howard and Angus 2009, esp. 250–2).

So, studies in the psychophysics of timbre are not yet adequate to play as rich a role in the timbre realism debate as the color solid plays in the color realism debate. Nevertheless, even without a complete account of the primitive features of timbre, studies such as Grey's provide evidence about the similarities between timbres that must be explained (or explained away) by an adequate physicalist theory of timbre. Indeed, one route to advance our science of the experience of timbre is to refine our hypotheses about the correlates of timbre in the stimulus. Perhaps it is difficult to see intuitive patterns in a timbre solid such as Grey's because a description of the stimuli in terms of spectral composition does not correctly characterize that aspect of the sound signal that our timbre experience detects and encodes. To draw another analogy with color: once we hypothesize that color experience detects surface reflectance properties (as opposed, say, to the spectral composition of light waves incident at the retina), we can use descriptions of stimuli in terms of surface reflectance to describe our psychophysical data and regularities in surface reflectance to motivate the design of new experiments (as, for instance, in the case of asymmetrical matching experiments). What then are the prominent scientific theories of the external correlates of timbre experience?

### 2.3 Two theories of timbre

Arguably, the scientific study of timbre begins with Hermann von Helmholtz's *On the Sensations of Tone* (1885). Helmholtz's investigation of sound and how we experience it integrated methods and theories from acoustics, psychoacoustics, and auditory physiology. To the already established view that sounds are waves in a medium, he added two key ingredients upheld in most subsequent scientific discourse: perceived sound is the proximal wave at the ear, and timbre is a property of this proximal wave. However, there are phenomenological reasons to doubt this story. Our auditory experience is not of sounds as mere proximal disturbances: we appear to hear distal events (I can hear the construction outside my window, for instance), and we experience sounds as maintaining identity while traveling through

a medium (I hear my neighbor's TV through the wall). Even more compelling is the corresponding epistemic intuition: we learn about distal events through their sounds. The insight that sounds convey information about the environment suggests a conception of timbre as a property of distal events, not of sounds *per se*; this view forms the basis for an alternative approach to the scientific study of timbre advocated by J. J. Gibson (1966). The remainder of this section elaborates these two scientific views in more detail before we turn to philosophical theories of timbre.

Helmholtz (1885) is responsible for initiating the characterization of the sound stimulus articulated in Sect. 2.1. He explicitly employs Fourier's Theorem to mathematically describe the sound stimulus incident at the ear. For Helmholtz, a first pass theory of timbre<sup>7</sup> identifies it with the overtones of the sound stimulus, i.e. the pattern of component waves that accompany the fundamental frequency. On this view, there is a continuity between timbre and chord, a continuity Helmholtz believed to be supported by psychophysical evidence—if one listens carefully to a single note struck on the piano, for instance, it will resolve itself into its component frequencies, and be perceived no longer as an isolated tone, but as a chord. Furthermore, this psychophysical theory conformed with Helmholtz's physiological theory of hearing, on which the frequency of each component of a complex wave excites a different position along the length of the cochlea, and the nerve fibre excited at that location determines the pitch component heard (the so-called "place theory" of sound perception). On this view, the ear essentially performs a Fourier transform on the incoming sound wave, and thus it seems natural to identify the timbral quality of the sound with the pattern of simpler waves derived through this process.

Helmholtz himself was already aware of inadequacies in this theory. Most obviously, the Fourier transformation is defined over a standing wave, which is homogeneous in time. Yet we rarely perceive such homogeneous sounds—the more typical cases are sounds which change dynamically in time, with harmonic components growing softer or louder over the course of the sound's duration. The natural extension of Helmholtz's view is to take timbre to be *some function* of the spectral components of the incident wave and their dynamics. Since Fourier's Theorem ensures that all possible sound stimuli may be described in terms of their component waves, this view has the advantage of ensuring that it covers all possible timbres, even if the function from the spectral composition of waves to timbres is unknown. It is perhaps for this reason that an updated Helmholtzian view along these lines appears typical amongst acousticians and psychophysicists.

Gibson (1966) accepts the general view of sounds as disturbances in a medium, but he emphasizes the *informational content* of these disturbances. What we learn about first and foremost on hearing a sound is not the proximal state of the medium,

---

<sup>7</sup> The German term is *Klangfarbe*, literally tone or sound "color"—etymologically encoding the analogies between timbre and color highlighted in the text. Despite the recommendation of colleagues, Helmholtz's translator Ellis refused to render this as "timbre," since it "is a foreign word, often odiously mispronounced, and not worth preserving" (24), preferring instead "tone quality." Nevertheless, "timbre" has survived as a technical term in both musicology and psychology, and is the standard modern translation of *Klangfarbe*.



but the distal source of the sound. In line with his general ecological perspective, Gibson argues that the timbre categories we hear should be identified with distal event types of importance to the organism on an evolutionary timescale. During sufficiently long periods of our evolutionary history, for instance, the distinctive timbral features of the snapping of twigs or the growl of a tiger allowed us to identify an event of interest (the presence of a predator) and its spatiotemporal location. Selection pressures instilled in us the perceptual power to directly detect these relevant mechanical disturbances as timbral categories.

Gibson's view has been influential, yet remains a minority position in the science of perception. One issue for the study of timbre in particular is that the Gibsonian tradition has yet to develop a comprehensive taxonomy of timbre in ecological terms. Consequently, even studies in ecological psychoacoustics (e.g. Neuhoff 2004) employ the apparatus of Helmholtz (Fourier analysis, spectral decomposition) when a precise characterization of the stimulus is required. Nevertheless, a lingering motivation for Gibson's approach, and one that has resulted in the basic ideas being rediscovered or remotivated many times, not least by recent philosophy of sound, is the phenomenological plausibility of its ontology: we don't experience sound qualities as occurring proximally at the ear, but as properties of distal events or objects.

### 3 What has timbre?

With the science of timbre under our belts, we can turn to the philosophy of timbre. If we are to seek a physicalist reduction of timbre, we must first ask what sorts of entities are the bearers of timbre. The general metaphysics of sound may influence our answer to this question; if sounds are the bearers of timbre, then our ontology of timbre will depend on our ontology of sound. Nevertheless, it is conceptually possible that sounds themselves are not the bearers of timbre, and thus that our ontologies of sound and of timbre come apart. Gibson's view takes this form: sounds are disturbances in a medium, but timbres are not properties of these disturbances, rather they are properties of sound sources.

Since the ultimate target of our discussion is timbre physicalism, I will not canvas the host of possible subjectivist, or anti-realist theories of timbre. Just as in the case of color, for instance, one might be eliminativist, arguing there are no actual timbre qualities in the world, and that our attributions of timbre are in widespread error; or one might be dispositionalist, arguing timbres are dispositions in external objects or events to cause particular sensations in us. For an extended argument against anti-realist theories of sound and its qualities, see O'Callaghan (2007).

