Article

A Defense of an Amodal Number System

Abel Wajnerman Paz [iD]

Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires C1425FQB, Argentina;
awajnerman@filo.uba.ar; Tel.: +054-11-39646630 or +056-9-89556222

Abstract: It has been argued that the approximate number system (ANS) constitutes a problem for the grounded approach to cognition because it implies that some conceptual tasks are performed by non-perceptual systems. The ANS is considered non-perceptual mainly because it processes stimuli from different modalities. Jones (2015) has recently argued that this system has many features (such as being modular) which are characteristic of sensory systems. Additionally, he affirms that traditional sensory systems also process inputs from different modalities. This suggests that the ANS is a perceptual system and therefore it is not problematic for the grounded view. In this paper, I defend the amodal approach to the ANS against these two arguments. In the first place, perceptual systems do not possess the properties attributed to the ANS and therefore these properties do not imply that the ANS is perceptual. In the second place, I will propose that a sensory system only needs to be dedicated to process modality-specific information, which is consistent with responding to inputs from different modalities. I argue that the cross-modal responses exhibited by traditional sensory systems are consistent with modality-specific information whereas some responses exhibited by the ANS are not.

Keywords: perceptual system; grounded cognition; concept empiricism; number representation

1. Introduction

Empiricist or grounded approaches to cognition are a family of theories which share the thesis that cognition depends on components of modality-specific systems for perception, action and emotion (e.g., [1–3]). Grounded theorists often assume that structures underlying conceptual processing are perceptual because they carry modality-specific information (e.g., [3–5]). For instance, in a task that requires thinking about dogs, grounded views would predict that the representations activated in the brain will carry auditory information about dog barks, visual information about dog shapes and/or olfactory information about dog scent, etc.

At the neural level, representations that carry modality-specific information are often identified by the fact that they track a given property only when it is presented under a specific modality. For instance, a visual dog representation is a structure that responds only or maximally to visually presented dogs. Structures of this kind are often found in visual areas dedicated to recognition tasks (visual face, object or action recognition, see [6] for a review on visual object recognition). This view of perceptual representations lines up with a popular characterization of perceptual systems proposed by Prinz according to which these are dedicated input systems [7]. That is, sensory systems process information received through a specific (proximal) input-type (such as wavelengths of light, frequency of molecular motion or molecular shapes). This implies that, for instance, if the visual system processes a representation that carries a piece of information I, then I was received through patterns of retinal stimulation produced by wavelengths of light.

Some authors claim that there is evidence indicating the presence of neural representations that do not carry modality-specific information and therefore are not modal (e.g., [4,8,9]). The relevant

studies show brain structures that exhibit cross-modal responses to a given property. Many of these studies focus on the ability of estimating the approximate numerosity of a given stimulus. Neural representations manipulated by the approximate number system (ANS) can respond to a given numerosity presented, for instance, in either visual or auditory format. This implies that these structures represent numerosity without carrying modality-specific information and therefore the approximate number system (or ANS) is a non-perceptual system. If this system underlies conceptual number processing, as current evidence suggests, then the grounded view is wrong at least regarding this conceptual capacity.

In this paper I will examine an argument against the amodal approach to the ANS advanced by Jones [10]. His proposal is two-fold. In the first place, he shows that the ANS exhibits features which seem to be characteristic of sensory modalities. In the second place, he claims that the property which is apparently inconsistent with sensory systems (namely, the ability to respond to inputs from different modalities) is not problematic because it is also exhibited by paradigmatic sensory systems. In Section 2 I will present different sources of evidence for the amodal approach, show that these have different degrees of strength and determine which kind of evidence is more relevant. In Section 3 I will develop Jones' first consideration and argue that none of the features of the ANS that he mentions is possessed by a paradigmatic perceptual system (i.e., the visual system) and therefore it is doubtful that these are characteristic (and therefore indicative) of sensory systems. In Section 4 I will propose that a sensory system only needs to be dedicated to process modality-specific information, which is consistent with responding to inputs from different modalities. I will argue the cross-modal responses exhibited by traditional sensory systems are consistent with modality-specific information whereas some responses exhibited by the ANS are not. The ANS is not perceptual because it is not dedicated to process (i.e., it does not process exclusively) modality-specific information.

2. The Cross-Modal Response Objection

Machery, Dove and others review different kinds of studies showing cross-modal responses to numerical stimuli [8,9,11]. In order to understand the role that these studies play in the present discussion one should notice that they have different implications. In this section, I will argue that they have different degrees of strength, depending on which grounded hypotheses they can rule out. In what follows I will present them in increasing order of strength and indicate which kind of study could be considered the main evidence against the grounded view.

In the first place, a behavioral study by Dehaene and colleagues determines whether a behavioral response to a numerosity-related task generalizes across significant variations in non-numerical physical parameters [12]. The experiments designed following this strategy show the same behavioral response to the same task in the face of variation in object features (such as size, color or shape), spatial location, modality (auditory or visual) and mode of presentation (simultaneous or sequential) [12] (p. 356).

