The Unity of Consciousness and Sensory Integration: Conference Report

By Kevin Connolly, Craig French, David M. Gray, and Adrienne Prettyman

This report highlights and explores five questions which arose from The Unity of Consciousness and Sensory Integration conference at Brown University on November 5th and 6th, 2011.

1. What Is the Relationship Between the Unity of Consciousness and Sensory Integration?

As you pick up your mug and drink from it, you perceive both visual and tactile features of the mug. You see the mug, and many of its visible features, such as its white color, its movement through space as you bring it towards your mouth, and so on. You also feel the smooth texture of the mug's surface as you grasp its handle, and as its rim makes contact with your lips. Your experience is multimodal in at least two respects. First, it has a multimodal phenomenological character—a proper description of what it is like for you to undergo the experience has to make reference to its visual and tactile aspects, that is, to the experienced visual and tactile features. Second, it is an experience which results from processing sensory information gathered from distinct sensory modalities, so it is also multimodal with respect to its causal origin.

In your experience of the mug, the information gathered from different sensory modalities are bound together (they are not merely co-present in an overall experience). You perceive the mug's distinct features as features of a single object (the mug). Your percept of the mug involves not just the co-presence of multimodal features, but their coherence: your experience involves *object unity*.

Such unity can be a feature of phenomenal consciousness. In the case in question, it is manifest in the phenomenological character of your experience: you perceive the mug—that single object—as bearing those various visual and tactile features. An exhaustive account of your perceptual phenomenology needs to mention the unity or coherence of the mug's features. It isn't sufficient to mention or list the features alone (just as it is not sufficient to represent the fact that John loves Jane by listing the components: John, the relation of loving, and Jane; one also needs to mention how those components cohere, or are, so to speak, bound together).

Multimodal perceptions such as your experience of the mug make use of multisensory integration, where that is understood as "the brain's ability to synthesize the information that it derives from two or more senses" (Stein et al., 2002, p. 227). How multisensory integration relates to the unity of consciousness depends upon the approach one takes to the unity of consciousness. One approach is the Hill-Bennett approach (which has its origins in Hill, 1991, Chapter 10). On this view there is no more to the unity of consciousness than whatever unity is derived from the obtaining of substantive unity relations. Substantive unity relations include object unity, which we have already encountered. Another such relation is joint access-consciousness, whereby two experiences are jointly available for use by higher-level cognitive functions such as reasoning or action. A third substantive unity relation is spatial unity, whereby two experiences represent objects or features as being part of the same space. There are other substantive unity relations as well. On the Hill-Bennett approach, the multisensory integration processes for object unity are relevant to the unity of consciousness just insofar as object unity is one of the substantive unity-making relations which (among others) can constitute the unity of consciousness.

Bayne and Chalmers (2003) and Bayne (2010) offer a different approach to the unity of consciousness. On their view, some unity of consciousness outruns whatever unity is provided by the obtaining of substantive unity relations. For example, suppose someone has a visuo-tactile experience of a mug, but also a feeling of elation at the same time. Her consciousness may exhibit, at that time, all sorts of unity (e.g., object, spatial, and so on). But in addition, on the Bayne-Chalmers view, it exhibits phenomenal unity (which is understood in subsumptive terms). That is, there is something it is like for her to have the visual-tactile experience *together* with the feeling of elation. According to Bayne and Chalmers, for such phenomenal unity, those separate experiences must be subsumed parts of a total conscious state.

On the Bayne-Chalmers view, multisensory integration is, on the face of it, less directly relevant to our understanding of the more general category of the unity of consciousness (where that includes subsumptive phenomenal unity) than it is on the Hill-Bennett view. Still, if one wants to understand total conscious states in terms of units of consciousness which are constitutively independent of the wholes which they comprise, then it may be that the best understanding of the basic units of consciousness is a multisensory understanding. In that case, multisensory integration will be relevant to a full understanding of total conscious states, in terms of what gets subsumed, even if not to the relation of subsumption. But to see whether this is the case, we first need to determine the basic units of consciousness.

2. Are Some of the Basic Units of Consciousness Multimodal?

In the *McGurk effect*, a subject views a video of a person saying one set of syllables (e.g. *ga-ga*), while the audio has been redubbed to a second set of syllables (e.g., *ba-ba*). The subject

experiences yet a third set of syllables, distinct from the first two sets (e.g., *da-da*) (McGurk and MacDonald, 1976, p. 747).