Considering only realist accounts compatible with physicalism, there are three broad candidates for the bearers of timbre: waves, material objects, and resonant events. Each potential bearer of timbre suggests a particular objective physical quality to which timbre might be reduced: for waves, the spectral composition; for objects, the disposition to resonate; and for resonant events, the component mechanical interactions. I consider each of these candidate bearers of timbre in turn, surveying briefly the arguments for and against. The most popular position appears

to be that timbres are properties of resonant events or processes, in part because similarities between timbres are purportedly better explained by similarities between resonant events than by similarities between waves. Nevertheless, I will question the plausibility of this argument, suggesting some potential counterexamples.

### 3.1 Waves

Are soundwaves, or disturbances in a medium, the bearers of timbre? If one accepts that sounds are themselves waves in a medium, and that timbre is a property of sounds, then the answer will be yes. I take there to be two main objections to this position.<sup>8</sup> The first identifies an apparent incompatibility between our attributions of stable location to sounds and the view that they are traveling waves. A second line of critique turns on the explanation of timbre constancy: no known similarities between the spectral compositions of waves predict our attributions of timbre constancy and similarity. Examining these objections in turn will motivate the possibility that distal objects or events are the bearers of timbre.

Before turning to criticisms of the view that timbre is borne by waves, it is worth briefly stressing its advantages. This is the view that conforms most closely with mainstream scientific practice. Furthermore, the trajectory of research originating with Helmholtz suggests a natural reduction basis for timbre: the spectral composition of the wave and its dynamics. From Fourier's Theorem, we know the spectral composition is well defined, and have strong reason to think it will be adequate to reduce all possible timbres. More generally, if our perceptual experience of timbre is determined entirely by the incident wave, we know that it may be characterized by *some* function of this wave, if not that suggested by Fourier *per se*.<sup>9</sup> Thus, it looks as if the two criteria for an adequate reduction, that the basis be a scientifically well-defined category, and that it exhaustively reduce all the relevant qualities, will be easily satisfied.

Pasnau (1999) acknowledges the wave view as both the scientific standard and the view most clearly captured in ordinary discourse about sound. Nevertheless, he argues that both our scientific and folk theories of sound exhibit "incoherence" in their attribution of locations to sounds. If sounds are disturbances in a medium, they should be everywhere, filling the surrounding air. Yet we easily locate sounds in space, at their sources—the bird chirped *over there*; the thud of the jackhammer is *outside my office*. Cases where we locate sounds throughout space—his harsh laughter *filled the room*—are atypical and illusory, "analogous to seeing colours in a hall of mirrors" (312). Consequently, in order to avoid widespread error in our apparent perception of sound locations, Pasnau advocates abandoning the wave view of sound and treating sounds as properties of sources, much like colors. Pasnau's argument seems to apply *mutatis mutandis* to the particular quality of

<sup>8</sup> These are not, however, the only objections to the wave view, see Casati and Dokic (1994).

<sup>9</sup> See for instance Palmieri (2012) for a discussion of the possibility that some function other than the Fourier transformation may best describe the relation between stimulus and the experience of sound.

timbre. We typically localize timbres like chirping or thudding, and only in rare, confusing cases do we perceive timbres as nonlocalized and filling space (e.g. the disorienting omnipresence of the oscillating hum of cicadas on a summer's night).

While Pasnau's claim that we are typically able to localize sound sources is incontrovertible, his phenomenological assertion that we do not experience sounds as traveling through a medium seems suspect. He insists "Sounds that were caused at a distance seem to be at a distance; they do not seem to be coming towards you, unless that which makes the sound is in fact coming towards you" (311). Yet it seems to me that the expressions we use to describe sounds traveling while sources remain stationary have phenomenal content: a guitar sounds different if it "washes over you," "fills the room," or is heard "through the wall"; likewise, words may be "blown away" by the wind, "lost" in a large hall, or "swallowed up" by the surrounding din. If these phenomenal differences in perceived sonic motion are not themselves properties of the sound, of what are they properties?

Pasnau puzzles that "if sounds are qualities of the air, then it is hard to explain how, in virtue of hearing those sounds, we also manage to hear the objects that make the sounds" (317–8). Yet this is no puzzle at all if the qualities of the air at issue bear directional and timbral information. This consideration appears to motivate defences of the wave view against the location problem. For instance, Sorensen (2008) points out that if we did experience sounds as expanding spheres of disturbance in a medium, i.e. as waves, pragmatic considerations would dictate that we locate sounds at their source—the center is the most informative point for locating a sphere (282*f*). O'Shaughnessy (2009) argues that sounds bear directional information, and thus allow us to locate their sources, but are not themselves at their sources. Tellingly, for him, the properties of sounds and their sources may be different, for instance, in duration—a short vibratory event, say a knock on wood, may produce a longer sound, for instance in a large auditorium, where it may reverberate.

Nudds (2009, 2010) is an advocate of the wave view who explicitly identifies timbres as properties of waves. He argues that our ability to detect and localize sound sources depends on the auditory system's capacity to analyze and group by plausible source the component frequencies in a complex soundwave incident at the ear. Since the relationship between sound sources and the patterns of frequencies they produce is lawlike, the auditory system can take advantage of regularities in the proximal soundwave to extract locational and categorical information about the source. Since the goal of this process is to provide information about sound sources, we can only make sense of it, and the regularities it exploits, by appealing to sources and their properties. Thus, this view has it that sounds are waves, timbres are properties of these waves, yet timbre categories can only be explained by appealing to the physical properties of distal sound sources.

One might think, however, that the failure of the wave view to explain timbre categories is evidence that timbres are not themselves properties of waves. This consideration motivates a second prominent argument against the wave view, which stresses the claim that the spectral composition of waves does not adequately explain the constancy in our attributions of timbre or its similarities. Anecdotally, we perceive timbre to remain constant across changes in the source that affect the

acoustic properties of the signal, i.e. its spectral composition. The soundwaves from a violin played in the same room as the listener, from one played down the hall, and from one played outside the window will be quite different in spectral composition, yet we perceive them as “the same,” or at least very similar, in timbre. What is the feature of the waveforms that is the same in this case? We as yet have no explicit account of what this could be. O’Callaghan (2007), for instance, cites the authority of Handel (1995) on this, who insists that “*no* known acoustic invariants can be said to underlie timbre” (89). Likewise, O’Callaghan and Nudds (2009) argue “[t]here are good reasons to doubt” that sounds “are determined by underlying acoustic features” (20):

Neither the sound of a car driving on a gravel road, nor the sound of wood striking wood, for example, corresponds to a simple or straightforward feature recognizable on the surface of the acoustic signal. Each is highly complex and probably requires mentioning features of its source to make its individuation intelligible. (20)

The fact that we can’t point to any systematic, or invariant, feature of a waveform that specifies its timbre plausibly counts against the identification of timbres with spectral composition.

These two lines of criticism against the wave view of sounds and timbres have motivated two prominent categories of alternative view. The location worries of Pasnau (1999) motivated him to identify sounds (*a fortiori* timbres) with properties of objects, since the locations of objects better explain our localization of sounds. The constancy worries of O’Callaghan (and many others) motivate the view that timbres are properties of resonant processes or events, as invariants in these events better explain the constancy in our attributions of timbre. Let’s address these views in turn.

