



BRILL

Multisensory Research 34 (2021) 219–231

brill.com/msr

Introduction to the Special Issue on Multisensory Perception in Philosophy

Amber Ross¹ and Mohan Matthen^{2,*}

¹ Department of Philosophy, University of Florida, 330 Griffin-Floyd Hall, Gainesville, FL 32611-8545, USA

² Department of Philosophy, University of Toronto, Maanjiwe nendamowinan (Rm. 6138), 3359 Mississauga Rd., Mississauga, ON L5L 1C6, Canada

European philosophers of the modern period generally acknowledged that the senses are our primary source of knowledge about the contingent states of the world around us. The question of modality was of secondary interest and was very little discussed in this period. Why? Because these philosophers were *atomists* about sense-perception, an attitude that makes multisensory perception impossible. Let us explain.

Atomists hold that all sense-experience is of ‘ideas’ — a somewhat over-simple, but still useful, way to think of these is as *images*. All ideas are ultimately composed of *simple* ideas. Atomists hold, moreover, that the intrinsic nature of a simple (or non-composite) idea is fully given by conscious experience of that idea, and in no other way. For example, *burnt sienna* is a simple idea because it is not composed of other ideas. Nothing about its intrinsic nature can be known except by experiencing it — a colour-blind individual cannot know what it is. It is, moreover, adequately and completely known when it is experienced; there is nothing more to know about it than is given by visual experience of it (see Note 1).

Now, on this account of simple ideas, distinctions among them cannot be analysed. For atomists, inter-modal distinctions, like all other distinctions among ideas, are primitive and based in experience. What, for example, is the difference between burnt sienna and the sound of a trumpet playing middle C? All that can be said is that they are experientially different from one

* To whom correspondence should be addressed. E-mail: mohan.matthen@utoronto.ca

another in a way similar to how Oxford blue and the sound of a flute playing middle D are different from one another. But since this difference is just a brute fact about experience, one cannot do much more than simply label it. Sensory atomists do not think that modality *explains* the nature of the colours or explains the difference between colours and sounds. Rather, modality is for them a grouping based on experienced similarity. Like *colour*, *visuality* is, for them, a similarity-based abstract feature that we find in the content of sense-perception — a ‘determinable,’ to use the terminology of W. E. Johnson (1921).

Atomists allow, of course, that we have empirical knowledge of the causal origin of these differences — almost everybody knows, for example, that visual ideas come from the use of the eyes and auditory ones from the ears. But such facts are not contained in the experiences themselves. For, as atomists insist, it’s logically possible that the idea of a trumpet playing middle C could have come from the eyes. Such origin-facts must be empirically ascertained — we come to know this in part by covering our ears and noting the diminution in volume. This adds nothing to our knowledge of the intrinsic natures of these ideas, atomists insist. For reasons like this, the atomist paradigm doesn’t admit the theoretical interest of facts about the modal source of simple ideas. Modality is of no more than taxonomical interest to the atomist. It does not *explain* phenomenology; it *is* phenomenology.

Because of this atomistic attitude, European philosophers of the modern era (i.e., from Descartes to the end of the twentieth century) have almost nothing of interest to say about sense-modality. We find the single important exception to this generalization is Molyneux’s Question, which has to do with whether the modalities intersect. Take this idea of the tactual type. Could it *also* be of the visual type?

Here’s how this famous incident went. In 1688, William Molyneux, an Irish physician and mathematician, wrote to John Locke to ask whether a newly sighted man would be able to distinguish a globe and a cube by vision alone. Could he, in other words, distinguish and identify shapes he already knew by touch using only his newfound sensory capacity. The question turns on whether visual experiences of the two solids reveal the same differences that visual experiences do. An atomist is committed to saying that *if* these shapes are simple ideas, there is no way of comparing the visual impression of shape with that provided by touch — the experiences are different, and so the ideas are as well. It was assumed, perhaps on this tacit assumption, that the answer must be negative.

This initial response to Molyneux’s Question assumes, however, that shape is a simple idea. But are they? Globes and cubes are spatially extended. Might they not be structured out of point-impressions? And might the structure not be the same in both senses? Following up on this possibility, Molyneux’s

Question led to much productive debate about whether there are differences of structure between the two shapes that would give the game away to a newly sighted man. For example, would the newly sighted man be able to discern the greater symmetry of the visual sphere and thus match it up with his touch-based idea? This said, Molyneux's Question seems to leave sensory atomism intact, diverting the debate about inter-modal comparisons to the arena of composite qualities.

