Sensorimotor signal mixer as a proposal for explaining pre-conceptual representation

Mauricio Pedroza-Torres [1] and Roberto Aguirre[1]

[1] *Center for Basic Research in Psychology, Universidad de la República, Tristán Narvaja 1674, Montevideo, Uruguay*

Abstract

The construction of concepts has always been a topic of research interest for the Cognitive Sciences and Artificial Intelligence. In this sense, some theoretical approaches to cognition advocate the need to define sets of pre-conceptual structures as the necessary basis for the formation of concepts of greater complexity or level of abstraction. When considering the internal representation of concrete elements as a starting point, the Recognition-by-Component theory postulates that the base blocks for the mental representation of objects would be a set of volumetric primitives such as spheres, cylinders, parallelepipeds called geons. Such theory describes the properties of these primitives, but without considering the leading role that signals to come from sensorimotor areas independent of vision could have in the formation of these mental representations. This research postulates the analogy of a sensorimotor signal mixer as a possible origin for internal representations of volumetric figures, deepening its genesis. For doing it, we carried out a case study on a set of 48 objects from everyday context, used in recognition of 3 volumetric primitives (sphere, cylinder, and parallelepiped). The results highlight the leading role of 3 sensorimotor signals: (i) the recognition of a second geometric contour, (ii) the texture of the object, and (iii) the hand proprioception required to interact.

Keywords

Objects recognition, Geon, Volumetric primitives, Signal mixer, Case study

1. Introduction

Language has historically been considered a central element in the origin of human civilization, understood as scaffolding for the social organization of forms, laws, and the content of thought [1, 2]. Additionally, natural language emerges as a strictly human capacity [3, 4, 5, 6]. The mind of human newborns develops to refer to and conceive objects as external to themselves [7, 8, 9]. Latter, as a prerequisite for the construction of concepts. According to the proposal that the concrete always precedes the abstract, this work seeks to delve into the mental genesis of volumetric primitives (geons). Our approach considers these primitives as fundamental representations, both for developing preconceptual structures such as image schemes [10] or perceptual symbols [11] and the subsequent formation of concrete and abstract concepts and their labeling in natural language [12, 13, 14, 15].

The hypothesis here conceives the geon as the output signal of a mixture of multiple interacting sensorimotor stimuli. This output signal is hosted by a neural structure (signal mixer) that, due to the refinement of everyday experience and presumably by a Hebbian learning mechanism [16], produces an element consistent enough to be valid as an internal model of representation [17]. As for the research

The Eighth Joint Ontology Workshops (JOWO'22), August 15-19, 2022, Jönköping University, Sweden
EMAIL: ptorres.pro@gmail.com (M. Pedroza-Torres); zimmer20uy@gmail.com (R. Aguirre)
ORCID: 0000-0002-5803-7637 (M. Pedroza-Torres); 0000-0001-8386-1380 (R. Aguirre)

approach, we propose the selection of a subset of representative sensorimotor stimuli in the interaction with elements of the environment. Then, through a case study, we weigh their relevance in the emergence of volumetric primitives (sphere, cylinder, or parallelepiped type).

2. A hypothesis about the origin of geons

2.1 What is a geon?

In daily experience, visual apparatus projects flat and volumetric objects. Meanwhile, the perceived shape will change gradually according to the relative movement between observer and object. However, interpreting different stimuli from the same object as separate entities does not seem to be a privileged characteristic by evolution [18].

In perceptual object recognition, the RBC (Recognition-By-Components) theory [19] asserts that input images are segmented into a set of simple geometric components called geons (for "geometrical ions"), such as bricks, cylinders, wedges, cones, and squashes, among others. Furthermore, these volumetric primitives are expected to have high viewpoint invariance, thus facilitating their recognition [20].

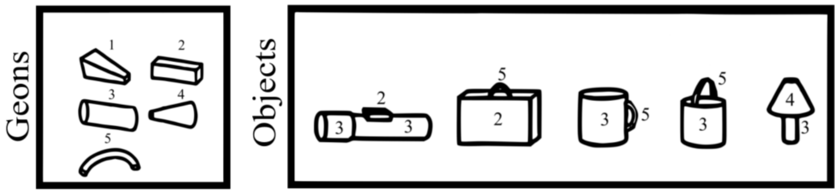

Figure 1: Biederman geons and objects composition

2.2 Geons and object recognition

Summarizing the main aspects of object recognition models, it is possible to address them from two sequential stages: 1) Pre-processing and 2) Recognition.