### 3.2 Objects

Although Pasnau (1999) instigates the view that sounds are properties of objects, he later repudiates it, endorsing a version of the event view addressed in the next section (Pasnau 2009). Nevertheless, the basic idea has been developed and defended by Kulvicki (2008, 2014). Pasnau’s original, cautious view was that “sounds either are the vibrations of ...objects, or supervene on those vibrations” (1999, 316). Kulvicki defends the more extreme position that sounds are *stable dispositions* of objects to vibrate (e.g. when struck). In many respects, the stable dispositions view seems even more plausible for timbres than for sounds themselves—we naturally discuss the timbre of a violin as if it is a stable property of the instrument, revealed when it is played, but persisting even when it is silent. Nevertheless, several facts about timbre individuation speak strongly against this view: timbres depend for their identity on changes in the sound over time; the same object may emit sounds of radically different timbres; and timbres may be determined by interactions between multiple objects. These considerations appear to invalidate stable dispositions of objects as eligible candidates for the physical reduction of timbre.

Kulvicki's stable dispositions view is motivated by analogies with color. Just as we learn about properties of an object (its color) when light bounces off it, Kulvicki argues we learn about the properties of an object (its sound) when it is struck or otherwise mechanically disturbed. What are the properties we learn about when an object is struck? Kulvicki emphasizes the natural modes of vibration of a resonant object. If one flicks a number of different wine glasses, they tinkle at different pitches—this is because the physical structure of each wine glass permits it to vibrate easily at some frequencies, but not at others. The modes or patterns of vibration at which an object easily vibrates are indeed a stable, dispositional property of the object, and they matter for the sounds that object produces when struck, rubbed, or otherwise disrupted—this is why the exact shape of a violin matters for the sound it produces.

Yet there is more to the sound that comes from an object when it is stimulated than just “its natural frequencies of vibration in their ordinary proportions with respect to one another” (2008, 5). The natural frequencies of vibration enter and exit prolonged sounds coming from an object at different times; they ramp up with the attack of the sound, and ramp down with its decay. Kulvicki argues that this pattern of attack and decay should also be understood as a stable disposition of the object, since it is determined by the object's physical structure (7). A purported virtue of this enriched view is its ability to explain the constancy in our individuation of object sounds: why does a violin sound the same whether played in the room with me or down the hall? Because the stable disposition of the violin to resonant in a certain dynamic pattern (for Kulvicki: its sound) stays the same.

While Kulvicki pushes the stable disposition view as a theory of the ontology of sound, it may at first appear even more appealing as an ontology of timbre. Many of the examples he discusses, for instance our talk of the “sound” of a bell or the “sound” of a violin, are most naturally understood as metonymic for the *timbre* of the bell or violin. We don't intend to refer to any particular pitch, loudness, or duration when we refer to the “sound” of a musical instrument, but rather to its sound quality, or timbre. If timbres are properties of objects, then Kulvicki's stable disposition view is the most promising candidate for their physicalist basis: the natural modes of vibration, and the natural pattern of attack and decay of these, determined by the physical structure of the object. Nevertheless, the stable properties of objects are not a rich enough basis for the correct individuation of the full array of possible timbre categories and their similarities.

One problem is that timbre identity is constitutively tied to changes in a sound over time. Kulvicki thinks he can account for this by appealing to the natural pattern of attack and decay determined by the object's structure. A problem with this solution, however, is that the same physically determined attack, decay, and overtones may contribute to quite different timbres. Consider, for instance, how the sound of a violin changes with the pressure with which it is bowed, or the sound of a trumpet changes with changes in the tightness of the lips as it is blown. In these cases, the natural modes of resonance determine the harmonics in the sound, the natural patterns of attack and decay determine how the resonant frequencies enter and exit the sound, yet still the sounds produced are in some sense different. Since the various violin notes or trumpet notes may be identical in loudness, pitch, and

duration, timbre is the critical factor that distinguishes them.<sup>10</sup> One way to make sense of the difference here is that, despite rough similarity in the pattern of attack and decay, the exact timing at which harmonics enter and exit the sound, and their exact relative strengths, differ for the various notes. The stable dispositions view seems to have no resources to account for this difference, as the stable dispositions of the object have not changed at all, while the timbre of the sounds it produces has.

This example is really just a special case of the more general point that the same object may produce a wide variety of timbres: a violin's strings may be bowed or plucked; its body may be tapped against or knocked; the palm of the hand may be dragged against its back to produce a loud squeak; it may be struck against a wall and shattered. Even if it is true that all these sounds share some component determined by the stable dispositions of the violin, they also differ in some way, and insofar as these differences are differences of timbre, then the stable properties of the violin do not explain them. In fact, I think examples such as this reveal just how heavily the stable dispositions view relies on musical examples for its plausibility; when we speak casually of the sound or timbre of a violin, the kind of consideration that motivates Kulvicki, it is really shorthand for the sound or timbre of a violin *played in the usual manner*—without reference to our usual mode of interaction with the object, a reference implicit in musical instrument examples, there is no way to identify *the* timbre of the object.<sup>11</sup>

A yet more general issue here is that timbres are typically determined by interactions between more than one object. For Kulvicki, the paradigmatic case is a short, sharp “thwack,” such that the thwacked object has the timbre, and the thwacking one merely reveals it. It is not clear how his view should extend to sound producing interactions that are radically different from thwacks, however. Consider the sound of a saw against a log: the distinct timbre of *sawing* is not obviously a

<sup>10</sup> One might think that increased pressure while playing the violin results in an increase in loudness, but this is not necessarily the case. Increased pressure from the bow creates greater friction with the string, impeding its movement. A slower bow at greater pressure may produce a note at the same loudness as a faster, lighter bow movement, yet the two will sound different. In the case of trumpets and other wind instruments, the sound is determined in part by the “embouchure” or tightness and shape of the lips while playing. Changes in embouchure can produce changes in pitch or loudness, as for instance on a bugle, where all such changes are so produced, but they may also produce differences of tone or sound quality while pitch and loudness remain stable. Examples such as these, where a change in performance technique produces a discernible change in timbre, but no change in pitch, loudness, or duration, are easy to find for any musical instrument (the only exception are those rare instruments where timbre is mechanically inaccessible to performer technique, e.g. a pipe organ).