In his talk on how to model the unity of consciousness, Tim Bayne proposed two different interpretations of crossmodal cases such as the McGurk effect. On a strictly causal interpretation, seeing the person mouth *ga-ga* causes you to hear *da-da* instead of *ba-ba*. According to this interpretation, integration occurs between processing in the two different systems (the auditory system and the visual system), but the result of that processing can be fully decomposed into an audio component and a visual component. So, while the processing is multisensory, the result is not intrinsically multisensory. On a *constitutive* interpretation, on the other hand, the *ga-ga* visual input and *ba-ba* auditory input give you an experience that is constitutively both audio and visual (not just a conjunction of an audio and visual experience). According to this interpretation, the perceptual state that results from the processing cannot be fully decomposed into two unisensory token states, one auditory state and one visual.

We can interpret other crossmodal cases constitutively or causally as well. In the motion-bounce illusion, subjects look at a computer display of two disks moving steadily towards each other until they meet. If the subject hears a sound at or around the point of convergence, the disks typically appear to collide and bounce off one another. If the subject does not hear a sound, the disks appear to cross through one another (Sekuler et al., 1997). According to a strictly causal interpretation, the motion-bounce illusion is a case where the sound simply causes you to have a certain visual experience (given the right visual input). According to a constitutive interpretation, on the other hand, it is a case where you have a constitutively audio-visual experience.

In the conference's first panel discussion, Susanna Siegel argued that whether we take a constitutive or causal interpretation of crossmodal cases will determine whether we hold that

some basic units of consciousness are fundamentally multimodal. If we hold a constitutive interpretation of the McGurk effect, for instance, then we hold that at least some of the basic units are audio-visual. A strictly causal interpretation, on the other hand, does not commit us to that.

Siegel mentioned (and Fiona Macpherson further bolstered) one piece of evidence in support of holding a causal interpretation of the McGurk effect. When you see someone mouthing the syllables *ga-ga*, the auditory input is *ba-ba*, yet you hear *da-da*. But you could always hear the syllables *da-da* and you could always see someone mouthing the syllables *ga-ga*. So it seems like there is not some new type of unit created by the McGurk effect. Rather, it seems like the results of the McGurk effect are decomposable into an auditory unit and a visual unit. After all, you could experience each of those types on their own, outside of the McGurk scenario (for a similar argument, see Connolly, 2011, pp. 126-128).

Siegel went on to suggest that whether we hold a causal or constitutive interpretation of cross-modal cases determines how we analyze the unity of consciousness. If we ask what conscious unity unifies, the answer to that question will depend on what the basic units of consciousness are, since they will be the most basic relata of the unity relations. Siegel pointed out that there are two options here: either all of the units of consciousness are unimodal, or else some of the units are not. If the basic units are strictly unimodal, then the building blocks unified in conscious experiences will be unimodal. Most philosophers just take it for granted that the most basic units of consciousness are unimodal. However, if at least some basic units are fundamentally multimodal, then at least some of the building blocks of unified conscious experiences will be multimodal.

3. How Should We Model the Unity of Consciousness?

Once we determine the building blocks of consciousness, we still need to determine how they are assembled. Most philosophers think that, at least in normal cases, the experiences of a single subject at a single time are phenomenally unified. As a first pass, two experiences of a subject are phenomenally unified just in case there is something it is like for that subject to experience them *together*. Much of the discussion at the conference centered on how to understand this relation.

On Tim Bayne's view, as noted above, phenomenal unity is understood in terms of *subsumption*: two experiences are phenomenally unified just in case they are subsumed by the same encompassing conscious state, where subsumption is an irreducibly phenomenal relation (2010, p. 16). During the conference panels and Q&A sessions, participants put forward various alternatives. First, along with Bennett and Hill, we might take phenomenal unity to be reducible without residue to the obtaining of one or more other "unity-making relations" among experiences, such as object unity or spatial unity (see section one). Second, we might understand phenomenal unity in terms of mere conjunction. Suppose, for example, that you simultaneously have a conscious experience of a sunset and a conscious pain. Alex Byrne suggested that in such a case there is no more to two experiences being phenomenally unified than that I experience them at the same time. Third, we could give some or other representationalist account of phenomenal unity. David Chalmers mentioned one possibility: two experiences are phenomenally unified just in case they represent their intentional objects as belonging to one and the same world. He added that the best representationalist view of phenomenal unity holds that such unity involves mere conjunction. You hear someone speaking and see a cup, for instance, and conjunction is responsible for the phenomenal unity of your world model.

