Visual spatial learning of complex object morphologies through the interaction with virtual and real-world data

Chiara Silvestri, René Motro, Bernard Maurin and *Birgitta Dresp-Langley*,
Laboratoire de Mécanique et Génie Civil (LMGC), Université Montpellier 2,
CC048 Place Eugène Bataillon, 34095 Montpellier Cedex 5, France

Conceptual design relies on extensive manipulation of the morphological properties of real or virtual objects. This study aims to investigate the nature of the perceptual information that could be retrieved in different representation modalities to learn a complex structure. An abstract and complex object was presented to two study populations, experts and non-experts, in three different representation modalities: 2D view; digital 3D model; real object. After viewing, observers had to draw some parts of the structure into a 2D reference frame. The results reveal a considerable performance advantage of digital 3D compared with real 3D, especially in the expert population. The results are discussed in terms of the nature of the morphological cues made available in the different representation modalities.

Keywords: design tools, perception, problem solving

In the digital age, the role of visual perceptual learning and communication is becoming more and more important. Computer driven image technology and virtual reality devices have replaced former written or spoken media of information exchange in a wide range of public and private domains, such as education, healthcare, and navigation (Katz, Kripalani, & Weiss, 2006; Stelzer & Wickens, 2006). In the fields of architecture, engineering and industrial design, digital modelling and visualisation tools are nowadays fundamental in most practices. In this context, virtual data have largely replaced former media of real-world learning, like hand drawings or scale models. Moreover, digital tools are at the heart of a new way of approaching the design process itself. Expressions like 'digital design' or 'digital architecture' are often used to describe contemporary architectural practices, because they extensively rely on digital tools for conceptual design, production and construction (Borgart & Kocaturk, 2007; Kolarevic, 2003; Oxman, 2006). In particular, numerical tools allow modelling and visualisation of complex shapes, difficult to represent by hand drawing through descriptive geometry. In the most advanced tools, geometrical modelling is based on manipulations of shapes

Corresponding author:
Chiara Silvestri
silvestri@lmgc.univ-montp2.fr

www.elsevier.com/locate/destud
0142-694X $ - see front matter *Design Studies* **31** (2010) 363–381
doi:10.1016/j.destud.2010.03.001

visualised on the screen; analytical information is becoming useless and is managed exclusively by the ‘black box’ of software’s algorithms. In this context, visual elaboration of 2D (on screen) spatial and formal qualities is fundamental, since these 2D configurations are the only accessible data. The conceptual design of new and complex object structures in the physical world relies on the designer’s learnt ability to effectively manipulate 3D representations. Digital 3D modelling, which is based on 2D screen communication to learn and manage formal properties of objects, raises deep questions about the consequent construction of mental representations of 3D qualities.

Only little is still known about the visual perceptual processes through which virtual world data may guide information processing in comparison with real-world data, especially in the domain of architectural design. ‘Virtual reality’ is often referred to in terms of ‘augmented reality’ because it reproduces enriched visual perceptual environments that are free from the information capacity constraints of the real world (Darken & Sibert, 1996). Experimental investigations into cognitive processes of learning and memory (Matheis, Schultheis, Tiersky, DeLuca, Mills, & Rizzo, 2007) have shown that learning through virtual 3D images is faster than learning through real-world objects. Real-world is limited by the constraint of partial observation in space and time, and therefore more heavily solicit the experience and prior knowledge of the learning individual. Indeed, for any limited length of time, the virtual medium may allow the generation of a much greater quantity of potentially relevant visual information compared with a real-world medium, especially in the case of architectural design (Borgart & Kocaturk, 2007). Experimental studies already demonstrated that a virtual 3D model is able to provide critical visual information to both experts and novice observers about the geometrical properties of objects, for example in the domain of curvature perception (Dresp, Silvestri, & Motro, 2007).

The objective of the work presented in this paper is therefore to improve the understanding on the cognitive processing of complex spatial information, particularly for different representation modalities and in relationship with the specific expertise of designers. Therefore, we performed a problem solving experiment to compare the performances on perceptual learning (shape reconstruction) of a complex and abstract object, explored either actively (physical object) or virtually (numerical model) by observers having different levels of expertise in geometry and numerical modelling. The chosen study object is considered as ‘unknown’, meaning it is not a familiar object (it is difficult to identify it, to give it a function, etc.); consequently the information used for its spatial processing is limited to morphological properties. The subjects have never seen this object before; they learn it during the experiment, the aim of our study being the understanding of this learning process through the resolution of a spatial problem solving task. The purpose is not the analysis of local aspects discrimination or identification of features; it is the understanding of global shape processing and, in particular, of the role of geometry based perceptual descriptors which are used to build mental representations of complex spatial configurations.