<sup>11</sup> c.f. Davies (2010), who argues that timbres are properties of musical instruments, but depend constitutively on the characteristic manner in which the instrument is played. Kulvicki himself has the resources to rescue his ontology of sound from these apparent counterexamples by abandoning the view that timbres are properties of sounds. For instance, Kulvicki (2014) defends the stable dispositions view from the accusation that sounds are individuated by durations by arguing that the phenomenal evidence is consistent with the claim that it is not sounds, but “merely the episodes in which sounds can be heard [that] have durations” (210). Kulvicki might likewise insist that it is not sounds themselves, but episodes in which sounds are heard that have timbres. Here, again, the ontologies of sound and of timbre would come apart; in this case, sounds would be properties of objects, but timbres would be properties of the resonant mechanical interactions that reveal those sounds—essentially the view discussed in the following section. (It is not clear to me that this defense will work for the example in the following paragraph, however.)

consequence of the resonant features of either the log or the saw. The closest example to this kind of interaction discussed by Kulvicki is that of a glass shattering against a stone floor, of which the stable dispositions view might say “the vibratory dispositions of the objects in question are changing quite quickly” (2014, 212). But this kind of explanation won’t work for the sound of sawing—the stable resonant dispositions of the saw and the log are pretty much the same before and after sawing; they sound the same when thwacked. Yet the sound of the sawing itself does not appear to involve, or depend at all, on these vibratory dispositions, but rather on the nature of the mechanical interaction between the teeth of the saw and a resistant solid.

These considerations all point to the conclusion that an analysis of timbres as stable dispositions of objects to resonate is untenable. In particular, a key criterion for an adequate theory of timbre, the correct individuation of timbre categories and their similarities, is not fulfilled. The counterexamples above suggest a more promising alternative: timbres are not properties of objects, but of resonant events or processes in which the objects participate. The same object (e.g. a violin) produces different timbres when it participates in different events or processes; likewise it produces similar timbres when it participates in similar events. The distinctive timbre of sawing is not a property of either the saw or the log, but of the mechanical interaction between them. Views such as this have come to dominate recent philosophy of sound.

### 3.3 Events

Worries about how sounds are individuated and located have driven many philosophers to the view that sounds are to be found in the vicinity of distal resonant events (e.g. Casati and Dokic 1994, 2009; O’Callaghan 2007, 2009; Roden 2010; Matthen 2010; Kubovy and Schutz 2010). Even once one has honed in on events as the general locus for an ontology of sound, there are a host of further metaphysical questions one might ask: are sounds events themselves, parts of events, or do they supervene on events? Must resonant events disturb a medium to subvene sound, or may sounds occur in a vacuum? How should we conceptualize events: as bounded spatiotemporal regions, as processes or mechanisms, as particulars or universals? Since our focus is on timbre physicalism, I think it safe to abstract away from most of these disagreements. Once one identifies timbres with properties of audible events, the natural bases for a physical reduction of timbre are the vibrating processes and resonating interactions that contribute to the production of sound.

The view that timbres are properties of events is very close to that of Gibson, and it is somewhat surprising that recent philosophy of sound has not made much contact with his work or ecological theories of audition.<sup>12</sup> One reason may just be that Gibson himself identifies sound with the ensuing wave, not the event (see, e.g. Casati and Dokic 1994, 15). Nevertheless, the basic insight behind Gibson’s

<sup>12</sup> Gritten (2012) bemoans the lack of attention to Gibson in this literature, advocating for more engagement with his conceptual framework (a prominent exception is Kubovy and Schutz 2010; c.f. Davies 2010).

project, that invariant aspects of auditory experience are best explained by invariants in the sound source, does appear to be rediscovered in these views. A second reason that event theories of sound have not engaged Gibson may perhaps be his resolute anti-physicalism. While Gibson advocated direct realism—we directly access distal objects and events in perception—he insisted that the ecological environment we so perceive could not be reduced in any interesting way to its physical properties (Gibson 1986). Nevertheless, many event theories of sound are explicitly presented as “physicalist.”

The Gibsonian point that invariant features of sound sources best explain the constancy in our attributions of timbres and their similarities has been offered as a major argument in favor of the event view. Recall the worry that the wave view could not explain similarities in our perception of timbre, as no known similarities between waves predict similarities in perceived timbres. In contrast, an event view may do better on this score, if similarities between events do successfully predict timbre categories. Plausibly, the events involved in playing two different notes on a violin, or of playing a note on a violin and one on a cello, comprise quite similar vibratory processes and resonant interactions. This similarity in event type may then explain the similarity in the corresponding timbres. A representative argument is that of O’Callaghan (2007): “timbre quality ...depends ...upon features of the source and the characteristic manner in which it disturbs a medium” because “[t]hat is ...what remains constant across changes to its determinate audible qualities. The uniformity of timbre across sounds and circumstances is best explained by constancy in factors beyond the attributes of waves” (89).

A closely related, but somewhat more nuanced position is expressed by Roden (2010), for whom timbres are “variable sets of physical features” of a “sound generation mechanism” (145). He takes our ability to discriminate timbres as evidence for some flavor of timbre realism, but presents his view as “a more modest physicalism” than O’Callaghan’s (144). Roden is motivated by results such as those discussed in Sect. 2.2, namely that the underlying factors determining our timbral judgments, when described in terms of wave properties, may be quite complex. He worries that this complexity may resist reduction to any simple property of the audible event. Thus, while “[i]n traditional musical contexts we distinguish timbres in terms of typical mechanisms of sound generation” (144–5), timbre categories in general may not correlate with any simply definable mechanical quality.

Timbral discrimination, then, does not plausibly ‘track’ a single type of physical feature ...but relatively idiomatic patterns of relations between such features. This is consistent with a qualified interpretation of timbral kinds as consisting of recurrent constellations of features of sound generation processes, but it need not entail an essential limit on what kinds of relationships between more basic physical features can be picked out through identification of timbres. This seems plausible given that we normally use timbre to track complex processes such as the crying of babies (or cats), the percussion of hail on corrugated iron, or the motion of a fan blade in an extractor—not basic physical properties. (145)



So, the prospective timbre physicalism of the event view may be characterized in strong terms—timbres reduce to characteristic forms of mechanical disturbance—or in much weaker terms—timbres “track ...idiomatic patterns” of such physical features. But are these arguments correct? Do invariant features of sound sources or events in fact do well at individuating timbre categories? There are many anecdotal reasons to doubt this claim. While the connection between sound source and timbre may be more intuitively robust than that between waveform and timbre, there are nevertheless mysteries about its exact instantiation.

If timbre is really best explained by the mechanical interactions that participate in audible events, then similarity between these interactions should be a strong predictor of timbre similarity. In particular, if two events are very different, they should produce different timbres, and if two events are very similar, they should produce similar timbres. Nevertheless, there are abundant counterexamples to the first claim, and considerations that speak against the second. For instance, there are many examples of apparent “timbre metamers”: sounds assessed as similar in timbre, yet produced by radically different mechanical interactions—the phenomenon of the babbling brook, for instance, where the rushing of water over stones produces a sound similar to a room full of people engaged in conversation. Famously, the crying of a baby, a muted trumpet, and an electric guitar played through a “wah” pedal all sound very similar in timbre, yet are generated by three very different processes. The art of foley (*ex post facto* recording of sound for film) may frequently rely on using similar mechanical interactions to generate sound effects (coconut halves for the clapping of horses’ hooves), but not always, for instance when crinkled cellophane is used to foley the crackling of a fire.