In the pre-processing stage, Biederman analyzes its relationship with biological models in object recognition [20], extracting the following main ideas:

1. All models of biological object recognition assume an input layer that can be approximated by a lattice of filters covering the visual field.
2. Some simple filters (typically modeled as a Gaussian damped sinusoid, termed a Gabor filter) occupied each node in the lattice.
3. Each filter is “tuned” to a particular orientation and spatial frequency at that position in the visual field.
4. At nearby nodes, there is considerable overlap in the coverage of the filters.
5. Each Gabor filter corresponds to a simple cell, and each node corresponds to a single hypercolumn in the initial region of the cortex that receives visual inputs (area VI).
6. A layer in the visual cortex (called the object layer) contains the units corresponding to the various object categories that the network is supposed to differentiate.

Meanwhile, regarding the recognition stage, the proposals are diverse. Some of the main modeling alternatives are:

- Theories that support the idea of mapping the output of the filters directly onto the object layer [21].
- Proposals that allow intermediate representations such as lines, surfaces [22] or assume construction of geometric primitives (Geon theory) [19].

- Approaches that conceive a coordinate space that preserves retinal proximities for the matching of input against stored representation [23].

Taking as reference the geon theory, it supports the idea of representation as a structural description, consisting of elements (such as parts) [24], and this paper seeks to propose a functional scheme suitable for such internal representations.

2.3 The hypothesis of a signal mixer

According to Marr [25], the previous and better-understood stage of recognizing contours comes before recognizing volumes. Marr stated that questions of psychological interest could be illuminated and perhaps even explained by neurophysiological terms. Contour recognition is a neural functionality of high evolutionary value because it is part of the vision as one of the senses (the other one is hearing) with the minor proximity requirement for categorizing environmental elements. Although there are other strategic stimuli for human survival, such as sound and color in the auditory and visual channels, respectively, the identification of contours plays a leading role in the recognition of food, predators, containers, sexual partners, and shelter.

When considering contour recognition as a privileged information channel, it would be expected that this fact will materialize, either in the spatial distribution of this function throughout the neural tissue or in the density of connections relative to its operation. In either of these two scenarios, the probability of interaction with other stimuli coming from different channels increased [20]. The present hypothesis formulates that the interaction of these multiple signals led to a neural structure of composition (or mixture) subordinated to contour recognition (as the main signal) but integrative (by bringing together the effect of contingent signals). This neural structure of composition would be the material substrate of geon.

According to Barlow [26], we can provide arguments for promoting the possibility that such a signal composition mechanism takes place in the neural tissue. Barlows expresses it as follows:

> Neurons do not loosely and unreliably remap the luminous intensities of the visual image onto our sensorium. Instead, they detect pattern elements, discriminate the depth of objects, ignore irrelevant causes of variation, and are arranged in an intriguing hierarchy. Furthermore, there is evidence that they give prominence to what is informationally important, can respond with outstanding reliability, and can have their pattern selectivity permanently modified by early visual experience. These patterns amount to a revolution in our outlook. It is now entirely inappropriate to regard unit acts as a noisy indication of more fundamental and reliable processes involved in mental operations. Instead, we must consider single neurons as the prime movers of these mechanisms. Neurons bring about thinking, and we should not use phrases like "unit activity reflects, reveals, or monitors thought processes" because the activities of neurons, quite simply, are thought processes. (p. 380).

We propose to think of the formation of geon as an underlying Hebbian learning process, in which simultaneous activation of cells leads to pronounced increases in synaptic strength between those cells, what Hebb [16] stated as follows:

> Let us assume that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability. ... When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased. (p. 62)

Figure 2 presents the analogy of geon as a composite signal. This figure shows the interaction of ten sensorimotor channels that compose an output signal (tracks initially assumed to have equal synaptic weight). The output signal (or geon) will be continuously subjected to a refinement process given by the subject's access to new experiences. So, the most consistent channels (channels with fewer novel signals) will take synaptic prominence in geon formation.

