Pointing as an Instrumental Gesture: Gaze Representation Through Indication

Massimiliano L. Cappuccio †
m.lorenzo@uaeu.ac.ae

Mingyuan Chu‡
mingyuan.chu@mpi.nl

Sotaro Kita *
s.kita@bham.ac.uk

ABSTRACT

We call those gestures "instrumental" that can enhance certain thinking processes of an agent by offering him representational models of his actions in a virtual space of imaginary performative possibilities. We argue that pointing is an instrumental gesture in that it represents geometrical information on one's own gaze direction (i.e., a spatial model for attentional/ocular fixation/orientation), and provides a ritualized template for initiating gaze coordination and joint attention. We counter two possible objections, asserting respectively that the representational content of pointing is not constitutive, but derived from language, and that pointing directly solicits gaze coordination, without representing it. We consider two studies suggesting that attention and spatial perception are actively modified by one's own pointing activity: the first study shows that pointing gestures help children link sets of objects to their corresponding number words; the second, that adults are faster and more accurate in counting when they point.

KEYWORDS: Pointing; Instrumental gestures; Joint attention; Counting; Gaze representation; Action simulation; Spatial cognition.

† Department of Philosophy, UAE University, United Arab Emirates.
‡ Neurobiology of Language, Max Planck Institute for Psycholinguistics, the Netherlands.
* School for Psychology, University of Birmingham, the United Kingdom.

1. Gestures facilitate cognitive tasks

We abundantly use gestures outside the communicative context, when nobody can see us performing them, and even when they don't have any directly manipulative/transformational scope. This is importantly true of finger pointing too, as the use of this gesture is not confined to communication. Why do we produce gestures in general, then, and pointing in particular? Some influential trends in psycholinguistics and philosophy of mind contend that hand gestures are not just the final motor byproduct of cognitive processing, or a supplementary aid to communication, but can also intervene to facilitate some cognitive tasks, soliciting the required thinking processes or simplifying the underlying computations (Chu & Kita, 2011).

To deepen this theory, we will focus on instrumental gestures, i.e. gestures whose distinctive effect is to actively facilitate or enhance certain thinking processes of the agent, regardless that her gesturing intervenes in solo or social tasks. For example, experiments show that subjects free to produce gestures perform better in tasks of information recollection (either linguistic, using lists of words, as in Goldin-Meadow, 2003, or visuospatial, using dots positioned in a grid net, as in Goldin-Meadow et al., 2001 and Wagner et al., 2004), learning of mathematical problem-solving strategies (Broaders et al., 2007; Goldin-Meadow et al., 2009), visuomotor transformation (e.g., mental rotations, see Chu & Kita 2008, 2011), and visualization for communication (Hutchins 2010). In communicative contexts, gesturing helps speaking by highlighting perceptually present information for speakers (Alibali & Kita 2010) and affecting fluency of speech (Rauscher et al., 1996), pauses (Graham & Heywood, 1975) and contents (Rimé et al., 1984).

How do instrumental gestures facilitate thinking? Crucially, the correlation between gesturing and performance increase is neither reducible to an incidental association between cognitive processing and behavioral production, nor to the transferring of information patterns from linguistic to visuo-motor memory: in fact, gestures facilitate problem-solving whether verbal information is required or not (Goldin-Meadow & Beilock, 2000; Chu & Kita, 2011), and provide rich perceptual-motoric information that can't be found in the accompanying speech (Goldin-Meadow & Beilock 2010); moreover, absence of gestures, whether imposed or spontaneous, correlates with poorer performances, confirming that the gestural activity can take over part of the computational work (Goldin-Meadow, 2003).

This is perfectly illustrated by the experiments with mental rotations, as there is evidence that participants use motor simulation to solve the object-matching task (see Schwartz & Holton, 2000; Wexler et al., 1998; Hegarty, 2004). The fact that the cognitive facilitation through motor simulation is bounded by the specific motor constraints of hand movements (Sekiyama, 1982) suggests that the problem-solving process is simplified because it is partly offloaded to the covert activity of the motor system and the overt activity of its motor effectors. The motor simulation can also be implemented by overt hand movements, namely hand gestures that simulate the shape of objects or their orientation in space in a physically controllable and directly perceivable manner, surrogating the covert visualization (in imagination) of the equivalent internal models (Chu & Kita, 2008, 2011; Ehrlich, Levine & Goldin-Meadow, 2006).

We endorse a particular interpretation of this conclusion. We assert that instrumental gesture is effective in cognitive tasks because it offers a model of an action in a virtual space of imaginary performative possibilities, i.e. a motor simulation that can be inspected and manipulated to predict the effects of an action or postdict its causes. In other words, we define as instrumental gestures all gestures that actively facilitate certain thinking processes of the agent by representing before him some of his possible actions. The content of this representation is actional and embodied in nature, as it provides a motoric simulation of the action execution, not a linguistic expression of its conceptual meaning: we call this view the Gesture-as-Action-based-Thinking hypothesis (GAH). Accordingly, instrumental gestures ("pretending to" rotate an object) can work as models of the corresponding goal-oriented actions (actually rotating the object) because the former represents the key motoric patterns of the latter. At the same time, an instrumental gesture significantly differs from a fully-fledged instantiation of that action: while an instrumental action aims to achieve a transformative goal in the physical environment, the gesture aims only to display a model of the possible effects associated to such goal. However, as a matter of fact, we can recognize actions from the corresponding gestures more or less like we identify real objects from their sketches. This analogical relation suggests that gestures seem specifically tailored to help some of our cognitive operations.

That said, one might still wonder why instrumental gesture might be better at facilitating thinking than action does (Goldin-Meadow & Beilock, 2010). Our answer is that GAH can best account for the fact that gestures do not only follow from, but also actively re-shape our problem-solving processes by strategically instantiating in flesh and bones some concretely perceivable representation of those very processes. The subject's performances are actually augmented by the fact that gestures, unlike actual actions, are not constrained by sensorimotor contingencies, and thus can solicit the production of rich representations of virtual possibilities of actions; because gesturing allows her to physically embody and directly inspect a perceptual representation of the general structure of these virtual possibilities, gestures acquire the function to model, complete, speed-up and make more precise her own thoughts. Later in our discussion we will show that pointing is a particular and significant form of instrumental gesture, one that can successfully exemplify how this happens.