Examples of radically different event types judged similar in timbre are found in the musical case as well. For instance, look again at Fig. 1; notice that the flute and violin are judged to be more similar to each other than either the violin is to other strings or the flute to other winds. Yet there are radically different processes going on here—the one involves the direct mechanical contact of a tense string and bow, as resonated through a wooden frame, the other the vibration of a column of air within a metal tube. Furthermore, arguments in favor of both object and event views imply the timbre attributed to a musical instrument should stay relatively fixed across changes in pitch. In fact, however, instruments that employ different sound production mechanisms may be confused readily in some parts of their pitch range, but not in others (Grey 1977, discusses this phenomenon with respect to confusion between bassoon and various brass instruments).

Are there cases where similar mechanical interactions generate radically different timbres? These are perhaps harder to find, yet there are still some prominent examples. We are all familiar with the danger of a very slight change in the angle at which chalk is used on a blackboard and the radical change in timbre that can occur from a gentle scratching to a piercing screech. More generally, it is not at all clear what features determine similarity between the mechanical processes that generate sounds in the first place. For instance, musical instruments are typically grouped by the gross features by which they produce sounds, but these do not always track similarity in timbre. A piano may be categorized as “percussion” since it generates sound through percussive events (hammers striking taut strings)—yet a piano

sounds nothing like a kettledrum; likewise, a violin bow may be rubbed against a variety of objects (a Tibetan prayer bowl, a musical saw, an electric guitar) and generate a great diversity of timbres.

These are cases where the gross mechanical features of the event seem a poor predictor for sound quality, while the detailed resonant features of the instrument, and the harmonics it produces, seem a much better predictor of timbre similarity. Yet the converse case may be found as well—simple changes to the gross features of a sound may predict timbre judgments far better than its spectral complexity. For instance, from both spectral and mechanical standpoints a violin and a clarinet are radically different. Mechanically, the violin produces sound through friction between the string and the bow, the clarinet through the vibration of a column of air induced through the vibration of its reed. Spectrally, the violin generates all harmonics above its fundamental at decreasing intensities, while the clarinet only generates odd numbered harmonics (in sound synthesis terms: violins generate “sawtooth” waves while clarinets generate “square” waves). Nevertheless, if you record a note from a clarinet and play it back by fading it gradually in, then gradually out, it will be judged very similar to, if not outright mistaken for, a violin. Thus, changes in volume envelope alone may radically affect our judgments of timbre category.

These considerations do not undermine the metaphysical coherence of a view that takes sounds to be events, and timbres properties of those events. They do, however, undermine (if not definitively so) the argument that invariant features of sound production mechanisms better explain constancies in our attribution of timbre. The crucial point is just that, while it may be a legitimate worry about the wave view that we do not have a simple theory of how to predict timbre from surface waveform, there is an analogous worry for the event view: we may have an intuitive grasp on the contributions made by features of events to timbre, but we do not have a detailed theory of how to predict timbres from those features.

#### **4 Does timbre reduce to properties of the sound event?**

To review the situation so far: we began with three candidates for the physical bearers of timbre: waves, objects, and events. The wave view takes timbres to be properties of the proximal wave incident at the ear. This view has two prominent advantages: it conforms to scientific practice, and Fourier’s Theorem guarantees that all timbres may be reduced to wave properties. Nevertheless, there are two marks against the wave view. One is the location problem: we perceive sounds, *a fortiori* timbres, as distally located. The second is the similarity problem: no known similarities between the spectral compositions of waves explain the similarities between timbres. The object view, identifying timbres with stable dispositions to resonate, solves the location problem, but it is demonstrably inadequate at individuating timbres. The event view appears to solve both the location problem and the similarity problem (modulo some worries raised in the previous section). What are the prospects for a timbre physicalism that reduces timbres to properties of events?

This section explores this question in more detail. I begin by revisiting color physicalism in order to highlight some analogies with wave timbre physicalism. These examples will set the standard of rigor that a successful event timbre physicalism must meet. I then discuss a plausible reduction basis for the event view: mechanical vibrations. Nevertheless, I ultimately conclude that there is a barrier to this reduction. In particular, sound sources dynamically interact with the surrounding medium. This means that there is no scientifically well-defined boundary between a distal audible event and the wave it generates. The upshot is that there is no legitimate scientific kind to which would-be event timbre physicalists might successfully reduce all timbres.

#### 4.1 Color physicalism and wave timbre

The gold standard for reductive physicalism about perceptual qualities is color physicalism (e.g. Byrne and Hilbert 2003). Color physicalism identifies colors with (classes of) surface spectral reflectance (SSR). Since SSR exhaustively describes the results of all possible interactions of a surface with light, we have strong reason to believe that it subvenes all possible (surface) color categories. These categories may not themselves have physical significance, i.e. scientific interest independent of their role in generating perceived color. Nevertheless, the space of all possible SSRs is precisely defined from a physical standpoint. An SSR is a function that assigns to each possible wavelength within a region of the electromagnetic spectrum a percentile representing the relative degree to which that wavelength is reflected; mathematically, this is just a function from a region of the real line to values in the interval  $[0, 1]$ . The space of all SSRs is just the set of all such functions. While the groupings of SSRs that correspond to color categories as we perceive them are not of primitive physical interest, SSRs themselves, and in general, bands of energy within this range of the electromagnetic spectrum, are objectively significant as they reveal underlying physical properties of objects—hence the use of spectral reflectance to analyze the elemental composition of meteorites (Gaffey 1976) or the quantity of chlorophyll, and thus photosynthetic activity, in plants (Myneni et al. 1995).

A reduction of timbre to the properties of waves identified by Helmholtz would have similar features. The space of all possible combinations of simple sine waves is well-defined, although the defining functions are more complex than those for SSR. Three numbers are needed to specify each sine wave: period, amplitude, and relative phase. Furthermore, as noted above, sounds change dynamically in time; as such, each sine composing a complex wave will need its own dynamic envelope—an attack, sustain, and decay—representable by a function from a region of the real line (length of the sound) to values in the interval  $[0, 1]$  (relative volume). Fourier's Theorem guarantees that this strategy is adequate to capture the complete range of complex waves, and thus of sonic possibilities. Just as with color, it may well be that classes of spectral composition assigned the same timbre do not have independent physical interest—this is implied by the similarity problem. Nevertheless, the Fourier decomposition of complex acoustic signals into their components is scientifically important for understanding the propagation of waves through media

entirely independent of our sense of hearing, for instance in the analysis of sonar data to map the seafloor and determine properties of the deep sea habitat (Brown et al. 2011).

So, both the reduction of colors to SSR and of timbres to spectral composition share these features: the reduction basis is (i) scientifically well-defined and of independent physical interest, and (ii) adequate to reduce all possible categories of the perceptual quality, despite perhaps (iii) the corresponding classes within that basis not themselves having physical significance. Features (i) and (ii) seem jointly necessary for the adequacy of a proposed physicalist reduction. If (i) is not satisfied, then it would seem that the reduction basis is not a legitimate physical kind.<sup>13</sup> If (ii) is not satisfied, then it would seem the proposed correspondence between some perceptual qualities and their physical correlates is not truly a reduction. (Of course, if (iii) is violated, and the perceptual qualities may be reduced to physical classes of independent interest, so much the better.)