This kind of evidence is problematic because the presence of switching costs in conceptual tasks involving different modalities has been used to support the grounded view (e.g., [13–15]). For instance, Pecher, Zeelenberg and Barsalou have shown switching effects during a verification task, in which participants are asked whether or not a particular property is true about a given category (e.g., CAT–has a head). They examined pairs of trials which were either from the same modality (LEAVES–rustling followed by BLENDER–loud) or different modalities (CRANBERRIES–tart followed by BLENDER–loud). They found longer reaction times for the second trial in a pair of different modality (switch) trials than for the second trial in a pair of same modality (no-switch) trials [13]. This supports the grounded view because the greater response delay in the switch second trial can be explained assuming that different modality-specific systems are active in each switch trial. In turn, performance in the second no-switch trial is facilitated because the same modality-specific system was already activated during the first no-switch trial. Following the same logic, the absence of increased response delay in switch trials during a number-related task implies that a single amodal system

is employed in different trials. Therefore, conceptual processing can be considered amodal at least regarding this particular domain. In order to examine the strength of this study it is important to bear in mind that, as the examples show, facilitation only requires activating the same system in both trials and not the same specific representations.

There is more direct evidence for the presence of a single number system that responds to numerical stimuli from different modalities. For instance, Piazza and colleagues show that a neural structure partially located in the right and left intraparietal sulcus is involved in estimation of numerosity and is independent of the stimuli's perceptual modality [16]. This study uses fMRI to investigate and directly compare brain responses to a numerosity estimation task and to an exact counting task, using both visual and auditory stimuli. They first segregate the functional structures involved in estimation and counting and then show that these brain structures do not require stimuli from a specific modality, since they perform these tasks using both visual and auditory stimuli.

However, I find the kind of evidence provided both by Dehaene and colleagues [12] and Piazza and colleagues [16] insufficient to rule out a relevant grounded hypothesis. These studies only show that the neural structure which underlies number-related abilities is not a modality-specific system (i.e., it is activated in response to stimuli from different modalities). This is consistent with the possibility that the specific representations recruited by this system are modality-specific. For instance, Barsalou claims that the perceptual representations which (according to his proposal) are required for conceptual tasks are processed (at least in part) by non-perceptual systems, such as association areas [3]. It could be the case that although the same system is activated in a pair of switch trials (and therefore the response delay is not increased in the second trial), the system employs different modality-specific representations in each trial. Recall that facilitation only requires activating the same system in both trials and not the same representations.

In order to rule out this possibility it is necessary to focus on the specific representations recruited for the relevant tasks. There is a wide variety of studies that determine the information carried by the cross-modal response of a neural representation by analyzing fMRI signals. In a recent study, Damarla and colleagues apply multivariate pattern analysis (MVPA) to fMRI signals in order to provide evidence for the presence of amodal representations of quantities of objects [17]. They show that a classifier trained to decode fMRI patterns evoked in frontal and parietal regions by visually displayed quantities can decode neural patterns evoked in the same regions by quantities of auditory tones, and vice-versa. The same neural structure responds with the same activity pattern to the same quantity presented either in visual or auditory format. This suggests that not only the system involved in number processing but also the specific representations are not modality-specific.

Nevertheless, there is a possibility that fMRI studies are not able to rule out. Given the temporal and spatial limitations of the blood-oxygen-level dependent (BOLD) signal, fMRI cannot detect the informational properties of individual neurons. This implies that fMRI is not fine-grained enough to exclude the possibility that a neural structure which responds, for instance, to visual and auditory stimuli with the same is constituted by closely placed unimodal (some visual and some auditory) neurons. To see this problem in more detail, it is useful to briefly consider some studies based on single-cell recordings.

Andreas Nieder has developed different studies focused on 'number neurons' [18]. A number neuron shows a maximum discharge rate to its preferred numerosity. When the presented number becomes more distant from the preferred one, the neuronal activity of single neurons progressively drops off, thus forming a peak-tuned response curve. That is, they have a bell-shaped (Gaussian) response function. Single-cell recordings performed by Nieder indicate that some individual neurons in the primate frontal and parietal association cortex respond to the same numerosity across different modalities [19]. However, these areas not only employ amodal representations but also rely on sets of distributed unimodal detectors. This means that the relevant areas actually employ both amodal representations and sets of intermingled modality-specific number neurons. fMRI data cannot be used

to distinguish between these two kinds of representations. Single-cell recordings provide the strongest confirmation that the ANS employs both modality-specific and amodal number representations.

Jones argues that perceptual systems can have cross-modal responses (even at the cellular level) [10]. I will discuss this idea in Section 4. However, even if his argument is successful, it would only determine that cross-modal responses do not imply that the ANS is not perceptual. It would not establish that the ANS is perceptual. In order to show this, Jones claims that the ANS exhibits many properties that are characteristic of perceptual systems [10]. In the next section I will examine this claim.

3. Typical Perceptual Features

Jones argues that there are different aspects of the ANS which suggest that this can be considered a perceptual system [10]. In this section I will argue that none of these features are possessed by a paradigmatic perceptual faculty (i.e., vision) and therefore it is doubtful that they can be considered hallmarks of perceptual systems. This means that they cannot be used to identify the ANS as perceptual.

In the first place, Jones claims that apprehension of numerosity is fast, direct, and automatic. I will consider speed and automaticity later, together with other features typically related to mental modules. Burr and Ross show that the representation of numerosity is direct in the sense that it does not depend on the representation of other perceptual properties. They support this idea by showing that adaptation to numerosity is independent of adaptation to other perceptual properties [20,21]. This independence suggests that numerosity representation does not require a previous stage related to some aspect of visual texture, such as density, orientation, shape, size or contrast of the stimulus.