2. Gestures have representational contents

The distinctive idea of GAH is that instrumental gestures facilitate a subject's thought process by displaying it externally, materially incarnating it in front of the subject and, possibly, of other observers, making it available for perceptual appreciation and overt (possibly public) consideration. This idea seems in line with the psychological definition of a gesture as a ′material carrier' of meaning (Vygotsky 1962), i.e. a physical realization of cognitive activity that has systematic effects: by this reason, the "actual motion of the gesture itself is a dimension of thinking [...], thinking in one of its many forms″ (McNeill 2000, p.98).

Instrumental gestures affect our cognitive operations precisely because, while dynamically integrated in the subject's brain-bodily unfolding (like actual actions), they are simultaneously exapted as vehicles of representational/symbolic content, i.e. models of virtual actions that gestures offer for explicit consideration and manipulation. Experiments can extensively document how, with protracted practice, gestures and mental rotation capabilities co-vary, generating novel problem-solving patterns. Both the frequency and the typology of the produced gestures vary with practice (Chu & Kita 2008): while object-movement gestures (representing the movement of an object by itself) increase, hand–object interaction gestures (pantomiming the agentive hand action on an object) decrease. This suggests that a new competence (based on previously unavailable abstract models of manipulation) is added to preexisting covert capabilities of mental rotation, and correlates with the gradual internalization of the models of the virtual actions. This hypothesis is strengthened by the evidence that verbal description of rotation in concurrent speech produces more agent-independent reports over time, and that the change of motor strategies is hampered when gestures are prohibited (ivi, p. 718). Such evidence confirms that gestures play an active casual role, as opposite to a merely expressive one, in solving the task of object matching: gestures do not simply explicate fully formed thoughts, they additionally mirror (and help complete) thoughts in statu nascendi, substantially contributing to shape them throughout their development.

Why should this contribution be representational in nature, even if only in a minimal sense? Instrumental gestures re-enact pre-specified body schemata as vicarious stand-ins for virtual actions or possible states of affairs: if we can intuitively differentiate between action and gesture this is precisely because we attribute a symbolic content to the latter. Such differentiation is perceptually possible because the normative conditions of gestures (the conditions under which gestures are successful), differ from the normative conditions of actions: an action is performed to accomplish a physical transformation, not to represent it; on the contrary, the corresponding gesture intends precisely to refer to that transformation, not to achieve it. Gestural movements suggest the effects of their corresponding action even if (and – to some extent - exactly because) it is evident that the performer doesn't intend to produce those effects: this requires complex capabilities of implicit projection and imaginary variation that allow visualizing possible sensorimotor contingencies that the gesture indirectly represents in a virtual space, rather than simply observing actual sensorimotor contingencies to which its movements would be directly adjusting to.

Endorsing a similar claim, Hutchins (2010, p. 97) notes that it is by bringing online such visualization skills that "gesture depicts a fictional event that facilitates reasoning", either reflexively (in the signaler herself) or communicatively (in its interlocutors). And Goldin-Meadow (2003, p. 186) asserts that gestures enrich our representational skills, either replicating the contents available in verbal format, or supplementing and nuancing them with information in the visuomotor format. The cognitive efficacy of gestures as overt virtual models is best accounted for by the concept of 'external representation' (Kirsch, 2010a): not too differently from other representational media (instantiated by speech, pictures, drawings, plastic arts, or – in a closer proximity with gestures - dance and pantomime), the gestural medium, instantiated by action movements, offers a symbolic content that is spatially localized, publicly accessible, and passively (perceptually) affecting the observers and the performer herself.

With respect to mental rotation tasks, Chu & Kita (2008) documents the symbolic function of instrumental gestures, which are not only produced to evoke bodily schemata of familiar actions (one-to-one simulation of the action's movements through the gesture's), but also to represent some increasingly general patterns of movement representing the context of the action. The model manipulated by the gesture doesn't need to be a fully-fledged motor re-enactement of the set of movements that implement an instrumental action; on the contrary, the gesture tends to symbolically exapt those movements to reference an entire typology of virtual events or features consistently associated to their goals (directions, trajectories, speed, or other general properties of objects and actions).

The symbolic function of gestures progressively becomes more important while the agent becomes more familiar with his gesture-based problem solving strategies. This happens in three stages the specific roles of which are empirically documented: in the first stage, gestures mimic the actions that would have been actually used to manipulate the objects (suffering from limitations by both the physical features of the object and the anatomical restriction of the body parts). In the second stage, gestures are de-agentivized: they don't reproduce familiar actions, but abstractly model them (body parts, especially the hand, represent the object, and the gestural trajectories represent the motion of the object). The problem solving strategy is detached: it is only limited by the anatomical restriction of the agent's body, rather than by the physical features of the object to be manipulated. In the third stage, the skill gained from the first two stages becomes internalized, as no overt bodily manipulation or representation is necessary anymore, and problem solving improves as it gets free from the limitations imposed by the sensorimotor contingencies. Summing up, the visible effects of de-agentivization and detachment result in more abstract, self-contained, gestural representations, culminating in fluent and efficacious internalization of the action simulation.

The gesture's incremental lack of precise correspondence with actual sensorimotor contingencies testifies that the agent is gaining more abstract and comprehensive internal models of potential situations (as recognized by McNeill, 2000, and recently confirmed by Goldin-Meadow & Beilock, 2010, p. 670). The abstract value of the gesture as a virtual model may increase with de-agentivization and detachment, even if the gesture's resemblance to its corresponding goal-oriented actions decreases. Interestingly, de-agentivization itself follows from the fact that the "inherent instability of motor execution may serve as a reservoir for different possible strategies" (Chu & Kita 2008, p. 721). Chu & Kita (2011) confirms that spontaneously produced gestures help solving spatial visualization problems (mental rotation task and a paper folding task); but as the subjects solve more problems, the spatial computation supported by gestures becomes internalized, and the gesture frequency decreases. The computational facilitation can be generalized to similar spatial visualization tasks, and persists even when, in subsequent spatial visualization problems, gesture is prohibited.