## 4.2 Vibrations?

Does the event view suggest a candidate basis for reduction that satisfies these two conditions: it is physically well-defined, and it subvenes all possible timbres? A plausible candidate is the set of mechanical vibrations that participate in an audible event. Unfortunately, accepting vibrations as the reduction basis of timbre requires abandoning an appealing feature of the event view: its solution to the similarity problem. To see this, let's look at the physics of musical instruments.<sup>14</sup>

Any body stiff enough to return to an initial position once displaced and possessing of inertia such that it might overshoot, and thus fluctuate around an equilibrium, will mechanically vibrate. Typical musical examples include bowed, plucked, or hammered strings, the wooden bodies of stringed instruments, drumheads, etc. Vibrating bodies, or oscillators, can be decomposed into their characteristic “modes,” or independent degrees of freedom, and oscillators may be physically linked such that they are “coupled,” or functionally interdependent. For instance, the different wooden components that make up a violin may be considered distinct vibrating bodies, but because of the complex physical interactions at their joints, as mediated also by the vibrating air trapped inside, the violin body as a whole exhibits complex modes of vibration.

How well do we understand the relationship between the gross mechanical events in which instruments participate and their subsequent patterns of vibration? Not very well. For instance, although some progress has been made (especially recently, with the aid of computers, e.g. Bretos et al. 1999) in simulating violin body vibrations with numerical methods, our physical understanding of a violin does not

<sup>13</sup> Note that this does not rule out disjunctive kinds; it merely requires that the set of disjuncts of such a kind itself be precisely definable in physical terms. The case of SSR illustrates this—a diverse set of microphysically distinct surface interactions result in the “reflectance” of some, but not all, of the light incident on a surface (for a detailed survey see Nassau 2001). Nevertheless, SSR provides a precise way to exhaustively characterize this set through its effects on the behavior of incident light.

<sup>14</sup> This section draws heavily on Fletcher and Rossing (1991), especially Chaps. 1 and 10.

permit an analytic derivation of patterns of vibration from its material and mechanical properties. Rather, empirical measurements of modes of vibration are made, and these are used to develop more accurate models of its physical structure.

In general, the physics of musical instruments proceeds by translating the mechanical interactions between components of different material and shape into vibrational modes, then combining these modes into overall vibratory response. Particular materials or gross mechanical interactions by themselves are not enough to identify the characteristic features of a complex sound event—the plucking of a string sounds radically different if it is connected to a guitar, a harpsichord, or a harp. The means of specifying these differences physically is through a piecemeal analysis of the different ways the motion of the string affects, and is amplified by, the vibratory properties of the instrument as a whole. Thus, the general theory, the one which subsumes complex mechanical process types that generate sound within a single framework, is the physics of vibration.

Is there reason to think the physics of vibration can provide a basis adequate to reduce all timbre categories? Well, we have reason to think it can provide a well-defined characterization of all possible complex vibrations. Since the Fourier Theorem applies to  $n$ -dimensional waves, and since vibrations are mathematically analogous to waves (both are periodic motions),<sup>15</sup> any complex vibration will be equivalent to some combination of primitive vibrations. Since the physics of vibration proceeds by analyzing and combining primitive vibrations (associated with each mode), it engages in an endeavor that will in principle generate all complex patterns of vibration.

I'll argue in the following section that this approach will not in fact successfully subvene all timbre categories. Nevertheless, let's pause for a moment and consider the consequences for an event view of timbre that reduces timbres to sums of simple vibrations. Timbre physicalism of this form would ensure reduction at the cost of sacrificing its solution to the similarity problem. Event theorists appeal to the gross mechanical features of audible events to explain similarities in perceived timbre—the striking of a hammer, the patter of hail on a tin roof, the bowing of a violin string—but these gross features are not in general reducible to, nor recoverable from, invariants in the pattern of component vibrations. To specify the distinctive timbre of hail on a tin roof in terms of some invariant in the relative degrees of the roof's (*and each piece of hail's!*) modes of vibration would be every bit as counterintuitive and unilluminating as doing so in terms of the relative degrees of

---

<sup>15</sup> There is actually a deep three-way analogy here, between mechanical, acoustical, and electrical systems. All three involve periodic fluctuations (vibrations, waves, and, e.g., alterations in current), and thus the corresponding physical theories are intertranslatable. However, since the primitive quantities are different in each case, and the details of each area are such that the behavior of one may map to the other in multiple ways, these mappings are merely *analogies* (rather than mathematical identities). Such analogies were originally employed (e.g. by Maxwell) to motivate intuitions about electromagnetism, but now that electrical systems are better understood, they are more frequently used to simplify analysis of mechanical systems. Another important area of application is at the various interfaces between mechanical, electrical, and acoustical phenomena: for instance when an electrical signal is translated into movements of a speaker cone, which themselves are translated into disturbances in the air (see e.g. Olson 1947, Chaps. 4 and 6).

component sines in the issuing sound wave (and astronomically more difficult). But if component vibrations are the physical feature to which timbres reduce, it seems this is exactly what the event physicalist must do.

### 4.3 Vibrations couple with disturbances

One might think that losing a solution to the similarity problem is an acceptable price to pay for the reduction of timbre to vibration. Perhaps the situation is analogous to the metamer problem for color—SSR similarities don't explain color similarities, but that does not defeat the identification of color with SSR. However, there is an important disanalogy between the case of color and the case of sound, and it is because of this disanalogy that the reduction of timbre to mechanical vibration ultimately fails. Light in the visible range of the spectrum does not interact substantively with a surface when reflecting from it, altering its reflectance properties.<sup>16</sup> In contrast, disturbances in the medium surrounding a mechanical process interact with that process substantively, altering its pattern of vibration and contributing to the emitted waves. Consequently, there is no physically principled way to draw a distinction between the contributions of mechanical vibrations and of nearby disturbances in the medium to determining a sound event, and thus timbre. Without such a distinction in hand, there is no principled way to delimit the audible event as a scientifically legitimate kind independent of the overall pattern of interaction culminating in a wave incident at the ear. If audible events are not scientific kinds, then they are not fit basis for a properly physicalist reduction.

My claim is that sound events (typically) involve an active coupling between disturbances in the medium and vibrations in the object. In order to understand what this claim amounts to, and its significance, it is important to distinguish it from two nearby facts that do not block timbre event physicalism. The first is just the observation that features of the medium may change the perceived quality of the sound from that determined by its source. The violin played next to me and the one played down the hall do sound similar, but they also sound in some sense different. An analogous phenomenon in color vision is that of haze, when particles in the air diffract some wavelengths of light, altering the color signal during its journey from surface to eye. Just as we attribute invariant colors to surfaces despite haze, we attribute invariant timbres to sources despite intervening changes in the sound signal—this does not undermine an identification of the timbre of the original sound event with properties of that event.