Putting the ingenuity of Molyneux's Question aside, the atomistic program served as an effective means of shutting down philosophical discussion of multimodal perception. Take an example that sensory scientists have found particularly revealing — the McGurk effect (McGurk and MacDonald, 1976). Here, a video of a person speaking the syllable /ga/ is dubbed with a soundtrack of somebody speaking the syllable /ba/. Subjects listening to the soundtrack with the video turned on generally report that they hear some visually and auditorily intermediate syllable: for example, /da/. But when they close their eyes, or when the video is switched off, these same subjects accurately hear the phoneme in the sound-track — /ba/. (If the audio track is switched off, no determinate phoneme is perceived — vision alone is insufficient for speech recognition in such short snippets, even for experienced lip-readers.) The alternation between /ba/ and /da/ is caused by the video-track being switched on and off.

It is natural to take this effect to show that our perception of phonemes integrates sensory inputs from vision and audition. We *seem to hear* a syllable /da/. But sometimes when we do, what we *see* is partly responsible. Most speech scientists think that this confirms something that has long been known about speech perception in everyday situations: namely, that when we engage in conversation in noisy environments, we are helped by many non-auditory contextual cues. Vision is among these (Sumbly and Pollock, 1954). So, as we now know, is the sense of touch: Gick and Derrick (2009) found that inaudible puffs of air applied to a subject's skin caused her to hear phonemes as more aspirated, thus transforming /b-/ into /p-/ , for example. Thus, most conclude, speech perception is multisensory.

Sensory atomists think that whatever the other merits of this finding, it tells us nothing about the intrinsic character of speech sounds. For them, if the experience of hearing /da/ in a McGurk video is the same as that of hearing an environmentally authentic utterance, then the two phonetic ideas are the same. And if they are the same, it must be a mistake to hold that the source of the input makes a difference to phonemes. Yes, /da/ can be produced in different ways, but since the consciously apprehended *idea* is the same, and moreover simple, its modality, understood as an intrinsic experienced feature, cannot be sometimes this, sometimes another. On this reasoning, /da/ is auditory no matter how it is produced — it is experienced in the auditory manner.

One might have thought that this disagreement between scientists and philosophical friends of sensory atomism would be an embarrassment to the latter. But in the scientifically blinkered world of mid-twentieth century linguistic/analytic philosophy, the conflict did not cause any lack of ease. The atomist's conclusion is reiterated by most important, indeed the only, original work on sense-modality in this period — H. P. Grice's (1962) 'Some Remarks on the Senses.' In this influential paper, Grice considers an imaginary creature that possesses two pairs of eyes, each 'more or less like our eyes,' and each used, as our eyes are, to experience colour, size, and shape. Despite these similarities, the creature experiences the same properties differently — when it experiences *blue* through one set of eyes, for example, it has a different experience than when it experiences it through the other. Grice argues that in such a case — and in our own case, by extension — the correct way to demarcate a modality is by how things are experienced — or, as he puts it, 'the way to describe our visual experiences is in terms of how things look to us.' This reinforces the atomist premise — though the creature experiences colour through similar sense-organs, it is experiencing different simple ideas. (In a way, the creature is like Molyneux's newly sighted man, experiencing through one pair of eyes the same properties as it does with the other — but experiencing them differently.) So, the nature of the sense-organ is not relevant to the nature of the quality experienced through it.

Come back now to the McGurk effect. For atomists, speech perception is unimodal-auditory because it is of the same sort as we experience when we *hear* speech. It does not seem to us that we see it or see-*and*-hear it. Ultimately, the quality of experience is what matters. So, since we experience the phonemes as if they are simple auditory qualities, they *are* auditory qualities. And here 'auditory quality' does *not* mean the same as 'quality that the ears are sensitive to.' Rather, it means 'quality that is experienced in the same manner as other qualities labelled auditory.' And that is the end of the matter, at least as far as sensory atomism goes. (Grice actually considers a case structurally somewhat similar to the McGurk effect, in which a perceiver to whom 'two coins look equal in size when only seen, feel equal in size when only felt, but look unequal and feel equal when *both* seen *and* felt.' He does not consider the possibility of cross-modal interaction as a possibility for resolving the case.)