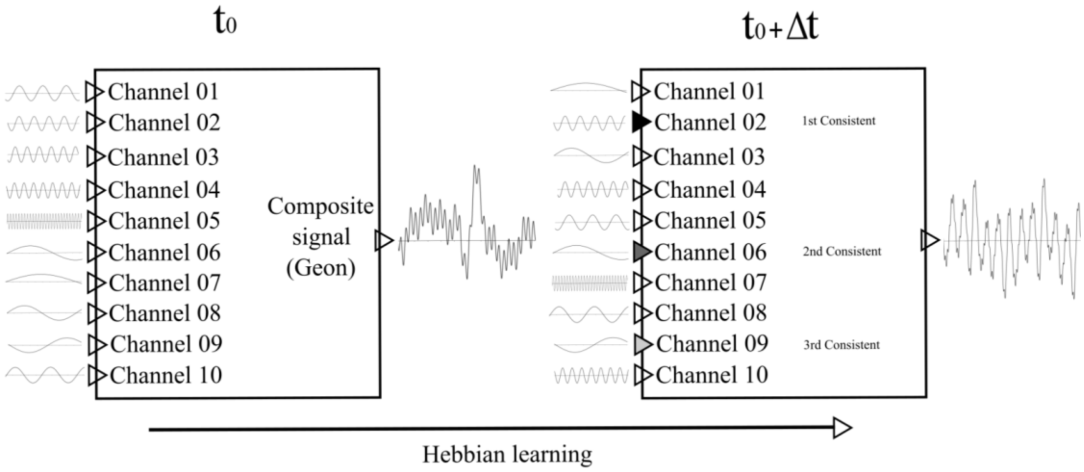

Figure 2: Neural substrate of geon as a signal mixer. In the example shown, the time course and Hebbian learning have favored (over the entire set of inputs) in descending order of importance the signals coming from channels 2, 6 and 9

3. Case study

As a case study, the interaction of a single observer (Male, 32 years old, right-handed) with 48 random everyday objects, 16 of a spheroidal nature (Figure 3), 16 cylindrical, and 16 parallelepipeds was documented.

TRIAL	OBJECT	DETAIL	COLOUR	2ND CONTOUR	TASTE	ODOR	SOUND	TEXTURE	HARDNESS	THERMAL STATE	HAND PROPIOCEPTION	BODY PROPIOCEPTION
1	Red ball	Plastic / diameter 10 cm	Red	Circular	NSTST01	NSODR01	Impact	Slippery	ShoreD 70	NSTHM01	Power Sphere	NSBPC01
2	Blue ball	Rubber / diameter 3 cm	Blue	Circular	NSTST02	Rubber	Impact	Vibrating	ShoreA 40	NSTHM02	Sphere 4-finger	NSBPC02
3	Orange	Fruit / diameter 9 cm	Orange	Fairly circular	Orange	Orange	NSSND01	Slippery	ShoreA 50	NSTHM03	Power Sphere	NSBPC03
4	Macadamia nut	Nut / diameter 1 cm	White-yellow	Fairly circular	Macadamia	Macadamia	NSSND02	Slippery	ShoreD 30	NSTHM04	Tripod	NSBPC04
5	Soccer ball	Leather / diameter 22cm	White-black	Circular	NSTST03	Leather	Impact	Vibrating	ShoreD 10	NSTHM05	Two-hand palmar	NSBPC05
6	Green apple	Fruit / diameter 9 cm	Green	NSCTR01	Apple	Apple	NSSND03	Slippery	ShoreA 50	NSTHM06	Power Sphere	NSBPC06
7	Lollipop	Candy / diameter 2 cm	Red	Fairly circular	Lollipop	NSODR02	NSSND04	Highly	ShoreD 20	NSTHM07	Stick	NSBPC07
8	Lamp	Light / diameter 30 cm	White-orange	Fairly circular	NSTST04	NSODR03	NSSND05	NSTXT01	NSHRD01	NSTHM08	NSHPC01	NSBPC08
9	Grape	Fruit / diameter 1 cm	Purple-black	Fairly circular	Grape	NSODR04	NSSND06	Slippery	ShoreA 20	NSTHM09	Tripod	NSBPC09
10	Party balloon	Rubber / diameter 20 cm	Blue	Ovoid	NSTST05	Rubber	Balloon	Vibrating	Shore00 40	NSTHM10	Two-hand finger pressure	NSBPC10
11	Tangerine	Fruit / diameter 6 cm	Orange	Ovoid	Tangerine	Tangerine	NSSND07	Slippery	ShoreA 50	NSTHM11	Power Sphere	NSBPC11
12	Peeled potato	Vegetable / diameter 5	White-yellow	Roughly	Potato	NSODR05	NSSND08	Highly	Shore00 40	Warm	Sphere 4-finger	NSBPC12
13	Monument	Concrete / diameter 6 m	Grey	Circular	NSTST06	NSODR05	NSSND09	NSTXT02	NSHRD02	NSTHM12	NSHPC02	NSBPC13
14	Moon	Full moon	White-grey	Circular	NSTST07	NSODR06	NSSND10	NSTXT03	NSHRD03	NSTHM13	NSHPC03	NSBPC14
15	Sun	Sunset sun	Orange-red	Circular	NSTST08	NSODR07	NSSND11	NSTXT04	NSHRD04	Warm	NSHPC04	NSBPC15
16	Bubble	Soap bubble / diameter 3	Transparent	Circular	Soap	Soap	Bubble burst	NSTXT05	NSHRD05	NSTHM14	Precision sphere	NSBPC16