A second observation is that disturbances in a medium can cause sound events. This also does not necessarily undermine an identification of sound qualities with properties of the event. For instance, a wave may interact with an object by inducing it to vibrate at its resonant frequencies, as when an opera singer shatters a wine glass by inducing it to vibrate with her voice. Resonance such as this is an essentially

---

<sup>16</sup> At least not typically, or at relevant time scales. Some substances are unstable in visible light, and some properties of objects are changed over long time periods of exposure to light (think fading of dyes in sunlight); nevertheless, these effects do not contribute to our real-time perception of color.

passive phenomena. While the wave initiates, and perhaps even drives it, the quality of the sound event itself is determined by the resonant properties of the object. So, even in a case of resonance induced by a disturbance in the medium, there is a principled boundary between properties of the sound event, and nearby properties of the medium.

The much more problematic case obtains when a disturbance dynamically couples with a mechanical process to constitutively determine properties of the sound event. A dramatic example of such dynamical coupling from outside the realm of sound is the 1940 Tacoma Narrows Bridge collapse. The original Tacoma Narrows Bridge was known for vertical (transverse) vibrations in high wind. On November 7, 1940, however, it exhibited a never before seen mode of vibration, twisting back and forth around its longitudinal axis. Subsequent analysis has shown that the cause was *not* mere resonance, but rather *aeroelastic flutter*—a phenomenon that occurs when coupling between a vibrating body and a fluid produces positive feedback. The eddies in the air caused by rotation of the bridge themselves caused a magnification in that rotation. Over the course of 45 minutes, this positive feedback induced the bridge's rotation to gradually increase, until eventually reinforcing cables snapped and it collapsed. The point of this example is that the bridge's behavior was determined more by the nature of its interaction with the surrounding air flow than by its intrinsic resonant modes, and thus the collapse can only be understood in terms of the dynamic interaction between a mechanical vibration and waves in the surrounding medium (Billah and Scanlan 1991).

Less dramatically, the physics of musical instruments depends also on dynamic interactions between mechanical vibrations and the surrounding medium. In the case of violins, for instance, some modes of vibration are mechanical, involving movement of the body, but some are acoustical, involving waves in the air cavity. These mechanical and acoustical fluctuations are coupled with each other, and any explanation of the sound of the violin must appeal to both.<sup>17</sup> The importance of such interactions becomes even more acute in the case of wind instruments. In some cases, such as the flute or trumpet, the initial mechanical event, the vibration of the lips, determines very little about the overall sound, while the resonant behavior of the trapped column of air is enormously important.

One might think that an easy way to circumvent these examples, and maintain the integrity of violin bowing and flute blowing as well-defined audible events, would be to treat the trapped air as a special case. The part of the medium filled with waves coupled to the mechanical vibrations of the object is internal to it, doesn't that give us a natural boundary between these waves and the rest of the medium? And won't that natural boundary serve to circumscribe the event? The problem with this approach is that it is precisely the continuity between waves occurring in the instrument's cavity and the surrounding medium that ensures the musical sound event is successful—only at low-impedance boundaries such as those at the f-holes of a violin or the bell of a trumpet is sound quality efficiently communicated. In

---

<sup>17</sup> For instance, Bretos et al. (1999) identify omission of the vibrations in the air cavity from their computational model as a primary source of the discrepancy between its behavior and empirically recorded violin body vibrations.

other words, there is no physically significant boundary between waves internal to the instrument and external to it, so the fact that part of the medium is enclosed by the instrument also does not serve to establish a physical boundary between sound event and traveling wave. From the standpoint of physics, there is a distinction between mechanical vibrations and disturbances in a medium, but not a distinction between those oscillations contributing to an audible event and those that merely constitute its effects. More generally, acoustics and the physics of musical instruments treat mechanical vibrations, near sound waves, and far sound waves as all homogeneously interacting. Analysis of the “sound” of a violin does not stop at its body, but at the other end of the concert hall, in the ear of the listener.<sup>18</sup>

## 5 Conclusion

What then are the prospects for timbre physicalism? The situation is subtly disanalogous to that of color physicalism. In the case of color, would-be physicalists were impressed with the constancy with which colors are assigned to distal surfaces and were able to find a well-defined physical kind, surface spectral reflectance, to serve as a reduction basis. The question of whether timbre, or sound in general, should be identified with proximal or distal physical categories is much more vexed. Our ability to spatially locate sound sources, and to individuate them by timbre, has motivated a form of timbre physicalism that identifies timbres with properties of distal resonant events. However, the most obvious candidates for the distal correlates of event timbre do not appear to form a physically well-defined category. In contrast, correlates of proximal wave timbre do appear to constitute a well-defined category, but reduction of sound to properties of the proximal signal would fail to account for apparent spatial and categorical aspects of sound perception.

I have considered the possibility that event timbre might be identified with either vibrations of rigid bodies, or with some combination of such vibrations and nearby waves. I argued that the first suggestion, while well-defined, does not subvene all relevant aspects of an audible event; and the second suggestion, if it is to include only properties of the distal sound event as intuitively understood, is not scientifically well-defined. This does not mean that other candidates for the physical reduction of timbre are not available. For instance, if the would-be event timbre physicalist can find some physical feature of vibrations and nearby waves that captures their intuitive unity in an audible event, and appropriately

---

<sup>18</sup> Compare: “The study of acoustics is greatly simplified by understanding the circumstances governing the flow of sound energy because instruments, ears, and rooms can all be viewed as networks of interconnected vibrating elements” (Loy 2007, 325).



distinguishes them from emitted waves, it would serve as a plausible candidate for timbre physicalism.<sup>19</sup>

A second option for those who would identify timbres with properties of events is to adopt a more modest, non-reductive realism. One possibility is ecological realism along the lines suggested by Gibson. An ecological theory of timbre, identifying it with event properties of interest to organisms on an evolutionary timescale, would better satisfy many of the intuitive arguments in support of timbre “physicalism,” without requiring any strong reductive program. Nevertheless, there are challenges for this approach as well: what is the full taxonomy of ecological audible events? What ecological features determine timbre similarity? These are topics on which philosophers of sound might fruitfully collaborate with musicologists and ecological psychologists.

**Acknowledgements** This paper has benefited substantially from the suggestions of an anonymous reviewer, as well as those of Casey O’Callaghan, Dmitri Tymoczko, and J. E. Wolff. I am also grateful for discussions at earlier presentations of this material at Queen’s University Belfast, the Southern Society for Philosophy and Psychology, the Edinburgh Philosophy Society, the Royal Musical Association Music and Philosophy Study Group, and the Glasgow Philosophy, Psychology, and Neuroscience Research Seminar.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Alluri, V., & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception*, 27(3), 223–242.
- Billah, K. Y., & Scanlan, R. H. (1991). Resonance, Tacoma Narrows bridge failure, and undergraduate physics textbooks. *Am. J. Phys.*, 59(2), 118–124.
- Bregman, A. S. (1990). *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

<sup>19</sup> One possibility: identify sound events with networks of tightly coupled oscillators. This would help enormously for musical examples, as the air in a violin’s cavity acts as a Helmholtz resonator, an oscillator mathematically equivalent to those that describe the simple harmonic motions of the violin’s body. What is the threshold for sufficiently tight coupling, however, and how would this account handle non-musical sound events, such as a car crash?