1 Perceptual learning

To analyse how the human perceptual system processes the visual spatial configuration of objects, numerous experimental studies have compared the precision and speed with which human observers recognise local aspects of 3D shapes under varying conditions of presentation or information given (see Norman, Norman, & Clayton, 2004, for an extensive review). These experiments offer important data about the perceptual early processing of spatial qualities. To analyse and study the perceptual learning processing in its whole it is however necessary to approach the problem from a more global position. Concepts elaborated in the field of Artificial Intelligence may be a useful reference for modelling complex cognitive processing.

1.1 Perceptual learning by agent–environment interaction

The perceptual learning of new visual spatial information can be approached on the basis of concepts of machine learning and skill acquisition by agents, such as the representation matching (Carpenter & Grossberg, 1991; Whitehead & Ballard, 1991) or the concept of partially observable worlds (e.g. Singh & Sutton, 1996). It is considered that any agent, man or machine, is capable of learning from its environment, virtual or real, on the basis of perceptual and entirely non-verbal actions.

A perceptual action describes a formal or mental operation/transformation through which states of the real or virtual world are matched to observations of the latter. A matching operation usually triggers an action, such as moving the cursor to a new position on the computer screen in a virtual reality situation, or pulling a handle to open a drawer of a filing cabinet in a real-world scenario. In the context of the present study, two types of matching operations will be considered: 'one-to-one' matching and 'many-to-one' matching (Whitehead & Ballard, 1991).

In one-to-one matching, a unique observation is matched to a unique world-state: this could correspond to the matching of a drawing with the real object it represents (Figure 1, top). In this case, a direct matching is possible since the visual spatial information does not need to be transformed to successfully perform the match. A single criterion (like the global shape of the object) or several criteria (like the shape, the colour and the texture) may be necessary.

In many-to-one matching, several observations are matched to a unique world-state: this could correspond to the matching of different views with a unique object (Figure 1, bottom). In this case, a direct match is not possible since the visual spatial information need to be transformed and elaborated. The criteria for a many-to-one match may be multiple (shape, colour, texture,

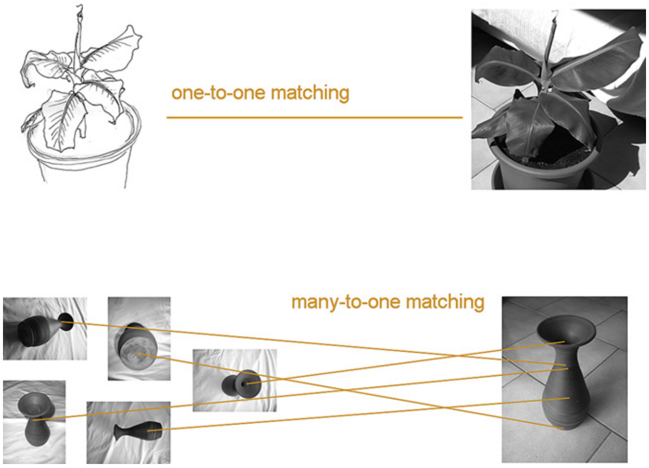

Figure 1 Representation matching: in 'one-to-one' matching (top) a unique observation is matched with a unique world-state; in 'many-to-one' matching (bottom) many observations are matched to a unique world-state

etc.) or single, such as in the case of a match on the sole basis of the global shape. Usually, successfully matching many observations to a single world-state requires knowledge (expertise) of appropriate criteria. More the matching is complex, more expertise is necessary to successfully perform the match.

Matching operations of a similar kind are involved in the basis of the high-level perceptual learning processes through which knowledge about spatial configuration of complex visual objects is made accessible in human cognition. The criteria for matching world-states represented in working memory to new observations, the so-called perceptual learning criteria, would correspond to what researchers in the field of machine learning have referred to as eligibility traces (Singh & Sutton, 1996). Eligibility traces are working memory data with a specific heuristic or diagnostic value, corresponding to representations of simple or complex events, transformations, or actions. Inevitably, the eligibility traces for working memory operations during perceptual learning of a non-familiar and abstract visual object are essentially morphological.