Figure 3: Experiment with spherical objects. The shaded cells in the figures are related to signals of less novelty, whereas the column with the most significant number of shaded cells would represent a channel of greater consistency.

In this proof of concept, the simultaneous interaction of stimuli involved in the recognition process was restricted to the following ten interacting channels in the signal mixer (whose possible values are shown in Table 1):

1. COLOUR: 12 colors, their possible combinations in pairs plus non-specific cases
2. 2ND CONTOUR: 17 contour values plus non-specific cases. It is worth noting that the design of the experiment takes the 1st contour as the primary evocation signal of the geon ("circular" as the central contour of sphere and cylinder, "rectangular" as the 1st contour of parallelepiped), while the 2nd contour would be the geometric primitive recognized when manipulating the angle of vision on the figure (for example, when rotating the sphere it will go from the 1st

contour (circular) to a 2nd contour (circular again), while in the cylinder it will go from a 1st contour (circle) to a 2nd contour (rectangular))

3. TASTE: Material or food-specific plus non-specific cases
4. ODOR: Material or food-specific plus non-specific cases
5. SOUND: Material or functionality specific plus non-specific cases
6. TEXTURE: 6 texture values [27] plus non-specific cases
7. HARDNESS: 21 hardness values [28] plus non-specific cases
8. THERMAL STATE: Material or context-specific plus non-specific cases
9. HAND PROPRIOCEPTION: 36 hand postures [29] plus non-specific cases
10. BODY PROPRIOCEPTION: Functionality specific plus non-specific cases

Table 1
Hypothetical channels and their possible values

COLOUR	2ND CONTOUR	HARDNESS	HAND PROPIOCEPTION	
Red	Circular	Shore00 10	Large diameter	Quadpod
Orange	Fairly circular	Shore00 20	Small diameter	Sphere 3-finger
Yellow	Roughly circular	Shore00 30	Medium Wrap	Stick
Green	Ovoid	Shore00 40	Adducted thumb	Palmar
Blue	Flat ovoid	Shore00 50	Light tool	Ring
Purple	Elliptical	ShoreA 10	Prismatic 4-finger	Ventral
White	Flat elliptical	ShoreA 20	Prismatic 3-finger	Inferior pincer
Grey	Triangular	ShoreA 30	Prismatic 2-finger	Two-hand palmar pressure
Black	Fairly triangular	ShoreA 40	Palmar pinch	Two-hand finger pressure
Brown	Roughly triangular	ShoreA 50	Power disk	Two-hand square support
Transparent	5 or plus sided polygon	ShoreA 60	Power sphere	***NSHPC
Golden	Roughly 5 or plus sided polygon	ShoreD 10	Precision disk	
*Combination in pairs	Rectangular	ShoreD 20	Precision sphere	
***NSCLR	Fairly rectangular	ShoreD 30	Tripod	
TASTE	Roughly rectangular	ShoreD 40	Fixed hook	
**Material or food specific	Thick line	ShoreD 50	Lateral	
***NSTST	Thin line	ShoreD 60	Index finger extension	
ODOR	***NSCTR	ShoreD 70	Extension type	
**Material or food specific	TEXTURE	ShoreD 80	Distal	
***NSODR	Highly slippery	ShoreD 90	Writing tripod	
SOUND	Slippery	ShoreD 100	Tripod variation	
**Material or functionality specific	Vibrating	***NSHRD	Parallel extension	
***NSSND	Highly vibrating	BODY PROPIOCEPTION	Adduction grip	
THERMAL STATE	Rough	**Functionality specific	Tip pinch	
**Material or context specific	Highly rough	***NSBPC	Lateral tripod	
***NSTHM	***NSTXT		Sphere 4-finger	