A second: O’Callaghan (personal communication) and an anonymous reviewer have both expressed optimism that identifying timbres with *dispositions* of events to disturb a medium would allow one to precisely specify its distal correlates in terms of features of the resulting wave. However, the considerations introduced here raise two challenges for this view: first, the discussion above appears to show that the notion of an audible event is not well-defined, and so the question of what sort of entity possesses these dispositional properties still remains. Second, as emphasized previously, if properties of the wave are used to identify the distal correlates of timbre, then it seems the event physicalist solution to the similarity problem must be abandoned. These challenges are not unsurmountable, but they are also by no means trivial.

- Bretos, J., Santamaría, C., & Alonso Moral, J. (1999). Vibrational patterns and frequency responses of the free plates and box of a violin obtained by finite element analysis. *J. Acoust. Soc. Am.*, *105*, 1942–1950.
- Brown, C. J., Smith, S. J., Lawton, P., & Anderson, J. T. (2011). Benthic habitat mapping: A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques. *Estuarine, Coastal and Shelf Science*, *92*, 502–520.
- Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *Behavioral and Brain Sciences*, *26*, 3–21.
- Carroll, J. D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of “Eckart-Young” decomposition. *Psychometrika*, *35*(3), 283–319.
- Casati, R. & Dokic, J. (1994). Chapter 3 (english translation). In *Philosophy of Sound*. Chambon, Nîmes. original title: *La philosophie du son*.
- Casati, R., & Dokic, J. (2009). Some varieties of spatial hearing. In M. Nudds & C. O’Callaghan (Eds.), *Sounds and Perception* (pp. 97–110). Oxford: Oxford UP.
- Churchland, P. M. (2007). On the reality and diversity of objective colors: How color-qualia space is a map of reflectance-profile space. *Philosophy of Science*, *74*, 119–149.
- Davies, S. (2010). Perceiving melodies and perceiving musical colors. *Review of Philosophy and Psychology*, *1*, 19–39.
- Fletcher, N. H., & Rossing, T. D. (1991). *The Physics of Musical Instruments*. New York, NY: Springer-Verlag.
- Gaffey, M. J. (1976). Spectral reflectance characteristics of the meteorite classes. *Journal of Geophysical Research*, *81*(5), 905–920.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton-Mifflin.
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, *61*(5), 1270–1277.
- Gritten, A. (2012). Review of sounds: A philosophical theory and sounds and perception: New philosophical essays. *British Journal of Aesthetics*, *52*, 430–434.
- Handel, S. (1995). Timbre perception and auditory object identification. In B. C. J. Moore (Ed.), *Hearing* (pp. 425–461). San Diego, CA: Academic Press.
- Hardin, C. L. (1988). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis, IN: Hackett.
- Helmholtz, H. (1954 [1885]). *On the Sensations of Tone*. Dover, Mineola, NY, 2nd edition.
- Howard, D. M., & Angus, J. A. S. (2009). *Acoustics and Psychoacoustics* (4th ed.). Oxford: Elsevier.
- Isaac, A. M. C. (2014). Structural realism for secondary qualities. *Erkenntnis*, *79*(3), 481–510.
- Kubovy, M., & Schutz, M. (2010). Audio-visual objects. *Review of Philosophy and Psychology*, *1*, 41–61.
- Kuehni, R. G., & Schwarz, A. (2008). *Color Ordered: A Survey of Color Order Systems from Antiquity to the Present*. New York, NY: Oxford UP.
- Kulvicki, J. (2008). The nature of noise. *Philosophers’ Imprint*, *8*, 1–15.
- Kulvicki, J. (2014). Sound stimulants: Defending the stable disposition view. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and Its Modalities* (pp. 205–221). Oxford: Oxford UP.
- Loy, G. (2007). *Musimathics: The Mathematical Foundations of Music* (Vol. 2). Cambridge, MA: MIT Press.
- Matthen, M. (2010). On the diversity of auditory objects. *Review of Philosophy and Psychology*, *1*, 63–89.
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, *58*(3), 177–192.
- Myneni, R. B., Hall, F. G., Sellers, P. J., & Marshak, A. L. (1995). The interpretation of spectral vegetation indexes. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(2), 481–486.
- Nassau, K. (2001). *The Physics and Chemistry of Color* (2nd ed.). New York, NY: John Wiley & Sons Ltd.
- Neuhoff, J. G. (Ed.). (2004). *Ecological Psychoacoustics*. San Diego, CA: Elsevier.
- Nudds, M. (2009). Sounds and space. In M. Nudds & C. O’Callaghan (Eds.), *Sounds and Perception* (pp. 69–96). Oxford: Oxford UP.
- Nudds, M. (2010). What sounds are. In D. Zimmerman (Ed.), *Oxford Studies in Metaphysics* (Vol. 5, pp. 279–302). Oxford: Oxford UP.
- O’Callaghan, C. (2007). *Sounds: A Philosophical Theory*. New York, NY: Oxford UP.

- O'Callaghan, C. (2009). Sounds and events. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and Perception* (pp. 26–49). Oxford: Oxford UP.
- O'Callaghan, C., & Nudds, M. (2009). Introduction: The philosophy of sounds and auditory perception. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and Perception* (pp. 1–25). Oxford: Oxford UP.
- Olson, H. F. (1947). *Elements of Acoustical Engineering*. New York, NY: D. van Nostrand Co., Inc.
- O'Shaughnessy, B. (2009). The location of a perceived sound. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and Perception* (pp. 111–125). Oxford: Oxford UP.
- Palmieri, P. (2012). Signals, cochlear mechanics and pragmatism: A new vista on human hearing? *Journal of Experimental & Theoretical Artificial Intelligence*, 24(4), 527–545.
- Pasnau, R. (1999). What is sound? *The Philosophical Quarterly*, 49, 309–324.
- Pasnau, R. (2009). The event of color. *Philosophical Studies*, 142(3), 353–369.
- Plomp, R. (1976). *Aspects of Tone Sensation*. New York, NY: Academic Press.
- Pressnitzer, D., Agus, T. R., & Suied, C. (2015). Acoustic timbre recognition. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of Computational Neuroscience* (pp. 128–133). Berlin: Springer-Verlag.
- Roden, D. (2010). Sonic art and the nature of sonic events. *Review of Philosophy and Psychology*, 1, 141–156.
- Sorensen, R. (2008). *Seeing Dark Things: The Philosophy of Shadows*. Oxford: Oxford UP.
- Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception*, 8(1), 95–106.