Sensory atomism shuts down the discussion of multisensory perception, and for this reason, the question of the sense-modalities is very little discussed in the modern period in European and English-speaking philosophy. In older times, however, the question was regarded as more interesting. Thus, both in ancient Greece and in India, the five external senses are a staple of the discussion of perception.

Philosophers in ancient Greece violated the methodology of sensory atomism by discussing the causal process of perception, which, according to them

is modality-specific. Aristotle, in particular, offered a detailed account of what he took to be the causal sequence originating in an external object, continuing through an intervening medium, and culminating in a perceptual act. Aristotle held that a sensory quality is one that ‘actualizes’ a sense organ by transmitting its ‘sensory form’ through a specific medium. For example, the colour, orange, is a quality that actualizes the eye’s capacity to sense orange by imparting its ‘sensory form’ to the sensitive part of the eye. It does this by first acting on ‘the transparent’ — a medium that occupies the space between the object and the eye. Without going into details about this process — see Matthen 2019 for a more detailed account and discussion — it is meant to be modality-specific. The five external modalities are distinguished from one another both by the medium specific to them and by the qualities that are able to act on them through the medium.

In Aristotle’s model, however, there is an additional problem that leads to a consideration of the possibility of multimodal perception. Since a sensible object makes the eye like itself, and since the act of perception must ostensibly identify a particular object, it seems impossible to perceive *two* distinct objects simultaneously. For if one object makes the eye orange, and this state of the eye constitutes a representation of *this* object being orange, then the eye is thereby rendered incapable of representing *another* object being any colour — its representational capacity is exhausted by the first object and its colour. Aristotle attempts to address this problem in *de Sensibus* 7 by proposing, as Pavel Gregoric (2007, 133) writes, that acts of perception must be ‘complex enough to grasp both objects,’ yet ‘sufficiently unified to produce a single act of simultaneous perception.’

A similar problem is articulated in Gautama’s Nyaya-sutras (a work that is at least 1900 years old, though it could antedate Aristotle), which defines perception as ‘cognition that arises from the contact of the sense organ and object.’ Gautama, like Aristotle, assumes a pairing of sense-organs and their objects. He does not, however, mention the medium as something that differentiates the modalities. (See Monima Chadda, 2010, for more information on Indian theories of perception.)

These more accommodating conceptions of perception raise interesting questions for which sensory atomists have no room. Consider once again how we know that *burnt sienna* is different from the trumpet’s middle C? Clearly not by either sight or hearing, since neither modality is able to entertain both simultaneously. Accordingly, Plato, Aristotle, and the Nyaya school all posit a common faculty that makes inter-modal comparisons — the Greeks posit a ‘common sense’ and the Indians a mental faculty they call *manas*. Aristotle and the Nyaya scholars then ask whether it is possible for this common faculty simultaneously to entertain ideas from more than one modality. To which, they answer negatively — the common faculty can access only one modality at a

time, or in Aristotle's case, an emergent mixture formed from two modalities (Gregoric, 2007, 134). Interestingly, this line of thought aligns with Spence and Bayne (2014), who argue on much the same grounds that 'consciousness' is not multisensory.

Consider, in the light of these ancient theorists (and also Spence and Bayne, 2014), the ventriloquist's illusion. This performer produces speech sounds with his mouth closed, while simultaneously flapping the mouth of a wooden dummy that he holds by his side. This produces the irresistible illusion of the dummy speaking. The explanation usually given is that our perceptual systems detect a coordination between the visually detected movements of the dummy's mouth and the auditorily detected speech sounds actually produced by the ventriloquist. As a consequence of this coordination, it locates the sound stream in the mouth of the dummy. Call this audiovisual 'binding.' (Perhaps this act of binding would satisfy Aristotle's requirement that simultaneous cross-modal perception demands a mixture.) Now, which system performs the binding act? Not vision, say both the Greeks and the Indians referenced above, because vision cannot hear the speech sounds. Not audition either, for symmetrical reasons. So, there must be a common faculty that does the work. The Nyaya philosophers say that this common faculty does its work by switching rapidly back and forth between visual and auditory experience, both of which it can access, but not simultaneously. Aristotle considers, but rejects this solution, at least partially on the grounds that it does not give an adequate account of musical harmonies. (It is significant that harmony is not systematized in Indian classical music — it's at best the product of voice-leading there.)