Note that a conjunctive view of phenomenal unity would rule out the possibility that a single subject has multiple streams of consciousness. In her discussion of split-brain subjects, Elizabeth Schecter argued for a view on which subjects can have multiple streams of consciousness (see Schecter, 2010). On her view, some distinct conscious experiences in a subject are not phenomenally unified. For example, the same subject could consciously experience a mug in the left visual field, and a pencil in the right field, while having nothing it's like to experience the mug and the pencil together. If this is a coherent possibility, it suffices to show that there must be something more to phenomenal unity than a conjunction of conscious experiences.

4. Is the Mechanism of Sensory Integration Spatio-Temporal?

So far, we have discussed various ways to model the unity of consciousness, but what about ways to model sensory integration? Consider a case of sensory integration in a single modality like vision. When presented with a visual array containing a red circle and a green square, one popular view is that the visual system binds the feature *red* to the circle (and not the square) in part because that feature and that shape are located in the same space at the same time. Likewise, co-location in space or time may explain how we integrate features into the same object or event across modalities. Space and time could potentially provide an amodal framework shared across the senses (a view suggested by Farid Masrour), and so would be well-poised to explain sensory integration. David Chalmers noted that several discussions from our meeting pointed to space and time as the "glue" of multi-sensory integration. Less metaphorically, they suggest that the mechanisms of sensory integration are universally spatial or temporal. Call this the Space-Time hypothesis (ST).

Whether or not ST turns out to be true, spatial and temporal mechanisms certainly play a significant role in sensory integration. David Eagleman gave one example of time's crucial role in sensory integration. He argued that we create multi-sensory events by calibrating our expectations about how synchronous stimuli are perceived across sensory systems (see Eagleman, 2008, p. 133). Furthermore, Pawan Sinha presented some evidence for the importance of spatial perception, particularly perception of motion, for integrating visual features into objects (and that account might be extendable to multimodal cases as well). Mohan Matthen went on to note that an *observer's* motion in space can also help to integrate features into objects, as when an observer walks around to locate a smell or sound source.

Some challenges arise for ST, however, when we consider objects of perception that are not obviously spatially located. Sounds, smells, and tastes are arguably aspatial (for a discussion of this, see O'Callaghan, 2011, pp. 147-49). If this turns out to be right, then the integration of auditory, olfactory and gustatory features into multisensory objects are probably not explicable in terms of a spatial mechanism. Berit Brogaard and Carolyn Dicey Jennings raised similar challenges for vision. In association synaesthesia, for instance, subjects report a strong association between, e.g., a number and a color, but the color is not spatially located. Other suggested examples of visual objects that are not spatially located include an undifferentiated visual ganzfeld, or a free-floating color flash. In response to these examples, proponents of ST argued that the latter two visual features are located in space and time: the many points of a ganzfeld are each localizable within the visual field, and likewise with the free-floating flash. Furthermore, association synaesthetes arguably do not have a sensory experience of color, as evidenced by the fact that they do not show a pop-out effect in visual search.

Notably, each of the proposed challenges to ST involves a sensory experience of an object that is not spatially localized. No one, however, proposed an example of a sensory experience that does not occur for a perceived duration of time. The absence of counterexamples may reflect a closer connection between sensory experience and perceived duration, as opposed to spatial location. Even if some sensory objects are not located in space, perhaps time provides the glue of multisensory integration.

5. How Should We Study Experience, Given Unity Relations?

In the conference's final panel discussion, Casey O'Callaghan outlined one major idea motivating the traditional approach to studying experience. According to proponents of the traditional approach, we can tell a story about vision (in a bottom-up way) which does not appeal to processes in any other sensory modality (see, for example, work by David Marr). According to this view, we can give an account of vision by paying attention to what happens when you stimulate the eyes. The visual system, on this account, is informationally encapsulated with respect to the other sensory systems, and the same can be said for each of the other sense modalities. The claim then is that our story about perceiving is just the sum of the stories for each individual sense modality.

O'Callaghan went on to argue that crossmodal cases seem to suggest that this traditional approach to perception fails. Such cases seem to show that perceptual systems are not informationally encapsulated with respect to one another, and that in order to fully understand each sense modality, we need to understand how that modality relates to other sense modalities.