However, even if number representation is direct, directedness is not a hallmark of perceptual representations and therefore it cannot be used to argue that the ANS is perceptual. Many (perhaps most) typical perceptual representations are not direct. Hierarchical models of perceptual processing (a family of models based on extensive neurophysiological data from perceptual areas) propose that at each stage of perceptual processing, the representation of a given feature is achieved by combining the representations of simpler features at earlier processing stages. The classic case is Hubel and Wiesel's proposal that the representation of an oriented bar in V1 is achieved by combining the representations of dots of light found in the lateral geniculate nucleus (LGN) [22]. In turn, cells in V4 respond to parts of borders by combining the representations of bars with different orientations and locations. Also, some perceptual representations are indirect without being mere combinations of simpler features. For instance, visual depth is detected by using (the representations of) phase and position of inputs from each eye. This is achieved by calculating the difference in phase and position between these monocular inputs (i.e., binocular disparity, which is inversely proportional to depth) [23]. This suggests that the representation of perceptual properties very often is achieved through the representation of different properties.

Jones also mentions variance of accuracy according to Weber–Fechner laws as a typical feature of perceptual systems. The law states that subjective magnitude varies logarithmically with respect to stimulus magnitude [20,21]. Accuracy in the ANS decreases logarithmically (roughly in line with the Weber–Fechner law) as the numerosity of stimuli increases. These variations are responsible for so-called distance and magnitude effects. That is, the reliability of numerosity discrimination decreases as the compared quantities become bigger or as the distance between them becomes smaller. The problem with this is that the fact that both the ANS and perceptual systems exhibit this behavior is consistent with two possibilities: It could suggest that the ANS is perceptual but it also consistent with the possibility that the ANS is amodal and therefore the Weber–Fechner law is a general principle for

the (modal and amodal) representation of magnitudes[1]. In this sense, it would be begging the question to say that the ANS is perceptual because it exhibits a Weber–Fechner behavior. The same fact can be used to argue that the law is not unique to perceptual representations but it rather applies both to perceptual and non-perceptual representations of magnitudes. In order to rule out this later possibility one would need an independent criterion for assessing whether the ANS is perceptual.

In the third place, Jones argues that the ANS satisfies all of Fodor's criteria for characterizing modular systems [24] (pp. 47–101). I will not challenge Fodor's characterization of modular systems (see [25]) nor argue that these criteria are also satisfied by non-perceptual systems (e.g., [26]). I will pursue a different line of argument. I will claim that there are good reasons to doubt that perceptual systems are modular. Specifically, I will show that a paradigmatic perceptual system, namely the visual system, does not satisfy any of Fodor's conditions.

A first criterion for modular systems is that these are fast. When we say that a modular system S_1 is faster than a non-modular system S_2, we simply mean that if both S_1 and S_2 receive their respective inputs, I_1 and I_2, at a time t_1, then S_2 would produce its output o_2 at time t_m whereas S_1 would produce its output o_1 at an earlier time t_{m-n}. However, it could be the case that although S_1 is faster than S_2 in this sense, a given external stimulus s systematically causes o_2 before causing o_1 (i.e., o_2 has a shorter response latency than o_1) because the pathway from s to o_2 is shorter than the pathway from s to o_1. For instance, this could be the case if S_1 uses o_2 as its input (that is, $o_2 = I_1$). In this situation S_1 would operate only after S_2 delivers its output.

The evidence mentioned by Jones (2015) does not support that the ANS is fast but rather only that it delivers its output earlier than non-perceptual areas. Piazza et al. (2004) mention the fact that the parietal neurons have a shorter firing latency than the frontal ones (about 100 versus 160 ms) [27]. Nieder and Miller point out that this difference in firing latency suggests that visual quantity may be represented first in the parietal cortex and then sent to the PFC where the representation is expanded (ia greater number of neurons convey numerosity information) and held online (i.e., in working memory) to gain control over thought and action [28]. The experiment by Piazza and colleagues confirms this idea because it did not involve any explicit working-memory demands and, as a result, significant number related activation was not found in prefrontal cortex, but only in the intraparietal region [27]. This means that the ANS is "faster" than the PFC only in the sense that it has a shorter firing latency.

I will not criticize Jones on the basis that the ANS is not fast in the same sense as modules are. Although shorter firing latency is not a hallmark of modular systems, perhaps it is a hallmark of perceptual systems and therefore an indicator that the ANS is perceptual. Given that non-perceptual systems often employ as inputs the outputs of perceptual systems, perhaps neurons in perceptual systems always have shorter response latencies than non-perceptual neurons. However, perceptual systems are not always fast in this sense. Perceptual latencies are longer when they are controlled by top-down pathways. For instance, Andreau and Funahashi showed that when visual information was provided through bottom-up pathways, the mean latency in pre-frontal cortex neurons (PFC, a non-perceptual area involved in higher cognition) was longer (144 ms) than the mean latency (73 ms) in inferotemporal cortex neurons (ITC, a visual area) [29]. However, when visual information was provided through top-down pathways during an associative memory task, latency of ITC neurons (178 ms) was longer than that of PFC neurons (see also [30]). That is, visual responses were slower than non-visual responses in tasks involving top-down control. Shorter relative latencies are not a hallmark of perceptual responses.