1.2 Virtual viewing and exploration by vision and touch

A potential factor in visual spatial processing of structure is the sensory modality through which the structure is explored. Earlier views, such those of Gibson (1962, 1963, 1966), considered visual and tactile exploration as equivalent media which make essentially the same kind of information available to an observer. Consequently, visual-plus-tactile exploration of a novel object would not provide more information than purely visual exploration. Other recent studies suggest that this may not necessarily be the case, like in shape recognition experiments where observers were found to recognise target shapes presented among other shapes significantly better when the target was previously explored visually and by hand rather than having been viewed only (Norman et al., 2004). Structural information processing may therefore be facilitated when observers are able to not only see but also touch and explore an object by hand.

Viewpoint dependence of spatial information processing in object recognition has been also explored with familiar objects: visual recognition was best when objects were viewed from the 'front', and tactile recognition was best when the 'back' was explored manually (Newell, Ernst, Tjan, & Bülthoff, 2001). The axis of rotation appears to be a critical factor determining viewpoint dependence in the visual modality. Viewpoint dependence is however abolished by perceptual learning, or through repeated interaction with virtual 3D environments (Christou & Bulthoff, 1999). Recent research shows that visual–tactile recognition could be viewpoint independent for non-familiar objects (Lacey, Peters, & Sathian, 2007), which is consistent with the idea that multimodal object representations are formed through cognitive processes beyond perception.

2 Study object

2.1 Tensegrity structures

The origins of tensegrity structures are found in the art domain, mainly by the works of the sculptor Snelson (1965) who developed this structural principle to create numerous original objects (Figure 2). They have interested architects and engineers for their spatial and mechanical characteristics, such as R.B. Fuller who imagined architectural applications. More recent applications have been realised, as presented in Figure 2 right (K. Kawaguchi, in association with a textile membrane). Tensegrity structures are complex spatial self-stressed structures composed by compressed bars and tensile cables, where the bars are never touching each other and are only connected to cables (Motro, 2003). R.B. Fuller proposed the definition of 'tensegrity' (contraction of the words 'tensile' and 'integrity'): tensegrity structures are defined as islands of compression in an integral sea of tension. This characteristic results in specific effects, giving the appearance of bars freely flowing in the air, almost chaotically organised. This generally leads to very troubling perception, contrasting with the spontaneous 'gravitational' image that we have of structures based on a continuous set of compressed elements from the top to the ground. This effect, together with the geometrical and mechanical complexity, requests necessarily a specific and rich perceptual processing to elaborate and

Figure 2 Tensegrity structures: three K. Snelson's sculptures, mast with tensile membrane

understand such structures, as next presented for one simple case that will be used to investigate these mental processes.

2.2 The 'simplex': a complex spatial structure

The simplex is the most elementary tensegrity structure, composed of three rigid bars assembled by nine tensioned cables and characterised by a three-fold symmetrical structure (see Figure 3). Even if it is considered as the simplest possible system, its elementary but complex spatial characteristics provide a pertinent object of exploration. Tensegrity structures, because of their specific principle, have complex geometrical configurations; we define as complex a spatial organisation where the links between the elements are multiples and impossible to reduce to a single elementary known configuration, like a prism or a cube.

2.3 Eligibility traces in the simplex structure

The well-defined capacity limits of attention and visual spatial working memory have been discovered on the basis of visual memory studies (Miller, 1956; Oberly, 1928; Parkin, 1999; Vogel, Woodman, & Luck, 2001), which have shown that any normally developed human adult is capable of attending to an average maximum of seven (±2) representations, and of retaining them in working memory for several minutes (Potter, 1993). Representations stored in working memory may correspond to visual fragments or whole visual scenes, single numbers or groups (often called 'chunks' or 'clusters') of numbers, as in a code, single words or whole sentences. The structure of these representations depends on the storage system that is solicited by a given task, and