We said that these ancient theorists have a more accommodating conception of modality than the sensory atomists. In what way? And what philosophical work does their conception do? To answer the first question, let's ask: what does the sensory atomist say about the ventriloquist illusion? Just that the voice seems to come from the dummy's moving mouth. The manner in which this illusion is achieved is a contingent fact about our experience, and as such it is not the philosopher's concern. This is a narrow-minded conception of the philosopher's task, but it is important to note that this is not where the ancient theorists disagree. For their focus is on the manner in which the coordination between the voice and the moving mouth is achieved — that is, by a faculty that is capable of entertaining both auditory and visual experiences. And this leads them to the second question: What philosophical work is accomplished thereby? By adopting an empirically substantive thesis about the modalities, they arrive at more interesting conceptions of multisensory integration than is available to sensory atomists. Within these conceptions, it is possible to ask questions like: How can we simultaneously be aware of ideas from different modalities? In short, they are led to consider questions that are suppressed in

the classic sensory atomism that dominated philosophy of perception from the seventeenth century right up to the end of the twentieth.

With these examples of conceptual richness in mind, we will briefly outline some other philosophical questions that could be raised about sensory atomism and multisensory perception. It is our contention that however these questions are answered, they are questions that don't naturally arise in the context of orthodox sensory atomism. It is worth unsequestering them, for each leads in its own way to productive questions about the proper analysis of the sensory qualities and capacities involved.

Field awareness vs. point awareness Sensory atomists assume that when a quality seems to be located at a point in the perceptual field, the content of the sensory state is point-specific. Suppose, for instance, I seem to see *red* instantiated in that maple leaf. Then the perceptual content of my visual state is that the maple leaf is red.

Sensory integration, including multisensory integration, suggests an alternative view. Consider the ventriloquist illusion again. Ostensibly, my perceptual state is a purely auditory experience of speech emanating from the dummy's mouth. But here's a plausible analysis of the effect that takes us in a different direction: my perceptual systems settle on a probable common cause for two coordinated effects, one visual and the other auditory by co-locating them in the dummy's mouth. If this analysis is correct, then one may think it plausible that the content of my perceptual state is more complex, involving an integrated amalgam of several inputs from audition and vision.

Coming back to the red of the maple leaf: we know that colour perceived at a point is the result of a comparative analysis of the entire visual field. (This is the basis for colour constancy.) In view of this, is it appropriate to deny that colour vision provides point-located awareness and to hold instead that it offers (or at least that it incorporates) awareness of the colour distribution across the entire visual field? And if this is so, what is the mental faculty that achieves field-integration? Does it have any of the attributes of Aristotelian 'common sense' or Nyaya *manas*?

Temporal Integration We seem to sense qualities *at a moment*. My experience *at this moment* of the world seems to be experience of how the world is *at this moment*. But this view has its difficulties when it comes to movement and change. Movement is defined by some one thing being in different locations at different times. If what I perceive is only how things are at a moment, then how do I perceive at a moment that something is moving? The answer must be that *clocked* sensory information about the past is incorporated into our awareness of the present. So, what appears to be visual information about the present ('NOW: that car is moving toward me at high velocity') is actually multisensory since an internal clock comes into play.

The (much-discussed) question then is this: how is temporally spread-out sensory information compressed and presented in perception of a single moment?

Spatial Integration Through Self-Movement Vision is spatially organized at the input end — the lens of each eye throws a two-dimensional image on the retina. The inputs to audition and smell are *not* spatial: audition separates the input out into tones, each of which sums up spatially distributed input; smell similarly analyses it by chemical properties. Yet, we do seem to be able to locate sounds and smells. How?

One part of the answer lies in the fact that we have *two* ears and *two* nostrils, resulting in small differences in the arrival times of stimuli. This gives information about the direction in which stimuli lie relative to us. Another part of the answer, and perhaps more important, is variation due to self-movement. We move our heads; we move about our surroundings. As we do so, the strength and direction of the signals received from various stimuli changes. The various stimuli can be mapped from these changes. But this requires proprioceptive information about one's own movement and information about timing. So, the important question whether hearing and smell yield spatial mappings turns on how we integrate these inputs.