A second characteristic related to modules is that they are ontogenetically determined: they develop in a predictable way in all healthy individuals through maturation, rather than learning and

---

[1] I am indebted to an anonymous referee for the idea that the Weber–Fechner law is plausibly a general principle for representing magnitudes.

experience. Although the development of modules is *triggered* by specific environmental events, their maturation is driven intrinsically by the organism. As Fodor puts it: "neural mechanisms sub serving input analysis develop according to specific, *endogenously determined* patterns under the impact of environmental releasers" (my emphasis) [24] (p. 100). In contrast, after an initial environmental input or releaser is presented, a non-modular system can be constantly readjusted in response to novel experiences caused by environmental shifts. It involves learning in this sense. Jones believes that ontogenetical determination be used to identify perceptual systems and argues that the ANS has this property [10]. The presence of a homologous system in macaques and the presence of similar capacities in infants suggests that the development of the ANS is ontogenetically determined [31]. However, the visual system involves learning mechanisms in the specified non-modular sense. Feedback connections in the ventral stream are involved in a relevant learning mechanism on which visual recognition depends. Hierarchical generative models propose that feedback projections carry top-down hypotheses in order to compare them with bottom-up sensory information. It has been proposed that these hypotheses are constantly corrected and refined by means of Bayesian learning rules (e.g., [32]). If, following Fodor, we understand the ontogenetical determination of mental modules as the idea that their representations are not constantly updated by experience then it is doubtful that sensory systems are ontogenetically determined.

One could argue that this indicates precisely that the visual system is not a single system but a set of different domain-specific systems. Jones suggest an idea along these lines [10] (p. 10). However, there is a reason why we characterize the visual system as a single system with multiple domains rather than as a set of disconnected systems. All of the mentioned ventral sub-systems dedicated to different properties are visual because they can only identify those properties through patterns of retinal stimulation, that is, they share a common input. Moreover, it is plausible that visual inputs impose similar constraints on different visual sub-systems, given that they implement very similar computations [23].

Additionally, characterizing a given sub-system in terms of its specific proximal input is crucial in order to distinguish it from a sub-system that detects the same property through a different input type. For instance, the sub-system in the ventral stream dedicated to shape detection can identify shapes only through patterns of retinal stimulation. Despite some common or shared circuits, this sub-system is different from the one dedicated to haptic shape recognition, which identifies shape only through touch receptors [34]. Moreover, it is known that when haptic and visual information are both employed in object recognition, they interact in specific ways in order to optimize performance (Ernst and Banks 2002). The idea of such interaction only makes sense if haptic and visual processing are distinguished[2]. These considerations suggest that it is reasonable to characterize the visual system as a single system with different sub-systems dedicated to different domains.

A third idea is that modules are mandatory or automatic. It is considered that numerical representation is automatic because subjects are susceptible to numerical priming effects [35]. A weak reading of automaticity as a criterion for perceptual systems could be that these are systems that involve some automatic processes together with controlled processes. It is clear that this condition is too permissive. Even non-perceptual abilities such as language processing involve a combination of

---

[2] I thank an anonymous referee for mentioning this fact about visual-haptic integration.

automatic and controlled processes. As Prinz points out, we form syntactic trees automatically, but sentence production can be controlled by deliberation [25]. A strong reading of this condition could be that perceptual systems only engage in automatic processing. The problem with this is that it is incompatible with a core assumption of concept empiricism. The reuse hypothesis is precisely the idea that we can acquire endogenous control of perceptual representations and activate them in absence of external stimuli in order to employ them in higher cognitive tasks (e.g., [3]). Empiricism implies that conceptual processing requires some degree of control over perceptual systems.

Four additional properties of modules are that they are localized, subject to characteristic breakdowns, inaccessible and encapsulated. This means that they are implemented by unique neural structures that can be selectively impaired and that these structures "don't let much information out and they don't let much information in" [25]. The ANS is often identified with a structure within the horizontal intraparietal sulcus (hIPS) and it can be selectively impaired by damaging this region [36]. Jones also points out that the ANS is informationally encapsulated because adults make similar performance errors to infants and animals [10]. It is also inaccessible, since we cannot use introspection to understand how it works.

An argument proposed by Jones (which I will examine in the next section) to show that cross-modal responses are consistent with perceptual systems is that these are multimodal systems [10]. The problem is that if this view is correct (and current neurocognitive evidence suggests that it is) then the mentioned four properties are not possessed by the visual system. Recent studies suggest that multimodal integration is a pervasive aspect of sensory processing which begins at its earliest stages. In the next section, I will argue that sensory systems are multimodal in a way in which the ANS is not and therefore multimodal sensory processing is consistent with the idea that the ANS is amodal. Now I will assess the implications of this characterization of sensory processing for the relation between perceptual systems and the mentioned four conditions.

We can consider first encapsulation and accessibility. The mentioned evidence suggests that the visual system has inputs from other sensory modalities and that other sensory modalities have visual inputs. For instance, Zhou and Fuster show that neurons in the somatosensory cortex have access to visual information [37]. In turn, parts of the early visual cortex respond to auditory and tactile, as well as visual, stimuli [38]. This implies that visual information is processed in non-visual systems and that non-visual information is processed inside the visual system. Therefore, the visual system is neither inaccessible nor encapsulated.

We now can examine localization and characteristic breakdowns. As Cappe and colleagues point out, the mentioned communication between different sensory systems is the result of strong neural connections between them [39]. For instance, there have been found direct projections from the auditory cortex to the primary visual cortex V1, as well as projections from extrastriate visual area V2 to the auditory area TPT (temporoparietal temporal area) in the superior temporal gyrus [40,41]. This implies that, unlike information processing within modules, visual processing is not localized but highly distributed. This also means that visual processing cannot be selectively impaired but rather damage in visual areas can affect information processing in other sensory areas that depend on visual inputs.