In the ensuing discussion, Ned Block pointed out that there is still a question as to how much we need to modify the traditional approach. Even if we acknowledge (as perhaps all vision

scientists do) that in order to understand visual processes, we need to know all sorts of things about the body (including about other sensory systems), it could still be that multisensory phenomena exist at the margins. This is not to deny that instances of multisensory phenomena are pervasive. It is to say that an account of multisensory phenomena might not be essential for explaining how individual perceptual systems work. If this is right, then perhaps the best way to study multisensory phenomena is to examine sensory systems by themselves, understanding how they work in isolation, and then to merge that account with an explanation of multisensory phenomena. Such an approach, Block pointed out, has worked extremely well so far.

But even if an account of multi-sensory phenomena is not essential for explaining each sense modality (and that remains to be seen), it still seems right that an exhaustive explanation of each sense modality needs to take into account multi-modal phenomena. Plausibly, cross-modal cases show that we can understand a part of experience fully (say, the visual contribution to an experience) only if we understand how that part relates to a greater whole (say, a total auditory-visual experience). This marks a departure from the traditional approach, according to which we can understand each sense modality in isolation from the others. It also raises the following question: how should we best study such phenomena in which an understanding of the parts requires an understanding of the whole? This question is even more pressing if it turns out that an account of multi-sensory phenomena is essential for explaining each sense modality. If that is the case, then it is essential for us to understand the whole in order to understand its parts.

The same sorts of considerations apply beyond cross-modal cases. Cases of synchronic phenomenal unity can be taken to show that in order to understand the parts of a multi-modal experience, we need to understand how those parts contribute to a single synchronic phenomenal unity. Arguably, the same sort of claim applies to at least some cases of diachronic phenomenal

unity. Take the case of listening to a melody (see Shoemaker, 2001, p. 65). One view is that we need to understand the diachronic phenomenal unity in such a case in order to understand the synchronic unity of each moment in an experience of a melody. But if this is right, then it goes against the natural inclination to study a phenomenon by breaking it up into its component parts, and then analyzing each of them. But then for such phenomena, what sort of a method should we use to study them?

References

Bayne, Tim (2010). *The Unity of Consciousness*. Oxford: Oxford University Press.

Bayne, Tim and Chalmers, David J. (2003). "What is the Unity of Consciousness?" In *The Unity of Consciousness: Binding, Integration, Dissociation*. Ed. by A. Cleeremans. Oxford: Oxford University Press.

Connolly, Kevin (2011). *The Role of Concepts in Perception.* PhD Dissertation, University of Toronto.

Eagleman, D. M. (2008). "Human Time Perception and Its Illusions." [Research Support, N.I.H., Extramural Review]. *Current Opinion in Neurobiology, 18*(2), 131-136. doi: 10.1016/j.conb.2008.06.002

Ernst, Marc O. and Heinrich H. Bülthoff (Apr. 2004). "Merging the senses into a robust percept." In: *Trends in Cognitive Sciences* 8.4, pp. 162–169. ISSN: 1364-6613. DOI: 10.1016/j.tics.2004.02.002. URL: http://dx.doi.org/10.1016/j.tics.2004.02.002.

Hill, Christopher (1991). *Sensations: A Defense of Materialism*. Cambridge: Cambridge University Press.

McGurk, H. and MacDonald, J. (1976). "Hearing Lips and Seeing Voices." *Nature*, 264, 746-748.

O'Callaghan, Casey. (2011). "Lessons from beyond Sound and Vision." *Philosophical Studies,* 153(1), 143-160.

Schecter, E. (2010). "Two Unities of Consciousness." *European Journal of Philosophy*. doi: 10.1111/j.1468-0378.2010.00439.x

Sekuler, R., Sekuler, A. B., and Lau, R. (1997). "Sound Alters Visual Motion Perception." *Nature*, 385, 308.

Shoemaker, Sydney (2001). "Consciousness and Co-consciousness." In *The Unity of Consciousness: Binding, Integration, Dissociation*. Ed. by A. Cleeremans. Oxford: Oxford University Press.

Stein, Barry E. et al. (2002). "Multisensory Integration". In: *Encyclopedia of the Human Brain*. Ed. by Editor-in-Chief: V. S. Ramachandran. New York: Academic Press, pp. 227–241. ISBN: 978-0-12-227210-3. DOI: 10.1016/B0-12-227210-2/00225-9. URL: http://www.sciencedirect.com/science/article/pii/ B0122272102002259.