Since our analysis targets only self-affecting gestures (and not, for example, attention-getting gestures, emblem gestures, and interactive gestures), it is possible to conclude that the specific patterns of similarity/difference between such gestures and the corresponding actions characterize their relation as functionally dedicated (inherently structured to performance enhancement), rather than extrinsic (generally associative): their correspondence serves a referential function (i.e., the gesture serves to vicariate the action, not vice versa), in that it provides a representational model (i.e., only the gesture, not the action, has a content) that enables the manipulation of contextual information relevant for facilitating certain cognitive tasks. With regard to the similarities between gestures and actions, the gesture of hand rotations overtly embodies at least some of the general patterns (motoric, topological, and morphological) that would be necessary to actually rotate the objects, even without actually needing to fully accomplish this transformation in real space (this would prevent the gesture from visualizing the details of the imaginary rotations). With regard to the differences, according to the intuition proposed by Werner and Kaplan (1963), and later developed by Piaget (1968), in development an action can begin to refer to its typical outcome only in virtue of some intervening distance (physical and symbolic) between the former and the latter (e.g., movements typically involved in a certain form of interaction become detached requests for that interaction). Does reference mediated by distance/detachment implies representation? There are different opinions on this regard.

For example, some enactive (anti-representationist) approaches to cognition deny that contentful representations are among the constitutive components of basic minds (Hutto & Myin, 2012). This claim doesn't conflict with our representationist approach to instrumental gestures, which accounts for advanced, symbolic forms of cognition, rather than basic ones. An (at least minimally) representational story seems necessary to account for the functional difference between goal-oriented actions and instrumental gestures: namely, only instrumental gestures can function as material carriers of symbolic thought, as they do not only solicit one's own thought processes, but specifically display them perceptually. While skillful, unreflective goal-oriented actions fundamentally are motor output adaptively modulated to match the perceptual reality, gestures (as well as ritualized embodied routines) work also as super-imposed models that actively change the perception of the context, transfiguring it into "a readily available problem-solving resource, one whose elements restructure the piece-finding problem and thereby reduce the information processing load placed on the inner mechanisms in play" (Wheeler in press). Representations seem required in that a content/vehicle distinction is in place: once they have developed a full symbolic distance from the corresponding actions, gestures do not simply re-enact the sensorimotor contingencies of those actions, but also invite to withdraw from them, while attending to the action context from a relatively detached perspective (Cappuccio & Wheeler, 2011; see also McNeil, 2000 p. 156, with reference to the Heideggerian account of 'signs' developed by Dreyfus, 1994).

3. Pointing: a case for structural de-agentivization and detachment

It is usually assumed that instrumental gestures can only represent goal-oriented hand actions, like rotations of three dimensional objects in space. But we propose that pointing is also an instrumental gesture, even if apparently it doesn't represent any such hand action. We claim that pointing illustrates very well how a gesture can enhance our thinking processes because the symbolic contents made available by it happen to facilitate certain cognitive tasks. Why is pointing an instrumental gesture? And what is it an instrument for?

Consider the extreme case of "abstract deictic gestures, which point to a seemingly empty location in front of the speaker or move as if to track a moving object" (McNeill 2000, p. 170). These gestures "could be related to the ability to orient our body parts (e.g., gaze and the hand) toward a target in the physical environment, and to the ability to track the target when it moves." According to Kita (2000, 170-171), pointing provides a peculiar case for GAH because, even if it doesn't need an "obvious enactment component" (i.e., it doesn't aim to produce direct physical manipulation of the external reality, like other instrumental gestures), it is "also formed by spatio-motoric thinking" (it re-configures the subjective spatial perception and the attentional state of both the producer and of the observers). Index-finger pointing has a prototypical function as an embodied vehicle of externalized, symbolic representation. The representational content conveyed by pointing is fundamentally twofold.

Primarily, it conveys geometrical information on gaze direction, i.e. a model of a virtual agent's system of attentional/ocular fixation/orientation (Cappuccio & Shepherd, 2013; Shepherd & Cappuccio, 2012). This view is in line with the individualistic accounts of the ontogenesis of pointing (Vygotsky 1962, Bates 1975), and is corroborated by the studies suggesting that the first instances of pointing are manifestations of the child's own attention during tactile exploration of close-by objects, and that only subsequently this gesture becomes social in its use, while infants become gradually aware of the responses given to this action by the adults (Carpendale & Carpendale 2010). By pointing, the agent maps the abstract information on her gaze/attention into the set of performative hand-based and postural motor skills that concretely allow her to align finger, hand, arm, torso and whole-body with the orientation of her eyes, hence facilitating future actions on the target. Through this mechanism, the pointing subject himself can look into the direction indicated by his own pointing finger. This situation invites him to split his attention between a certain direction and, simultaneously, the index finger (signifier) that symbolizes attention towards the very same direction: this division is also a duplication of the concrete pointing subject into an agent who looks while looking at himself looking. Pointing is a symbolic replay (a predictive/postdictive model) of gaze and posture direction that refers to the very action of gazing while being categorically irreducible to it: this distance is structural because an index finger can replicate the visual orientation of the eyes but obviously, unlike the eyes, doesn't produce its own sense of vision. Its function is parallel to but incommensurably different from looking and, therefore, inherently abstract. In the virtual space drawn by pointing, the index-finger extension acts as a virtual stand-in that informs the observers (including the signaler himself) of the mere possibility of a certain direction of sight, implicitly inviting them to align their gaze accordingly.

Secondarily, the representational content conveyed by pointing encompasses also the symbolic meaning that is associated to the designated object for the fact that the signaler intentionally highlights it (Bruner 1983).