Finally, Fodor's shallowness condition depends on informational encapsulation. The thesis that modules are shallow comes down to the idea that they are confined, in virtue of their informational encapsulation, to encoding properties which can be inferred from the properties that their specific transducers detect. For instance, the visual system can only represent what can be reliably inferred from shape, color, local motion, etc. The shallowness of modular outputs is determined by the representational limits imposed by informational encapsulation [24] (p. 97). If, as we saw, sensory systems receive information from other systems then their outputs are not constrained in this way. That is, they are not shallow.

However, there is a possible characterization of shallowness that is independent of encapsulation. Prinz affirms that something shallow if it implements a computationally cheap mechanism in the

sense that its outputs are very simple and therefore they require very little processing [25]. However, it would be inaccurate to say that, for instance, visual recognition in the ventral stream involves a simple process or output. First, the mentioned hierarchical models of the ventral stream suggest that visual recognition is achieved through the concatenation of many layers of information processing from the retina, through the LGN, V1, V2, V4 and the inferotemporal cortex (IT) (see [6]). That is, the mechanism that produces the output of this pathway is not cheap but is rather constituted by many successive computational operations. Additionally, the output itself is not simple either. It has been argued that one of the main objectives of this pathway is precisely to form a complex (and therefore selective) representation of visual objects. For instance, the last visual area of the ventral pathway, the IT [42], produces face representations that require combining many elements (e.g., many oriented bars at different positions and locations) from earlier processing stages [23].

This implies that the visual system does not satisfy any of the conditions that characterize modular systems. As I mentioned, given that the visual system is a paradigmatic sensory system it is doubtful that perceptual systems are modular. Therefore, the fact that the ANS satisfies Fodor's conditions does not imply that it is perceptual.

4. Varieties of Cross-Modal Responses

Jones claims that if the ANS can be considered perceptual according to the criteria I discussed in the previous section then it would not be problematic for the grounded view in general but only for a stronger version proposed by Jesse Prinz [10]. The form of concept empiricism defended by Prinz implies that perceptual systems are "dedicated input systems" [7] (p. 115–117). As we saw, this means that these are systems dedicated to process information obtained through a specific input type (such as wavelengths of light, frequency of molecular motion or molecular shape). Given that the ANS manipulates representations of numerosity that are independent of input modality it cannot be characterized as a dedicated input system.

However, Jones argues that Prinz' proposal is undermined by the studies I mentioned in the previous section, which provide evidence for many multimodal neurons and intermodal connections in early stages of sensory processing. This compels us to revise the idea that sensory systems are dedicated to a specific input type. Nevertheless, I will argue that we do not need to jettison the dedicated-input proposal altogether but we can rather formulate a new version which is consistent with the mentioned studies. The cross-modal responses that have been found in sensory processes are consistent with some form of modal input processing.

Before characterizing perceptual representations, it is useful to have in mind different response profiles than can be found within the brain areas we are considering. In the first place, unimodal neurons are those that are only triggered by stimuli from a single modality. Unimodal neurons can be either completely unaffected by stimuli from a different modality or *modulated* by them. The process of modulating a given neural response can be characterized by comparing it with the process of triggering or driving a response. The response r of a neuron (e.g., variations in its spike rate) is driven or triggered by a given variable v_1 (for instance, the spike rate of another neuron) when producing specific variations in the value of v_1 causes specifics variations in r. In contrast, when variations in a variable v_2 do not cause variations in r by themselves but rather change the input–output relation between r and its driving variable v_1 we can say that v_2 modulates this input–output relation (e.g., [43]). For instance, dopamine neurons modulate a given postsynaptic neuron by changing the state of some of its receptors (e.g., glutamate or NMDA receptors). When this postsynaptic neuron is affected by dopamine molecules, its responses to a pre-synaptic driving neuron are potentiated (that is, the same pre-synaptic driving input causes a stronger post-synaptic response) (see [44], pp. 124, 125). Dopamine neurons change the input–output relation between two neurons. A unimodal neuron could be triggered by inputs from one modality but also modulated by stimuli from a different modality.

In the second place, conjunctive and disjunctive neurons can both be *triggered* or driven by stimuli from different modalities. The hallmark of a conjunctive neuron is that its response to the

simultaneous presence of stimuli from different modalities is significantly stronger than its response to either stimulus presented alone (see [45]). In contrast, the maximal response of a disjunctive neuron can be produced by a stimulus from a single modality.

Cappe, Rouiller, and Barrone affirm that multimodal responses found in early perceptual areas involve cross-modal *modulation* [39]. In primary sensory areas such as V1 and A1, multisensory interaction is predominantly of modulatory influence with no (or weak) response to non-specific sensory stimuli.

Based on this fact, we can say that a given neuron *n* is dedicated to a specific input Type I if and only if it is only triggered or driven by instances of I. This is consistent with the possibility that n is modulated by inputs that are not instances of I. Furthermore, Allman and colleagues argue that neurons which exhibit cross-modal modulation show that some degree of multimodal integration can be performed by uni-modal neurons [46]. Typically, multimodal integration is performed by conjunctive neurons, which are often considered multi-modal. Allman and colleagues claim that neurons with cross-modal modulation constitute an earlier and still unimodal stage of multimodal integration [46].