Cross-modal Binding You hear sounds; you see objects. You also hear sounds as coming from objects. For example, when listening to an orchestral performance, you might hear a flute somewhat off to the left of the ensemble. But then you pick out the flautist and see her head bob as she plays her instrument. Suddenly, the sound of the flute is heard coming from a much more confined location than it did before. It doesn't come merely from the left; it comes from the flautist. (Compare the ventriloquist illusion.)

What is the ontology of this phenomenon? Is there unified audio-visual perception of objects producing sounds?

Emergent Qualities. Sense qualities that have traditionally been treated as simple turn out in many cases to be mixtures that have some degree of phenomenological unity. Colour is a prime example of this. Philosophers in the atomist tradition take colour to be a prime example of a simple idea. Yet it was known even to Aristotle that colour is a composite appearance. Today, it is commonplace to analyse perceived colour into three intensive components that can vary more or less independently — values along blue-yellow, red-green, and black-white dimensions. Burnt sienna is a unitary colour, but it can perceptually (not physically) be resolved into strong components of red and black, together with a weak admixture of yellow. Qualities in other modalities are also analysable, though less systematically. (The orderly three dimensionality of colour is an exception.) Smell, for example, is a unitary quality that

arises cumulatively from multiple olfactory receptors. A single complex odoriferous substance like cheese might excite dozens of receptors. Yet, it has a unified odour; it is one odour, not a collection of odours existing side by side.

Multisensory emergence is a particularly interesting case. Here are two examples.

(i) *Solidity*. When you hold a hard object between your fingers — a fork or knife or pen — it feels solid. How is this feeling of solidity different from what you would feel if outward pressure was *separately* applied to the corresponding parts of your hand? — For in the latter case, you feel disconnected points of pressure, not solidity. The answer is that you feel counter-acting and reciprocal pressure when you hold something — you push down on a pen with your middle and index fingers to steady it against the resistance of the writing surface, and then the thumb must provide compensating force to stop the pen from being pushed down. This interplay between the fingers that grip the pen are responsible for the feeling of solidity — but you are aware of the interplay not by touch but by proprioception. The feeling of solidity, seemingly simple, arises, in short, from the interplay of touch and proprioception.

(ii) *Flavour*. The taste buds in the tongue provide only limited information about flavour. More information is provided by vapours from the oral cavity rising over smell receptors — a sensory process called ‘retronasal’ olfaction to mark the entry to smell from the mouth instead of the nose. Information is also provided by the trigeminal nerves in the face which are affected by irritants like mustard and stimulants like the fizz in soda. A single flavour emerges out of all of these influences. The taste of a fizzy drink arises then from an interplay of taste-buds, olfactory receptors, and pain receptors. But what we experience is unified — it is one flavour. Should we regard this single flavour as multisensory? We are inclined to think that this is a problem even for sensory atomists.

Several of these questions, and many others, are addressed by the papers in this volume. The notion that much of perceptual experience is multisensory is accepted by all parties involved; at the same time, the claim that much of perceptual experience is multisensory assumes a distinction between several individual senses, such that some experiences employ a combination of these senses, while other experiences may be unisensory. Yet the challenge of finding appropriate criteria for individuating the senses has not been resolved. Individuating the senses by phenomenology alone reintroduces the initial problems that plagued the atomists. Phenomenology seems to play *some* essential role in sense individuation, but the question of its more precise role remains. Several papers in this volume address the question of sense individuation, as we will explain below.

Our common notion of *sense* assumes that there are several distinct senses, independent from one another while working in concert with each other. But

the prevalence of multisensory perception in conscious experience threatens this ordinary notion of the senses as distinct and independent and puts pressure on the intuitively defined boundaries of individual sensory modalities. Taking on an issue similar to that addressed by Garzorz and Deroy (this volume), Casey O’Callaghan resolves the tension between our intuitive notion of sense and the phenomenon of multisensory perception with an account of the senses as capacities. ‘What each sense shares and what differentiates one sense from another’, he writes, ‘is the manner in which those capacities are exercised. Each way of perceiving involves a distinct type of information gathering, individuated by the information it functions to extract and the medium from which it does so’.