Pointing is one of children's first symbolizing devices in the "joint attentional frame" described by Tomasello (1999). The signaler's attention is not only incidentally communicated by his eye direction, it is also intentionally represented by his pointing finger. The signaler's attentional state affects the subjective perception of the indicated target in all the recipients that are aware of the signaler's pointing and can understand its declarative intention, recognizing in this target a focus of joint attention and symbolic representation. The intersubjective meaning of this common focus is actively modified by the very act of declarative pointing on the basis of the common ground of skills, knowledge, and cooperative goals that the co-attenders share: e.g., an empty chair can be highlighted as either a symbol of the person who usually sits on it or as an invitation to sit on it, depending on the contextually relevant information that the co-attenders share; a physical location in the ecological space of situated agents is transfigured into a formal indicator of a position detached by them, and identified by a metric system of geometric coordinates.

Therefore, the instrumental action that pointing represents is primarily the motor action of ocular re-orientation that targets a certain point in space, and secondarily the imaginary affordances associated to that point on the basis of the background information shared by the signaler and the co-attenders. Its goal is to realign the co-attenders' perceptual field and allow them (including the signaler herself) to see the world under "a different angle", either in a strictly sensorial sense (update the visual perspective, moving the eyes to a new perceivable target) or in a symbolic one (update the associated imaginative content, modifying its representational meaning in function of the joint attention established between co-attenders).

Now, in a communicative scenario, one can solicit a co-attender to bring an object by pointing at it; but, while dealing alone with the physical reality, one cannot displace an object by simply pointing at it with his finger. Like other instrumental gestures, pointing can't be used to directly manipulate or physically transform any object; additionally, unlike other instrumental gestures, also the action that it represents does not aim to produce physical transformation/manipulation: the "eye-action" for which pointing stands in (ocular re-orientation, either voluntary or automatic through saccadic movements) is essentially observational in nature. Another important difference between pointing and the other gestures is that a simple vector with a fixed dimension is sufficient to represent the eye-action in a schematic form via pointing, while the complexity of the hand-actions represented by the other gestures involve a high number of degrees of freedom. The non-manipulative, vector-representing function of pointing makes this gesture one of the most paradigmatic cases of GAH: unlike other instrumental gestures, pointing doesn't just occasionally produce the effects of de-agentivization and detachment; on the contrary, it structurally relies on them to produce its indicative function. Pointing specifically serves to indicate objects in the extra-personal space and highlight them as symbolic referents in a potentially contemplative, virtually detached dimension, not as poles of real actions affording immediate manipulation in a practical dimension.

4. Neuroscience and phenomenology of pointing

Neurofunctional data (see Ulloa and George, this journal issue) suggest that pointing gestures and eye gaze elicit similar attentional responses (Langton & Bruce, 2000), and share a basis of information instantiated by common neuronal structures, mainly the superior temporal sulcus (Sato et al., 2009). This is significant in relation to the distinction between peripersonal and extrapersonal space (the spatial dimension targeted by pointing), as we know that the perception of peripersonal space is modulated by the vocabulary of motor actions mapped in the premotor areas (Rizzolatti et al., 1988), while the extra-personal space is mainly mapped by intraparietal areas related to gaze direction and attention fixation: that is why the manipulation of extrapersonal space "is mainly subserved by oculomotor circuits, in which spatial information arises from neurons whose receptive fields are coded in retinal coordinates" (Neppi-Mòdona et al., 2007).

This neurofunctional dissociation reflects the different behavioral and experiential significance of far and near space, which respectively correspond to the domains of contemplative and manipulative actions. Both neuropsychological (Milner and Goodale 2006) and phenomenological (Kelly, 2001) accounts have stressed the functional differentiation between pointing and goal-oriented hand actions such as reaching for grasping, theorizing that they are controlled by two autonomous cognitive architectures ("ways of vision", or "ways to see" according to Jacob & Jeannerod, 2003) respectively dedicated to the semantic and pragmatic modalities of spatial processing: on the one hand, a "vision for perception", encoding metric coordinates (objective distances), and representational in character (decoupled from immediate interaction and regulated by agent-neutral, context-independent geometrical norms); on the other hand, a "vision for action", adaptively responsive to motor opportunities embedded in practical contexts, and dispositional in character. A precursor to this view can be found in Merleau-Ponty's *Phenomenology of perception* (1945), which builds on a famous neuropsychological case studied by Goldstein and Gelb (1918): according to these classical studies, pointing exploits the first way of vision, because it requires spatial information encoded in terms of positions to transfigure the target objects into stand-ins for symbolic reference and detached representation; manipulative hand-action processes exploit the other way of vision, because they need spatial information in terms of operative situations to respond to ecological affordances by means of pre-reflective embodied skills.

One of the crucial differences at stake is that the representational/symbolic contents play a defining role only for the positional organization of space, and don't have any evident weight in motor intentional behaviors familiarly produced by skilled agents in a space of situation (Kelly 2002). Only the former involves attentive consideration of the distance - both physical and symbolic - that separates the subject and the object of the action: while objects in the peripersonal space pre-reflectively afford possible hand-actions that are familiar to the agent in terms of practical know-how, objects referenced in the extrapersonal space via pointing principally invite contemplative awareness of the bare fact that they are objectively present. This representational and symbolic function of pointing invites two or more co-attenders to establish a particular attentional structure, as previously remarked: it highlights the indicated target as a focus of interest for their joint attention and factual knowledge (common ground), rather than as a usable object affording immediate manipulation.

It seems clear that pointing does have a representational content like other instrumental gestures; but is this sufficient to include pointing in the set of instrumental gestures (like hand rotations), in accord with GAH? Two objections might be deployed to resist this inclusion. The first intends to show that the format of the representational content of pointing is linguistic, rather than motoric/actional; the latter that pointing doesn't have any representational content at all. In particular, the first objection contends that, because the function of pointing is often associated to language, the meaning of the indicated objects is highlighted by the speech that accompanies the gesture. If this is true, then a symbolic meaning can be conveyed by pointing through the linguistic system alone, without involving the simulation of actions. The second objection contends that, unlike other gestures, a pointing index finger does not have any representational function in that it does not serve to simulate any manipulative or goal-oriented action.