Notes:
* The combination or presence of color pairs is considered without specifying their proportion
** Signals highly dependent on the object and difficult to locate on an orderly scale
*** NSXXX stands for a nonspecific stimulus signal

4. Results and discussion

Proceeding to count the number of Novel Signals per Channel (NSC) and defining a Low Plurality Estimator (LPE) as follows:

$$LPE = \frac{1}{NSC} \quad (1)$$

In Table 2, it is possible to evidence the channels of greater consistency (Higher LPE), those of less relevance for the formation of a volumetric primitive, and the Most Repetitive Signal (MRS) resulting from the case study.

Table 2

Estimation of channels of high relevance for geon formation

		COLOUR	2ND CONTOUR	TASTE	ODOR	SOUND	TEXTURE	RIGIDITY	THERMAL STATE	HAND PROPIOCEPTION	BODY PROPIOCEPTION
Spherical objects	NSC	14	5	16	15	16	8	13	15	11	16
	LPE	0.07	0.20	0.06	0.07	0.06	0.13	0.08	0.07	0.09	0.06
	MRS		1) Circular				2) Slippery			3) Power sphere	
Cylindrical objects	NSC	13	5	16	16	15	6	10	14	10	16
	LPE	0.08	0.20	0.06	0.06	0.07	0.17	0.10	0.07	0.10	0.06
	MRS		1) Roughly rectangular				2) Slippery			3) Large diameter	
Parallelepiped objects	NSC	11	2	16	15	15	4	9	14	8	16
	LPE	0.09	0.50	0.06	0.07	0.07	0.25	0.11	0.07	0.13	0.06
	MRS		1) Fairly rectangular				2) Slippery			3) Two-hand palmar pressure	

The most relevant channels for geon formation in descending order of importance were: 2nd contour, texture, and hand proprioception.

Firstly, the second contour rose in all three cases as the most representative channel in the formation of geon. Someone could argue that the novelty in the signal would tend to be elevated since each variation in the angle would offer a different signal for the second contour. However, it is here where the RBC theory provides theoretical support for invariant detection of properties of edges such as curvature, collinearity, symmetry, parallelism, and cotermination in a two-dimensional image [19].

Second, the texture channel becomes the next in importance, giving clues regarding the common association humans establish between smooth (or slippery) surfaces and perfect volumetric shapes.

Finally, hand proprioception appeared as the third channel in relevance. In each case, it associates with the most common signal for manipulation of such shapes in a daily environment: Power sphere grasp (Sphere), Large diameter grasp (Cylinder), and Two-hand palmar pressure grasp (Parallelepiped). Moreover, it showed a particular link with the gestural language used by a human trying to describe a given volumetric primitive.

5. Conclusions

This paper shows the hypothesis of a sensorimotor signal mixer as a formal explanation of the processing of pre-conceptual structures. In the Cognitive Sciences of the last 40 years, this proposal is relevant because the evidence about the role of sensorimotor information in processing concepts [12, 14], and the emergency of approaches into the frame of embodied cognition [10, 11]. The theoretical explanation was developed in three stages: (i) the main role of concrete objects in the construction of concepts, (ii) the RBC as alternative for explaining the perceptual object recognition, and (iii) a hypothetical signal mixer capable of composing geons, as an output signal from channels of cognitive relevance in the experience of the conceptualizer.

The paper exposes a case study for illustrating and testing how the signal mixer might work. As a results, a set of sensorimotor channels related to tactile experience add to contour recognition as prominent channels in preconceptual processing. The results suggests that a signal mixer model might provide an adequate formal explanation for being more deeply explored and elaborated.