O’Callaghan’s approach allows us to distinguish between unisensory capacities, multisensory capacities (those that *can* be deployed in a way that relies on more than one sensory manner), and distinctively multisensory capacities. The last of these are capacities that *cannot* be exercised originally by a single sensory manner. Flavor and balance perception are the examples of distinctly multisensory features, which employ distinctively multisensory capacities. The same theory of sense can be used to draw parallel distinctions in the domain of perceptual *episodes*, thus providing a single principled of both the nature of *sense* as well as unisensory and multisensory perceptual phenomena. O’Callaghan’s account also provides us with answers to certain *hard cases* regarding the senses, including how to understand the individuation of visual sense for Grice’s four-eyed creature.

In ‘Why There Is a Vestibular Sense, or How Metacognition Individuates the Senses’, Garzorz and Deroy ask whether the vestibular system should be counted as a sense. The question is similar to O’Callaghan (this volume); Garzorz and Deroy offer criteria for identifying a distinct sense that goes beyond the connection between a sense and the utilization of particular organs. Their proposal for individuating senses adds a metacognitive criterion- they argue that subjects have the capacity to *monitor* information generated by distinct organs, whether that information is unisensory or integrated multisensorily. This capacity, they claim, can be exhibited and measured by performance on specific metacognitive tasks.

Their proposal is worked out in detail for the vestibular system, and taking this as their model, Garzorz and Deroy offer *metacognitive awareness* of information generated by particular sense organs as a novel criterion for sense-individuation in general. The addition of an explicitly metacognitive criterion would establish personal criteria in addition to physiological criteria for individuating senses, in a way that Garzorz and Deroy claim would accurately reflect the contribution that each sense makes to how we experience the world at the subjective level.

In the popular imagination, *synesthesia* may be one of the most iconic forms of multisensory perception. While synesthesia encompasses a diverse array of phenomena, broadly speaking all forms of synesthesia are perceptual or sensory episodes in one modality (or cognitive stream) which have been triggered by stimulation of a different modality (or without sensory stimulation entirely). In ‘Synesthesia as (Multimodal) Mental Imagery’, Bence Nanay defends the claim that *all* synesthesia is a kind of mental imagery, including synesthesia not induced by sensory stimulation. This appears to conflict with several popular accounts of synesthesia- in particular, those that treat synesthesia as *distinct* from mental imagery. Nanay’s proposal avoids this conflict by treating mental imagery as an *especially broad* category, one which includes all early perceptual processing not triggered by corresponding sensory stimulation in the relevant sense modality. This allows mental imagery to be conscious or unconscious, voluntary or non-voluntary, and to have a variety of subjectively experienced qualities.

Treating all instances of synesthesia as forms of mental imagery (so defined) is consistent with recent evidence of artificially induced synesthesia, and illuminates significant similarities between synesthetic ‘projectors’ and ‘associators’. It can account for multimodal and unimodal synesthesia, as well as sensory substitution and synesthetic events that occur in the absence of sensory stimulation whatsoever.

While Nanay’s theory allows for sensory substitution to be interpreted as a form of synesthesia, in ‘Mixing up the Senses: Sensory Substitution Is Not a Form of Artificially Induced Synaesthesia’, Kirsch, Job, and Auvray ask whether *sensory substitution* experiences sufficiently fulfill the criteria for synesthetic episodes, in particular artificial synesthetic experiences. Focusing on criteria such as specific inducer–concurrent pairings and their idiosyncrasy, as well as automaticity and consistency over time, Kirsch et al. significantly narrow the range of experiences that qualify as synesthetic. While Nanay takes synesthetic experiences to be *sensory episodes in one modality which have been triggered by stimulation of a different modality*, Kirsch et al. have more stringent criteria.

Sensory Substitution Devices operate in a variety of ways, and research into the phenomenology of the sensory experiences produced by SSDs has not yet revealed uniform experiences across subjects. In particular, the phenomenology of SSD use reported by sighted and non-sighted subjects is significantly different in each case. Though some SSD subjects report a phenomenology that is consistent with the criteria for synesthesia (per Kirsch et al.), the evidence does not show that the criteria are met in other cases. The dissimilarities between *sensory substitution* and *synesthesia*, in regards to the essential criteria Kirsch et al. propose, make a compelling case that sensory substitution

should not be treated as an artificially induced form of synesthesia, despite promising first appearances.