Were these objections correct, pointing would be just a communicative, non-instrumental gesture; it could definitely affect the communicative context, but it would never play an active role in externalizing or enhancing the agent's thinking processes. Pointing wouldn't be a legitimate instance of GAH: it would hardly provide any representational contribution to non-linguistic tasks and, even if it did, its symbolic meaning wouldn't obtain from the representation of a motor action, as our hypothesis requires. But it seems that the assumptions supporting both objections are false: against the first objection, we stress that pointing is a form of re-enactment that actively creates or discloses new linguistic meanings, rather than passively mapping pre-existing linguistic meanings into the available indicated objects; against the second objection, we note that it is impossible to appropriately model the representational, content-productive function of pointing without recognizing that this gesture is a virtual re-enactement of a particular kind of motor action, i.e. eye movements whose goal consists exactly in controlling the observational, non-manipulative function of gazing behavior.

5. The role of pointing for language acquisition is constitutive, not derived

To respond to the first objection, we remark that - even if pointing largely occurs in communicative contexts, and is often accompanied by speech - its primary function can exist independently of both language and communication in general. On the contrary, pointing seems more of a precursor of language, and an important one, because it scaffolds joint attention in both development and evolution. Such coordination of attention is a way into declarative communication, not only its outcome (as suggested by the evidence that gestures facilitate speakers highlighting perceptually present information, see Alibali & Kita, 2010). On the other hand, language could hardly mature a conventionalized semantics without the mediation of a flexible and universal deictic designator of objects and actions, such as pointing: in development, reduction of pointing frequency is a symptom for the diagnose of severe impairments of communicative and social capabilities (American Psychiatric Association 1994); in evolution, non-verbal deixis did not derive from verbal deixis, but plausibly played a role in the acquisition of both joint attention and language (Leavens, Racine, & Hopkins, 2009). Without an appropriate triadic attentional scaffold that allows reciprocity, mutual knowledge, and nomination of individual objects, communication remains confined to asymmetric, affective, audience under-specified, vocalizations and bodily signals (Shepherd & Cappuccio, 2013, building on Tomasello, 2008). Pointing is one of the animal communicative signals with the highest potential to motivate this transition (Cappuccio & Shepherd, 2012).

Hence, pointing's function can't be linguistic at root because it precedes and sustains the very acquisition of language, scaffolding nomination in joint attention experiences that symbolically unite visual and vocal cues. This way, pointing highlights external representational contents (the objects that it makes publicly salient for mutual consideration, either in perception or in imagination) and makes them available as shared linguistic objects. Pointing contributes to nomination, but – like other gestures - it is also consistently produced in non-linguistic contexts to draw virtual objects in an imaginary space and help individual or shared cognitive processes related to symbolic visualization, memorization, and manipulation. For example, according to Hutchins (2005, 1567-68), the Micronesian navigators used to point to the stars to fixate the memory of an imaginary map of their travel, mentally superimposing certain key locations on the spatial landmarks provided by the star points. This way, highlighting the position of the stars as material anchors, the navigator "creates a model of the voyage that he can see and manipulate from his point of view on the deck of the canoe". In a rather similar way, according to Kirsh (2010b) Irish river dancers and choreographers use their hands and gestures, and occasionally pointing-like finger movements, to mark well-defined positions and iconically represent certain dancing steps during rehearsal: using their gestures as a small-scale replica of their full-body movements, they "create a simplified or abstracted version – a model" of the dance phrases they have to perform. In these cases, whether gestures are accompanied by speech or not, their role is to reorganize the representation of the physical space, transposing the relevant spatial cues from a situational to a positional framework.

This re-organization, highlighting relevant objects of individual or shared interests, and setting up a hierarchy of perspectively oriented levels of relevance in the perceptual experience, might become serviceable to associate pointed objects and linguistic contents, and fixate linguistic and social norms in either private or public memory. This is consistent with the view that gestures don't simply accompany, but also affect different aspects of linguistic production (Graham & Heywood, 1975; Rimé et al., 1984; Rauscher et al., 1996). If gestures can sustain language production this is probably because they may also have a non-linguistic function, in line with the Information Packaging Hypothesis:

> the process of informational organization is helped by representational gestures. It is helped because the production of representational gestures involves a different kind of thinking (namely, spatio-motoric thinking) from the default thinking for speaking (namely, analytic thinking). Spatio-motoric thinking provides alternative organizings of information that are not readily accessible via analytic thinking [...] The Information Packaging Hypothesis stipulates that what generates a gesture is spatio-motoric thinking, which has a raison d'être independent of speaking. Consequently, it is expected that gestures and speech have a certain degree of independence from each other (Kita, 2000, pp. 166-170).

Even before or without accompanying speech, pointing holds an important function in reorganizing our thoughts.

6. Pointing doesn't only produce, but represents gaze coordination

To respond to the second objection, we stress that pointing is a gesture with a peculiar representational function: namely, it doesn't represent immediate manipulative effects in the physical world, but attention-states and knowledge-states (including those of the agent who points). This is particularly evident in the declarative/informative use of pointing, when its expected communicative effects primarily depends on sharing some specific symbolic content, i.e. the common knowledge that the co-attenders share (Tomasello 2008).

Consider well-formed instances of declarative/informative pointing that successfully align the signaler's and the recipient's gaze direction: they do not only bring about actual coordination of attention between the signaler and the recipient(s), they concurrently impose a perceptually vivid and immediately shared representation of the very situation of coordination that they bring about. In other words, declarative/informative pointing is not just a strong visual stimulus to coordinate attention, but also one that symbolically represents the fact that attention is actively being coordinated by it (for a thorough analysis, see Cappuccio and Shepherd 2013, building on the concept of "self-involving situations" of joint attention proposed by Peacock 2005). This happens because pointing is not just an ostensive signal (one that displays a certain communicative intention, like other natural signals), but also a self-ostensive one. It is a very peculiar signal that represents its intention to display a certain communicative intention, thus stressing its declarative valence.