References

[1] L. H. Morgan, Ancient Society. MacMillan & Company, London, 1877.

[2] B. L. Whorf. Language, Thought and Reality, MIT Press, Cambridge, MA, 1956.

[3] J. A. Fodor, The Language of Thought, Harvard University Press, Cambridge, 1975.
[4] R. C. Berwick, N. Chomsky, Why only us: Language and evolution, MIT Press, 2016.
[5] N. Chomsky, Syntactic structures, Mouton, The Hague, 1975.
[6] D. Bickerton, Language and species, Chicago University Press, Chicago, 1990.
[7] J. Piaget, Intellectual evolution from adolescence to adulthood, Human development 15.1 (1972) 1-12.
[8] S. Carey, E. M. Markman, Cognitive development, in: Cognitive Science, 201-254, 1999.
[9] M. Tomasello, Culture and cognitive development. Current Directions in Psychological Science 9.2 (2000) 37-40.
[10] M. M. Hedblom, Image Schemas and Concept Invention Cognitive, Logical, and Linguistic Investigations, Springer Nature, Switzerland, 2020.
[11] L. W. Barsalou, Perceptual symbol systems. Behavioral and Brain Sciences 22.4 (1999) 577-660.
[12] A. M. Borghi, F. Binkofski, C. Castelfranchi, F. Cimatti, F. C. Scorolli,, L. Tummolini, The challenge of abstract concepts. Psychological Bulletin, 143.3 (2017) 263.
[13] D. Casasanto, R. Bottini, Spatial language and abstract concepts. Wiley Interdisciplinary Reviews: Cognitive Science 5.2 (2014) 139-149.
[14] A. M. Borghi, A future of words: Language and the challenge of abstract concepts, Journal of Cognition 3.1 (2020).
[15] G. Jakus, V. Milutinović, S. Omerović, S. Tomažič, Concepts, Ontologies, and Knowledge Representation, Springer Briefs in Computer Science, 2013.
[16] D. O. Hebb, The Organization of Behavior, Wiley & Sons, New York, 1949.
[17] J. J. Koenderink, A. J. van Doom, The internal representation of solid shape with respect to vision, Biological Cybernetics 32 (1979) 211-216.
[18] K. Kirkpatrick-Steger, E. A. Wasserman, I. Biederman, Effects of geon deletion, scrambling, and movement on picture recognition in pigeons, Journal of Experimental Psychology: Animal Behavior Processes 24.1 (1998) 34-46. doi:10.1037/0097-7403.24.1.34.
[19] I. Biederman, Recognition-by-Components: A Theory of Human Image Understanding Psychological Review 94.2 (1987) 115-147.
[20] I. Biederman, Geon theory as an account of shape recognition in mind and brain, The Irish Journal of Psychology 14.3, 314-327, (1993). doi: 10.1080/03033910.1993.10557936.
[21] J. Buhmann, J. Lange, C. von der Malsburg, J. C. Vorbruggen, R. P. Wurtz, Object recognition in the dynamic link architecture: Parallel implementation of a transputer network, in: Neural Networks for Signal Processing, 121-159, 1991
[22] D. G. Lowe, The viewpoint consistency constraint, International Journal of Computer Vision 1 (1987) 57-72.
[23] T. Poggio, S. Edelman, A network that learns to recognize 3-dimensional objects, Nature 343 (1990) 263-266.
[24] J. E. Hummel, I. Biederman, Dynamic binding in a neural network for shape recognition, Psychological Review 99 (1992) 480-517.
[25] D. Marr, Vision, San Francisco, 1982.
[26] H. B. Barlow, Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology?, Perception 1.4 (1972) 371–394. doi:10.1068/p010371.
[27] J. Faucheu, B. Weiland, M. Mathieu, A. Witt, P. Cornuault, Tactile aesthetics : textures that we like or hate to touch, Acta Psychologica, Elsevier, 2019.
[28] B. Schwartz, Hardness of Rubber Rollers, Imperial Rubber Products INC, 2012.
[29] T. Feix, J. Romero, H. B. Schmiedmayer, A. M. Dollar, D. Kragic, The GRASP Taxonomy of Human Grasp Types, Human-Machine Systems, IEEE Transactions, 2015.