In ‘Visual Self-Motion Feedback Affects the Sense of Self in Virtual Reality’ Schettler *et al.* investigate the impact of perceived motion in virtual reality avatars on a subjects’ sense of self-identification with those avatars. Evidence from their study suggests that observing corresponding (rather than incongruent) motion in avatars creates a more significant phenomenology of self-identification than several other types of correspondence, in particular congruent representations of biological sex in an avatar.

The ability to interpret an alien body as an extension of one’s self based on visual feedback has been utilized for decades in Mirror-Box therapy for pain in phantom limbs and other forms of chronic regional pain. More recently, researchers have explored whether particular perceived motions in virtual reality avatars will also serve as an effective treatment for these types of pain. The study by Schettler *et al.* examines the extent to which congruence in perceived motion between an individual and an avatar overrides other types of incongruence. They take their results to suggest that the strength of the influence of perceived congruent motion over other self-identification factors (here, identification with biological sex), show that certain bottom-up feedback may override top-down interpretations of self-representation.

In ‘The Role of Attention in Multisensory Integration’ Christopher Mole examines an apparent paradox concerning the relationship between attention and multisensory integration that has puzzled the multisensory research community for the past several years. In the very same experiments, data seem to show that integration occurs prior to and affects attention (when analyzed according to one method), and that attention occurs prior to and affects integration (when analyzed according to a different method). Neither interpretation of the data seems more appropriate than the other, and yet they seem to be at odds with each other.

Mole’s solution to this puzzle involves reconsidering how we interpret the events studied in these experiments; the event of attending to a stimulus and the event of integrating multisensory information. The paradox, according to Mole, arises only if we understand these events as having naturally fixed end-points; as terminating when a selection is made or a button is pushed. Mole suggests that ‘this appearance of paradox can be removed if we can find a way to resist the idea that there must be fixed temporal relations between the instances of these processes.’ Attending and integrating are process that could continue indefinitely, and the point at which we consider them completed is artificially imposed on them- an artifact of the studies themselves. Understanding the different metaphysics of different types of events allows us to resolve this paradox without fundamentally revising our concepts of attention or multisensory integration.

Note

1. Not all philosophers endorse the association between atomic ideas and simple experience. The great positivist, Rudolf Carnap, tried to ‘construct’ simple ideas out of temporally extended, all-inclusive streams of consciousness, or ‘elementary experiences,’ as he called them. The system of his *Aufbau* maps on to that of sensory atomism in interesting ways. But he apparently didn’t take the priority of ‘elementary experiences’ to be ontologically fundamental feature of his system, and in the Preface to the Second Edition, he signals a preference for Mach’s rigidly atomistic system. ‘It is certainly interesting that the restriction to a single basic concept is possible,’ he writes. ‘However, nowadays this procedure appears to me to be too artificial.’ (We are grateful to Jennifer Nagel for helpful discussion of Carnap.)

References

- Chadha, M. (2016). Perceptual experience and concepts in classical Indian philosophy, in: *The Stanford Encyclopedia of Philosophy (Spring 2016)*, E. N. Zalta (Ed.). <https://plato.stanford.edu/archives/spr2016/entries/perception-india/>.
- Gick, B. and Derrick, D. (2009). Aero-tactile integration in speech perception, *Nature* **462**(7272), 502–504.
- Gregoric, P. (2007). *Aristotle on the Common Sense*. Clarendon Press. Oxford.
- Grice, H. P. (1962). Some remarks about the senses, in: *Analytic Philosophy First Series*, R. J. Butler (Ed.). Basil Blackwell, Oxford.
- Johnson, W. E. (1921). *Logic Part I: Propositions and Relations*. Cambridge University Press. Cambridge.
- Matthen, M. (2019). Is the eye like what it Ssees? A Critique of Aristotle on sensing by assimilation, *Vivarium* **57**(3-4), 268–292.
- McGurk, H. and MacDonald, J. (1976). Hearing lips and seeing voices, *Nature* **264**(5588), 746–748.
- Spence, C. and Bayne, T. (2015). Is consciousness multisensory?, in: *Perception and Its Modalities*, D. Stokes, M. Matthen and S. Biggs (Eds), pp. 95–132. Oxford University Press, New York.
- Sumby, W. H. and Pollack, I. (1954). Visual contributions to speech intelligibility in noise, *J. Acoust. Soc. Am.* **26**(2), 212–215.