Conjunctive neurons could also be present in sensory areas. Cappe, Rouiller, and Barrone claim that connections found between different modalities might underlie a representation of this kind [39]. This is suggested by the fact that behavioral responses in monkeys to multisensory stimuli were faster than those triggered by unisensory stimuli. This is consistent with the general response profile of known multi-modal neurons. These neurons have been mostly found in the superior colliculus (SC). Individual SC neurons are maximally responsive to the simultaneous presence of stimuli from different modalities appearing at the same location.

A multisensory SC neuron has multiple receptive fields, one for each sensory modality to which it is responsive. A receptive field defines the area of sensory space (e.g., visual space, auditory space, somatosensory space) where a given stimulus will be effective in activating the neuron. For multisensory neurons, their different receptive fields are in spatial register with one another. For example, a visual-auditory neuron with a visual receptive field in central space will be responsive to auditory cues in a roughly overlapping region of auditory space. The response of SC neurons to cross-modal stimuli is typically stronger than one evoked by either of the modality-specific stimuli [45].

These conjunctive neurons cannot be considered modal according to the criterion I proposed above for neurons with cross-modal modulation because they are triggered (and not merely modulated) by inputs from different modalities. A more adequate criterion is that a representation is modal if and only if it preserves information about the modality of its triggering stimuli. This is something that conjunctive neurons share with unimodal neurons (including unimodal neurons with cross-modal modulation). If we know that a neuron only responds to visual stimuli, its signaling at time *t* will carry the information that its preferred stimulus is presented in visual format. That is, it will carry information about the specific modality of the triggering stimulus. For instance, if a neuron in subject S's brain only responds to visually presented dogs, its signaling at time *t* carries the information that a dog was visually presented to S. As we saw, SC conjunctive neurons respond to inputs from different modalities that are presented at the same time. Suppose that a SC neuron is triggered only when a visual and an auditory stimulus are simultaneously present (e.g., a dog image and a dog bark). When this structure is triggered at *t*, we can tell that a visual (and also an auditory) stimulus is present. Its responses can carry information about the modality of its inputs in this sense. As I mentioned in Section 1, the idea that modal representations are representations that carry modality-specific information is an assumption shared by different grounded theorists (e.g., [3–5]).

In contrast, disjunctive representations will preserve no information about the modality of the triggering input. As I mentioned, these neurons can be triggered by stimuli from different modalities but either one of them is sufficient to elicit the (maximal) response. For instance, we saw that the maximal response of a number neuron is triggered by a given numerosity when it is presented in either visual or auditory format. Either a visual or an auditory stimulus can trigger a neural response at time

t by itself. This means that the response of a disjunctive number neuron at *t* carries the information that a stimulus *s* presented at *t* has a given numerosity but it does not carry information about *s'* modality (i.e., the modality cannot be decoded from this response). Given that these disjunctive responses are incompatible with modality-specific information, the frontal and parietal neurons that have this response profile constitute amodal number representations.

We can reformulate the idea that a neural structure is a dedicated-input system as the idea that this system is dedicated to process (i.e., it only processes) modality-specific information. Given that a system can be dedicated to process different kinds of modality-specific information (e.g., auditory and visual information) being sensory is consistent with being multimodal (i.e., responding to different kinds of inputs). Neurons in traditional sensory systems exhibit cross-modal responses which are consistent with modality-specific information and therefore this version of the dedicated-input approach implies that they are perceptual. In contrast, the ANS processes representations that do not carry modality-specific information. It cannot be considered a perceptual system because it is not exclusively concerned with modality-specific information. Thus, a version of the dedicated-input approach which is consistent with the evidence mentioned by Jones can be used to show that the ANS is not a perceptual system.

5. Conclusions

I have defended the amodal approach to the ANS against two complementary arguments proposed by Jones [10]. In the first place, I argued that the fact that the ANS is modular does not imply that it is perceptual because it is doubtful that perceptual systems are modular. This is because a paradigmatic perceptual system (i.e., the visual system) satisfies none of the conditions for modular systems. Additionally, I argued that the kinds of cross-modal responses that sensory systems exhibit are different from those exhibited by the ANS. Specifically, the representations in traditional sensory systems have cross-modal responses which carry modality-specific information. If we characterize perceptual systems as systems dedicated to process modality-specific information, we can say that traditional sensory systems are perceptual without denying that they are multimodal. In contrast, the ANS is not perceptual because it does not process exclusively modality-specific information.

These arguments imply that at least some relevant cognitive processes cannot be performed by recruiting only perceptual resources. Numerical cognition requires the activation of an amodal system. This would undermine strong versions of the grounded approach (e.g., [7,47,48]), which claim that modality-specific systems are necessary and sufficient to support conceptual processing [49]. It is relevant to notice that many grounded theorists currently endorse a weak version of the approach, which claims that concept-related tasks require both modality-specific and modality-independent structure [3,50–55]. Barsalou has recently claimed that strong versions constitute a usual distortion of grounded cognition [3]. If this is so, the fact that the ANS is amodal would pose no real threat for this view.