What is the specific representational contribution that pointing offers for cognitive self-stimulation (i.e., augmentation of one's own attentional and perceptual processes) processes? Since pointing vicariates the signaler's gaze direction and represents it in front of her, its primary representational function is to make visible in a stylized symbolic format what couldn't be otherwise visible to oneself in a pre-symbolic context: one's own eyes. This is how pointing scaffolds and speeds-up our processes of attention re-direction and fixation by actively representing them in front of us, not unlike other instrumental gestures do. Since ocular movements are legitimate actions, having the goal to re-direct vision and solicit attention shifts, it is true that there exists a goal-oriented action of which pointing is a symbolic model. However, such action is eye-based, not hand-based, and has contemplation as its goal, not physical manipulation. Additionally, often pointing is produced not only for the sake of controlling one's own attention, but concurrently to offer a public, openly visible, and socially effective representation of it that is symmetrically accessible by all the observers and simultaneously affect them: a "triadic virtual agent", a pre-formed template for gaze coordination and cooperation that, through development, is progressively employed to initiate the situation of joint attention according to the general model of ontogenetic ritualization (Tomasello, 1999).

However, it is crucial that also when aiming at social interaction pointing doesn't simply serve to directly produce gaze coordination, but also to modify the rich meaning of the very social context in which gaze coordination intervenes (Shepherd, 2010). Through pointing, the signaler enjoys the opportunity to manipulate not only the co-attender's gaze direction, but also her imagination, by displaying a symbolic model of their joint attention towards an object of common knowledge. This allows an infant to request symbolic forms of engagement in joint attention that are crucial for the acquisition of complex narratives and mentalistic skills through the feedback provided by the caregivers (Franco & Butterworth, 1996).

7. Pointing helps us think

Pointing is sufficient to modify the pointer's spatial perception and attention, and this explains why this gesture can and actually is employed in an instrumental way: to represent certain cognitive processes in order to monitor them and enhance them. Two experimental studies focusing on counting skills confirm that pointing primarily plays an instrumental role in facilitating numerical thinking, and that this facilitation relies on an actional competence, not a linguistic one: Alibali and DiRusso (1999) found that pointing gestures help children link sets of objects to their corresponding number words (when counting ″one″ ″two″ ″three″, etc.); Carlson et al., (2007) showed that adults are faster and more accurate in counting when they can point. The first study suggests that pointing plays a constitutive role, not just an expressive one, in articulating the early developments of analytic, attentional/perceptual, and linguistic operations: therefore it doesn't presuppose, but actively scaffolds the visuo-spatial manipulation of the corresponding linguistic/analytic contents. The second study suggests that the nature of this scaffold essentially relates to the motoric enactment of the pointing gesture: simple sequential fixation of the objects to be counted doesn't grant the same performances that are achieved when fixation is concurrently represented as an explicit symbol in front of the subject through his own index finger.

A notable aspect of this dynamic is that the pointing subject can reflectively control his own attentional processes through a duplicate of his ongoing observational activity. Pointing hence represents a stand-in for the corresponding series of acts of ocular redirection; the benefits received from monitoring these acts affect capabilities such keeping track of what has been counted, individuating objects, focusing on a particular object, anchoring number words to objects, with overall increased command of the abilities to visualize, double-check, re-organize, concentrate, and parse in time/space the task of counting.

Different hypotheses could explain why this facilitation occurs. A promising hypothesis is that the process of counting is not enhanced by the mere perception of the pointing finger, but also by the execution of its motor schema in correspondence with the observation of the counted object. Therefore, even without visual and proprioceptual feedback (even without perceiving one's own finger), the activation of the motor programs related to the relevant skills could be sufficient to encode the counted objects in first person body coordinates, facilitating the representation of their position in a manner that is easier to remember and manipulate. This possibility seems more than plausible, as two reports confirm that the motor programs for finger movements stored in the pre-motor cortex play an active role in counting tasks: a TMS study (Sato et al., 2007) correlates the excitability of hand muscles with representation of numerals; and a fMRI study (Tschentscher et al., 2011) shows that the motor system activations during mental counting are consistent with individual finger counting habits, and suggests an "intrinsic functional link between finger counting and number processing". Evidence of blind-born infants who spontaneously accompany their speech with pointing in abstract space (Iverson et al., 2000) suggests that these motor routines are consistently organized and fully operative early, in human development, and can be elicited regardless of the lack of visual feedback.

Is it possible that the performative-motor competence of pointing, more than the subject's perception of her own index finger, affects the perception of what pointing shows, loading it with a new symbolic meaning? This hypothesis seems in accord with Arbib (2005), which proposes that human embodied communication (gesture, pantomime, dance) derives its representational power from an advanced (human-specific) version of the "mirror system" that matches observed gestures with familiarly executed actions. It is significant that representational gestures have been associated to activity in the brain area where the mirror neuron system is located (for reviews see Cartmill et al., 2012; Willems & Hagoort, 2007): a subject internally simulates a model of the action that corresponds to the gesture he is observing; therefore – building on his actual motor expertise - he can immediately recognize the meaning of the represented action goal even if the pantomimed movements don't achieve any actual goal. This implies that motor competences (motor schemata of familiar actions) can be actively reused not only for intention understanding and action-goal recognition, but also for symbolic interpretation of imitational performances and pantomimic communication.

If Arbib's hypothesis is correct (regardless of its evolutionary ramifications) then also index pointing - like other instrumental gestures, and pantomime – might convey to its recipients a representational content that is always motoric and never just linguistic in nature. Such content is represented by the observing subject because, while he observes the instrumental gesture of pointing (possibly his own pointing), he mentally simulates in a virtual space the corresponding instrumental action of gaze/attention fixation and eyehead-posture re-configuration. This motor information might be fused with or mapped into the motor information necessary to produce finger-hand-arm-torso coordination in pointing: if this is the case, then the simulation of attention redirection is triggered by the very recognition of the agent's intention to point with the finger towards a distal target, and could work – for the observer (including the agent himself) - as a symbolic model that facilitates the recognition of any similar attentional redirection, the prediction of its effects, or the postdiction of its motivations.