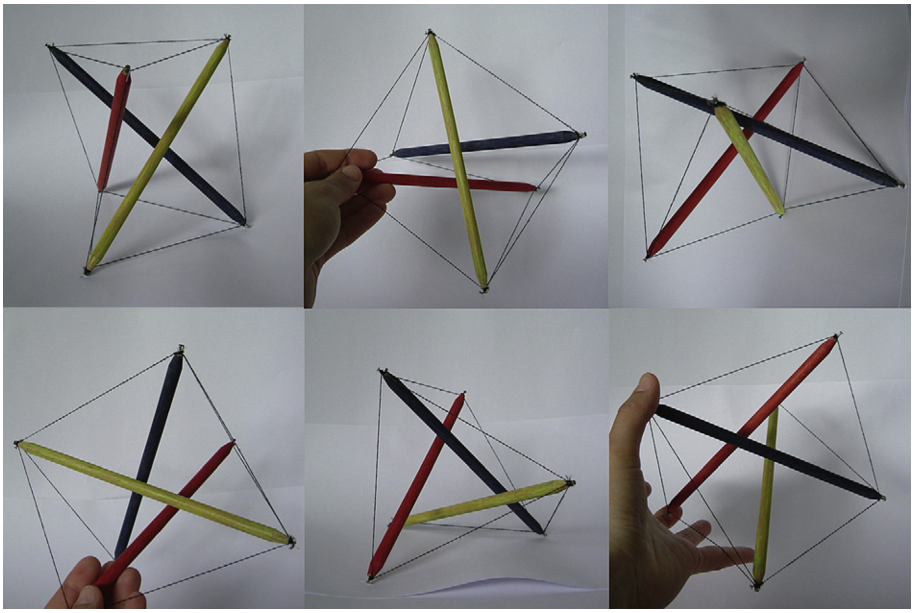

Figure 3 'Simplex' tensegrity structure composed of three bars and nine cables on the information processing strategies that may be activated. Regarding the functional architecture of visual spatial working memory, it is widely assumed that this particular form of memory involves a central executive and a so-called 'visual spatial sketchpad' (Baddeley, 2003) as the main storage system.

A visual spatial memory sketchpad of the simplex may be formed on the basis of either some elementary perceptual parts (two triangles, three polygons, and three central bars), or on the basis of some geometrical descriptors which define its structure in terms of perceptual relations between, or perceptual operations on, structural components. To be managed by the working memory's limited capacity, the spatial characteristics of this object must be represented with less than seven perceptual chunks of information.

Experimental studies already demonstrated the presence and importance of elementary geometrical concepts in perceptual and cognitive processing (Kanisza, 1980). Visual spatial problem solving data collected from an isolated Amazonian indigene group (Dehaene, Izard, Pica, & Spelke, 2006) provide compelling evidence that elementary geometry is part of universal core knowledge present in all normally developed human adults, regardless of gender, upbringing, or schooling.

Relying on elementary geometrical concepts, we made some hypothesis on the ways in which the elements of the simplex (three bars and nine cables, arranged in a complex way) can be perceptually related to each other, with the lightest cognitive weight:

- Two sets of three lines connect three oblique bars at their ends and they form two triangles. They seem to be identical (but differently orientated in space) and equilateral (all the lines with the same length) (Figure 4i).
- The three oblique bars seem identical and symmetrically arranged around a virtual axe (three-fold symmetry) (Figure 4ii).

Elementary forms, repetition and identity are fundamental factors in visual (virtual or real) processing of shapes and objects (Kanisza, 1980). Structural organisation, like symmetry, is also a fundamental factor; experimental studies already demonstrated that human observers may be able to derive representations of additional views from the single virtual sample view on the basis of symmetry transformations (Vetter, Poggio, & Bülthoff, 1994). We suppose therefore that those two geometrical chunks of information (the triangles and the symmetry of the bars) might be the perceptual eligibility traces of the simplex, which means the critical perceptual data (spatial relationships) necessary to build a mental representation of its spatial configuration. Eligibility traces are the fundamental tools used to perform the matching operations necessary to accomplish the task of our study. The conditions under which the simplex is viewed and/or explored by observers, who have never seen it before,