However, as Mahon points out, the problem with weak grounding is that it is not different or inconsistent with an amodal view ([55], see also [56] for a similar point). No amodal theorist denies that conceptual processing requires some degree of interaction between modal and amodal systems. Taking this into account, it follows that the ANS poses a serious challenge for the grounded view. It forces us to endorse a form of grounding which is too weak to be distinguished from the alternative approach to cognition.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Barsalou, L.W.; Kyle Simmons, W.; Barbey, A.K.; Wilson, C.D. Grounding conceptual knowledge in modality-specific systems. *Trends Cogn. Sci.* **2003**, *7*, 84–91. [CrossRef]

2. Barsalou, L.W. Grounded Cognition: Past, Present, and Future. *Top. Cogn. Sci.* **2010**, *2*, 716–724. [CrossRef] [PubMed]

3. Barsalou, L.W. On Staying Grounded and Avoiding Quixotic Dead Ends. *Psychon. Bull. Rev.* **2016**, *23*, 1122–1142. [CrossRef] [PubMed]

4. Reilly, J.; Peelle, J.E.; Garcia, A.; Crutch, S.J. Linking somatic and symbolic representation in semantic memory: The dynamic multilevel reactivation framework. *Psychon. Bull. Rev.* **2016**, *23*, 1002–1014. [CrossRef] [PubMed]

5. Binder, J.R. In defense of abstract conceptual representations. *Psychon. Bull. Rev.* **2016**, *23*, 1096–1108. [CrossRef] [PubMed]

6. DiCarlo, J.J.; Zoccolan, D.; Rust, N.C. How Does the Brain Solve Visual Object Recognition? *Neuron* **2012**, *73*, 415–434. [CrossRef] [PubMed]

7. Prinz, J.J. *Furnishing the Mind: Concepts and Their Perceptual Basis*; MIT Press: Cambridge, MA, USA, 2002.

8. Dove, G. Beyond Perceptual Symbols: A Call for Representational Pluralism. *Cognition* **2009**, *110*, 412–431. [CrossRef] [PubMed]

9. Machery, E. Concept Empiricism: A Methodological Critique. *Cognition* **2006**, *104*, 19–46. [CrossRef] [PubMed]

10. Jones, M. Number Concepts for the Concept Empiricist. *Philos. Psychol.* **2016**, *29*, 334–348. [CrossRef]

11. Machery, E. The amodal brain and the offloading hypothesis. *Psychon. Bull. Rev.* **2016**, *23*, 1090–1095. [CrossRef] [PubMed]

12. Dehaene, S.; Naccache, L.; Le Clec'H, G.; Koechlin, E.; Mueller, M.; Dehaene-Lambertz, G.; van de Moortele, P.F.; Le Bihan, D. Imaging unconscious semantic priming. *Nature* **1998**, *395*, 597–600. [CrossRef] [PubMed]

13. Pecher, D.; Zeelenberg, R.; Barsalou, L.W. Verifying different-modality properties for concepts produces switching costs. *Psychol. Sci.* **2003**, *14*, 119–124. [CrossRef] [PubMed]

14. Van Dantzig, S.; Pecher, D.; Zeelenberg, R.; Barsalou, L.W. Perceptual processing affects conceptual processing. *Cogn. Sci.* **2008**, *32*, 579–590. [CrossRef] [PubMed]

15. Collins, J.; Pecher, D.; Zeelenberg, R.; Coulson, S. Modality Switching in a Property Verification Task: An ERP Study of What Happens When Candles Flicker after High Heels Click. *Front. Psychol.* **2011**, *2*, 10. [CrossRef] [PubMed]

16. Piazza, M.; Mechelli, A.; Price, C.J.; Butterworth, B. Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Res.* **2006**, *1106*, 177–188. [CrossRef] [PubMed]

17. Damarla, S.R.; Cherkassky, V.L.; Just, M.A. Modality-independent representations of small quantities based on brain activation patterns. *Hum. Brain Mapp.* **2016**, *37*, 1296–1307. [CrossRef] [PubMed]

18. Nieder, A. The neuronal code for number. *Nat. Rev. Neurosci.* **2016**, *17*, 366–382. [CrossRef] [PubMed]

19. Nieder, A. Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11860–11865. [CrossRef] [PubMed]

20. Burr, D.; Ross, J. A Visual Sense of Number. *Curr. Biol.* **2008**, *18*, 425–428. [CrossRef] [PubMed]

21. Ross, J.; Burr, D.C. Vision senses number directly. *Nat. Preced.* **2009**. [CrossRef] [PubMed]

22. Hubel, D.H.; Wiesel, T.N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* **1962**, *160*, 106–154. [CrossRef] [PubMed]

23. Mély, D.A.; Serre, T. Towards a Theory of Computation in the Visual Cortex. In *Computational and Cognitive Neuroscience of Vision*; Cognitive Science and Technology; Springer: Singapore, 2017; pp. 59–84.

24. Fodor, J. *The Modularity of Mind*; MIT Press: Cambridge, MA, USA, 1983.

25. Prinz, J.J. Is the Mind Really Modular? In *Contemporary Debates in Cognitive Science*; Stainton, R.J., Ed.; Blackwell: Oxford, UK, 2006; pp. 22–36.

26. Carruthers, P. *The Architecture of the Mind: Massive Modularity and the Flexibility of Thought*; Oxford University Press: Oxford, UK, 2006.