It is perhaps too early to claim that an internal simulation of this kind mediates these cognitive tasks. But there is empirical evidence that the neuronal circuits in lateral intraparetial area dedicated to ocular redirection are recruited both for controlling one's ocular movements and for recognizing similar movements in the others', and that their causal contribution actively facilitates gaze following (Shepherd et al 2009). In other words, we know that the motor programs sub-serving ocular saccades, gaze orientation, and attention fixation, contribute to recognizing and interpreting the others' gaze behavior, discriminating among different social variables, and facilitating gaze following behaviors and attention coordination. If these motor programs were activated during the observation or production of pointing, this would suggest that the recognition of this gesture is accompanied by an internal simulation of the ocular movements in the direction indicated by pointing. This possibility seems plausible because the recognition of the direction of attention is partially processed by the same structures whether the direction is expressed by eyes or gestures: gaze shifts, arrows, and pointing fingers produce similar activations in the superior temporal sulcus area, and elicit similar brain activation in parietal, frontal, and occipital cortices (Sato et al., 2009; Materna et al., 2008).

8. Conclusions

Finger pointing significantly exemplifies how a gesture can modify an agent's spatial/perceptual experience, and enhance elements of his cognitive performance, even when it doesn't symbolize any physical manipulation. The cognitive benefits produced by the pointing gesture do not consist in anticipating the physical effects of a manipulative action (as in hand-rotations), but the consequences of the overall reconfiguration of the agent's posture, gaze orientation, and attentional attitude. This anticipation modifies her spatial experience and projects a symbolic meaning on it: by means of this reconfiguration, pointing draws a virtual line between far (observational) and near (interactional) space, and allows the subject to perceive the surrounding objects in a detached way, for example highlighting their intrinsic properties of being countable, or anchored to objective positions in an imaginary map. Subsequently, it can invite spatial reference, nomination, or visual imagination in the intersubjective space of position shared by the co-attenders who engage in joint attention.

It is at least in principle possible that in all these cases, and probably in others, the relevant action models manipulated by the pointing subject to facilitate his cognitive task are not only derived from the perception of the pointing gesture that he produced, but also from a concurrent simulation of the corresponding eye movements, mediated by the mirror neurons for gaze direction in the lateral intraparetial area. So, while a subject is intentionally pointing to help himself in a task of counting, or in anchoring a particular position to a mental map, it is possible that he is exploiting an internal simulation of his own gaze redirection. This allows him to project a particular valence (for example declarative/informative) to this gesture: phenomenologically, he doesn't simply fixate an object, he also actively dedicate his attention to the very fact that he is fixating it, explicitly and reflectively controlling his focus of attention. This might produce a qualitative reconfiguration of the spatial and bodily experience associated to attentional fixation toward a certain position/object targeted by the eyes, loading this very experience with symbolic meanings that are not reducible to direct perception. This is possible because we know that the internal structures recruited for the interpretation of the others' eye directions don't work in isolation, but tune their simulation function with the overall embodied and social background of one's experience.

GAH is a comprehensive and efficacious model to account for both the differences and the correlations between goal-oriented actions and instrumental gestures, and to explain the beneficial role that instrumental gestures play in cognitive tasks; we have argued that pointing is a particular case accounted for by this model, one that manifestly shows how the cognitive benefits of instrumental gestures do not always necessarily depend on representing physical manipulation.

ACKNOWLEDGEMENTS

The research of the first author was supported by a Fulbright Visiting Scholar Fellowship and developed in 2012 during a period of research visit at the University of Memphis.

REFERENCES

Alibali, M. W., DiRusso, A. A. (1999). The function of gesture in learning to count: more than keeping track. *Cognitive Development*, 14, 37-56.

Alibali, M. W., Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture*, 10(1), 3-28.

American Psychiatric Association (1994). *DSM-IV: Autistic disorder*. Washington, DC: American Psychiatric Association.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences,* 28(2):105–124.

Bates, E., L. Camaioni, and V. Volterra. 1975. Performatives Prior to Speech. *Merrill-Palmer Quarterly*, 21:205–226.

Beilock, S.L., Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science*. Advance online publication.

Broaders, S. C., Cook, S. W., Mitchell, Z., and Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology* 136. 539–50.

Bruner, J. (1983). *Child's Talk: Learning to Use Language*, New York: Norton.

Cappuccio, M., Shepherd, S. (2013, in press), Pointing Hand. Embodied Symbols and Joint Attention in Z., Radman (ed.), *The Hand. An Organon of the Mind*, MIT Press.

Cappuccio, M., Wheeler, M. W. (2011). The Sign of the Hand: Symbolic Practices and the Extended Mind, *Versus,* 113, 33–56.

Carlson, R., Avraamides, M., Cary, M., Strasberg, S. (2007). What Do the Hands Externalize in Simple Arithmetic? *Journal of Experimental Psychology: Learning Memory and Cognition*, 33(4), 747–756.

Carpendale, J.I.M., Carpendale, A.B., (2010). The development of pointing: From personal directedness to interpersonal direction. *Human Development*, 53, 110-126.

Cartmill, E. A., Beilock, S., and Goldin–Meadow, S. (2012), A word in the hand: action, gesture and mental representation in humans and non–human primates, *Phil. Trans. R. Soc.* B 367, 129–143.

Chu, M., Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, 137(4), 706–723.

Chu, M., Kita, S. (2011), The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General*, 140(1), 102–115.

Cole, J., Gallagher, S., McNeill, D. (2002). Gesture Following Deafferentation: A Phenomenologically Informed Experimental Study. *Phenomenology and the Cognitive Sciences* 1 (1), 49–67.

Ehrlich, S., Levine, S. C., Goldin–Meadow, S. (2006). The importance of gesture in children's spatial reasoning. *Developmental Psychology*, 42(6), 1259–1268.

Franco, F., and Butterworth, G., (1996). Pointing and Social Awareness: Declaring and Requesting in the Second Year. *Child Language*, 23, 307–336.

Goldstein, K., Gelb, A., (1918). Psychologische Analysen hirnpathologischer Falle auf Grund von Untersuchungen Hirnverletzer. *Zeitschrift fur die gesamte Neurologie und Psychiatrie*, 41, 1–142.

Goldin–Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*, Harvard University Press, Camb. MA.

Goldin–Meadow, S., Beilock, S. L. (2010). Action's influence on thought: The case of gesture. *Persepctives on Psychological Science*, 5, 664–674.

Goldin–Meadow, S., Nusbaum, H., Kelly, S., Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, 12, 516–522.

Goldin–Meadow, S., Cook, S. W., Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, 20(3), 267–272.

Graham, J. A., Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Psychology*, 5, 189–195

Grice, H. P. (1957). Meaning. *Philosophical Review*, 66, 377–388.

Hegarty, M. (2004). Mechanical reasoning as mental simulation. *Trends in Cognitive Sciences*, 8, 280–285.

Hostetter, A. B., Alibali, M. W. (2008). Visible embodiment: Gesture as simulated action. *Psychonomic Bulletin & Review*, 15(3), 495–514.

Hutchins, E. (2005). Material anchors for conceptual blends, *Journal of Pragmatics,* 37, 1555–1577.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science*, 2(4), 705–715.

Hutto, D. D., Myin, E. (2012), *Radicalizing Enactivism. Basic Minds without Content*, Cambridge MA: MIT Press.

Jacob, P., Jeannerod, M. (2003). *Ways of Seeing: The Scope and Limits of Visual Cognition.* Oxford: Oxford University Press

Langton, S.R.H., Bruce, V. (2000). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, 6(5), 541–567.

Leavens, D. A. (2004). Manual Deixis in Apes and Humans. *Interaction Studies,* 5, 387–408.

Leavens, D., Racine, T.P., and Hopkins, W.D. (2009). *The ontogeny and phylogeny of non–verbal deixis.* In: The prehistory of language. Oxford University Press, 142–165.

Materna, S., Dicke, P.W., Thier, P. (2008), The posterior superior temporal sulcus is involved in social communication not specific for the eyes. *Neuropsychologia* 11, 46, 2759–65

McNeill, D. (ed., 2000). *Gesture as Action*, University of Chicago Press.

McNeill, D., Bertenthal, B., Cole, J., Gallagher, S. (20050. Gesture-first, but no gestures? Commentary on Michael A. Arbib. Behavioral and Brain Sciences, 28, 138–39

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*, Paris: Gallimard, Phenomenology of Perception, trans. Donald A. Landes (New York: Routledge, 2012)

Milner, A. D., Goodale, M. A. (2006). The visual brain in action (2nd ed.). Oxford:

Oxford University Press.

Neppi–Mòdona, M., Rabuffetti, M., Folegatti, A., Ricci, R., Spinazzola, L., (2007). Bisecting lines with different tools in right brain damaged patients: The role of action programming and sensory feedback in modulating spatial remapping. *Cortex* Volume 43, Issue 3, 397–410.

Kelly, S.D. (2002). Merleau-Ponty on the Body, *Ratio* (new series) XV 4 December, 376–391.

Kirsh, D. (2010a). Thinking with external representations AI & Soc 25, 441–454.

Kirsh, D. (2010b). Thinking with the Body, Proceedings of the 32nd Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Peacocke, C. (2005). Joint Attention: Its Nature, Reflexivity, and Relation to Common Knowledge. In Elian, N., Hoerl, C., McCormack, T., and Roessler, J. *Joint Attention: Communication and Other Minds, Problems in Philosophy and Psychology,* Oxford University Press.

Piaget, J. (1968). Six psychological studies. New York: Random House.

Povinelli, D.J., Bering, J., Giambrone (2003). Chimpanzee ‘pointing’: Another error of the argument by analogy? In S. Kita (Ed.), 35–68, Pointing: Where language culture and cognition meet. Lawrence Erlbaum.

Schwartz, D. L., Black, J. B. (1996). Shuttling between depictive models and Abstract Rules: Induction and Fallback. *Cognitive Science*, 20(4), 457–497..

Rauscher, F. H., Krauss, R. M., Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, 7, 226–230.

Rimé, B., Schiaratura, L., Hupet, M., Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker′s nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion*, 8, 311–325.

Rizzolatti, G., Camarda, R., Fogassi, M., Gentilucci, M., Luppino, G., and Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Exp. Brain Res.* 71, 491–507.

Rowlands, M. (2006), *Body Language: Representing in Action,* Cambridge MA: MIT

Press.

Sato, M., Cattaneo, L., Rizzolatti, G., Gallese, V. (2007). Numbers within our hands: modulation of corticospinal excitability of hand muscles during numerical judgment. *J Cogn Neurosci.,* 19(4), 684–93.

Shepherd, S. V., Klein, J. T., Deaner, R. O., Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9489–9494.

Shepherd, S. V. (2010), Following Gaze: Gaze-Following Behavior as a Window into Social Cognition, *Frontiers in integrative neuroscience*: 4, March.

Shepherd, S. and Cappuccio, M. (2012), Sociality, Attention, and the Mind's eye. In Seemann, A.(ed.), *Joint attention. Perspectives and Developments*, Cambridge MA: MIT Press 2012.

Schwartz, D. L., Holton, D. L. (2000). Tool use and the effect of action on the imagination. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1655–1665.

Sekiyama, K. (1982). Kinesthetic aspects of mental representations in the identification of left and right hands. *Perception & Psychophysics*, 32, 89 – 95.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition.* Cambridge MA: Harvard University Press.

Tomasello, M. (2008). *Origins of Human Communication.* Boston, MA: MIT Press.

Tschentscher, N, Hauk, O., Fischer, M. H., Pulvermüller, F. (2012). You can count on the motor cortex: finger counting habits modulate motor cortex activation evoked by numbers. *Neuroimage.* Feb 15;59(4),3139–48.

Vygotsky, L. S. (1962/1986). *Thought and Language* (A. Kozulin 1962, trans.). Cambridge, MA: MIT Press.

Wagner, S.M., Nusbaum, H.C., Goldin-Meadow, S. (2004). Probing the Mental Representation of Gesture: Is Handwaving Spatial? *Journal of Memory and Language*, 50, 395–407.

Werner, J., Kaplan, B. (1963). *Symbol formation.* New York: Wiley.

Wexler, M., Kosslyn, S. M., Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, 68, 77–94.

Wheeler, M. W. (2005), *Reconstructing the cognitive world: The next step.* Cambridge, MA: MIT Press.

Willems, R. M., Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain Lang.,* 101(3), 278–289.