27. Piazza, M.; Izard, V.; Pinel, P.; Le Bihan, D.; Dehaene, S. Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron* **2004**, *44*, 547–555. [CrossRef] [PubMed]

28. Nieder, A.; Miller, E.K. A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7457–7462. [CrossRef] [PubMed]

29. Andreau, J.M.; Funahashi, S. Primate prefrontal neurons encode the association of paired visual stimuli during the pair-association task. *Brain Cogn.* **2011**, *76*, 58–69. [CrossRef] [PubMed]

30. Tomita, H.; Ohbayashi, M.; Nakahara, K.; Hasegawa, I.; Miyashita, Y. Top-down signal from prefrontal cortex in executive control of memory retrieval. *Nature* **1999**, *401*, 699–703. [CrossRef] [PubMed]

31. Nieder, A.; Freedman, D.J.; Miller, E.K. Representation of the Quantity of Visual Items in the Primate Prefrontal Cortex. *Science* **2002**, *297*, 1708–1711. [CrossRef] [PubMed]

32. Lee, T.S.; Mumford, D. Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **2003**, *20*, 1434–1448. [CrossRef] [PubMed]

33. Poggio, T.; Serre, T. Models of visual cortex. *Scholarpedia* **2013**, *8*, 3516. [CrossRef]

34. Yau, J.M.; Kim, S.S.; Thakur, P.H.; Bensmaia, S.J. Feeling form: The neural basis of haptic shape perception. *J. Neurophysiol.* **2015**, *115*, 631–642. [CrossRef] [PubMed]

35. Bahrami, B.; Vetter, P.; Spolaore, E.; Pagano, S.; Butterworth, B.; Rees, G. Unconscious Numerical Priming Despite Interocular Suppression. *Psychol. Sci.* **2010**, *21*, 224–233. [CrossRef] [PubMed]

36. Ashkenazi, S.; Henik, A.; Ifergane, G.; Shelef, I. Basic numerical processing in left intraparietal sulcus (IPS) acalculia. *Cortex* **2008**, *44*, 439–448. [CrossRef] [PubMed]

37. Zhou, Y.-D.; Fuster, J.M. Visuo-tactile cross-modal associations in cortical somatosensory cells. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9777–9782. [CrossRef] [PubMed]

38. Sathian, K.; Zangaladze, A. Feeling with the mind's eye: Contribution of visual cortex to tactile perception. *Behav. Brain Res.* **2002**, *135*, 127–132. [CrossRef]

39. Cappe, C.; Rouiller, E.M.; Barone, P. Multisensory anatomical pathways. *Hear Res.* **2009**, *258*, 28–36. [CrossRef] [PubMed]

40. Falchier, A.; Schroeder, C.E.; Hackett, T.A.; Lakatos, P.; Nascimento-Silva, S.; Ulbert, I.; Karmos, G.; Smiley, J.F. Projection from Visual Areas V2 and Prostriata to Caudal Auditory Cortex in the Monkey. *Cereb. Cortex* **2010**, *20*, 1529–1538. [CrossRef] [PubMed]

41. Falchier, A.; Clavagnier, S.; Barone, P.; Kennedy, H. Anatomical Evidence of Multimodal Integration in Primate Striate Cortex. *J. Neurosci.* **2002**, *22*, 5749–5759. [CrossRef] [PubMed]

42. Gross, C.G. Inferior temporal cortex. *Scholarpedia* **2008**, *3*, 7294. [CrossRef]

43. Silver, R.A. Neuronal arithmetic. *Nat. Rev. Neurosci.* **2010**, *11*, 474–489. [CrossRef] [PubMed]

44. Konradi, C.; Cepeda, C.S.; Levine, M. Dopamine—Glutamate Interactions. In *Handbook of Experimental Pharmacology*; Springer: Singapore, 2002; Volume 154, pp. 117–133.

45. Stein, B.E.; Stanford, T.R.; Rowland, B.A. Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* **2014**, *15*, 520–535. [CrossRef] [PubMed]

46. Allman, B.L.; Keniston, L.P.; Meredith, M.A. Not Just for Bimodal Neurons Anymore: The Contribution of Unimodal Neurons to Cortical Multisensory Processing. *Brain Topogr.* **2009**, *21*, 157–167. [CrossRef] [PubMed]

47. Gallese, V.; Lakoff, G. The Brain's concepts: The role of the Sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* **2005**, *22*, 455–479. [CrossRef] [PubMed]

48. Glenberg, A.M.; Gallese, V. Action-based language: A theory of language acquisition, comprehension, and production. *Cortex* **2012**, *48*, 905–922. [CrossRef] [PubMed]

49. Mahon, B.Z. What is embodied about cognition? *Lang. Cogn. Neurosci.* **2015**, *30*, 420–429. [CrossRef] [PubMed]

50. Binder, J.R.; Desai, R.H. The neurobiology of semantic memory. *Trends Cogn. Sci.* **2011**, *15*, 527–536. [CrossRef] [PubMed]

51. Hauk, O.; Tschentscher, N. The Body of Evidence: What Can Neuroscience Tell Us about Embodied Semantics? *Front. Psychol.* **2013**, *4*. [CrossRef] [PubMed]

52. Kiefer, M.; Pulvermüller, F. Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex* **2012**, *48*, 805–825. [CrossRef] [PubMed]

53. Ralph, M.A.L. Neurocognitive Insights on Conceptual Knowledge and Its Breakdown. Available online: https://www.escholar.manchester.ac.uk/uk-ac-man-scw:206214 (accessed on 14 April 2018).

54. Meteyard, L.; Cuadrado, S.R.; Bahrami, B.; Vigliocco, G. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* **2012**, *48*, 788–804. [CrossRef] [PubMed]

55. Zwaan, R.A. Embodiment and language comprehension: Reframing the discussion. *Trends Cogn. Sci.* **2014**, *18*, 229–234. [CrossRef] [PubMed]
56. Leshinskaya, A.; Caramazza, A. For a cognitive neuroscience of concepts: Moving beyond the grounding issue. *Psychon. Bull. Rev.* **2016**, *23*, 991–1001. [CrossRef] [PubMed]