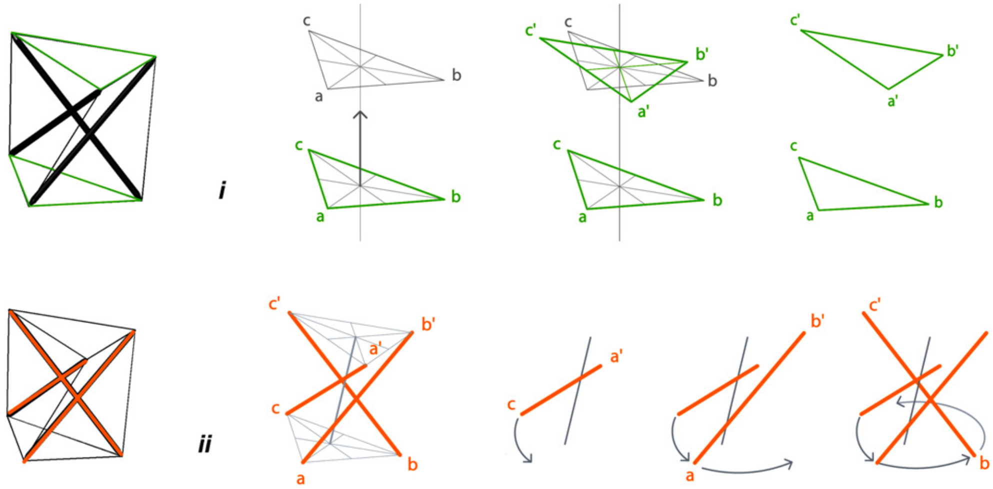

Figure 4 The two most probable eligibility traces in the simplex structure: two triangles on the top and bottom that seem equilateral and identical but differently oriented (i) and three bars, connecting the corners of the triangles, which appear to be identical and symmetrically organised around an axe (ii)

could determine whether perceptual eligibility traces are effectively made accessible to the process underlying the formation of a mental representation of the object.

3 The experiment

3.1 Objectives

Using concepts and arguments introduced above, we performed an experiment to find out what helps individuals with different specialised skills to process a visual spatial structure they have never seen before. The objective is to obtain data to support the hypothesis that representation modality has an influence on the nature of the spatial information that is accessible for perceptual and cognitive elaboration and, by consequence, on the learning process of the spatial configuration of objects (as previous studies suggest, Ho & Eastman, 2006). In particular, we are interested in the representation modalities currently used in the design field (physical model, 2D projection, numerical model). Moreover, we wanted to see if the expertise level (familiarity with geometrical concepts and numerical modelling tools, expertise specific to designers) has any influence on this process.

The simplex structure was therefore used as study object in this visual spatial working memory experiment where observers had to reproduce parts of on 2D sheets of paper, by drawing them from memory immediately after viewing/exploring the object in one of different representation modalities. Recent works have shown that the act of drawing is a powerful means of accessing, activating and consolidating knowledge representations of object properties in the memory structures of the brain, involving the most important functional regions for learning and communication, such as the Brodman area (region of the cortex, Harrington, Farias, Davis, & Buonocore, 2006). To successfully draw a novel object (or parts of it) from memory involves cognitive processes of attention and capacity-limited working memory.

3.2 Variables

The study is based on two main variables, which we wanted to analyse in their reciprocal interaction: the study population (experts/non-experts) and the exploration modality (2D/real 3D/virtual 3D).

Three different *exploration modalities* are therefore proposed to subjects (Figure 5):

- *Exploration modality 1: single 2D view*
  Axonometric projection on a sheet of paper.
- *Exploration modality 2*: *real 3D object* (visual-plus-tactile exploration)
  Physical model of the simplex made of three identical wooden bars (length 17 cm; diameter 0.5 cm), and tensioned textile strings (length 12 cm).
- *Exploration modality 3: digital 3D model* (multiple 2D viewing on the screen)
  Digital model generated with AutoCAD software, enabling multiples views on the screen through the '3D orbit' option (press and hold mouse button to move structure freely, release button to maintain a selected view).

Two secondary variables were also introduced to test the extent to which processes of visual attention could influence performance in the spatial memory task given. One variable was the colour of the bars of the structure. Colour is known to act as a powerful attracting or distracting force in processes of selective attention (Yantis & Jones, 1991). To test whether variations in selective attention could influence our study task, we presented to all observers two versions of the structure (for each exploration modality), one with three plain black bars, one with bars of three different colours.

The other secondary variable was the spatial reference frame given to the observers for drawing their solution. Two different reference frames were given to each observer, who therefore drew the bars from memory twice after each of the two views (Figure 5):

- *Positional reference frame*, containing only topological information (position of points in space: end points of the three bars).
- *Relational reference frame*, containing positional information (end points of the bars) but also some information on the relation between the points (all the cables are drawn).

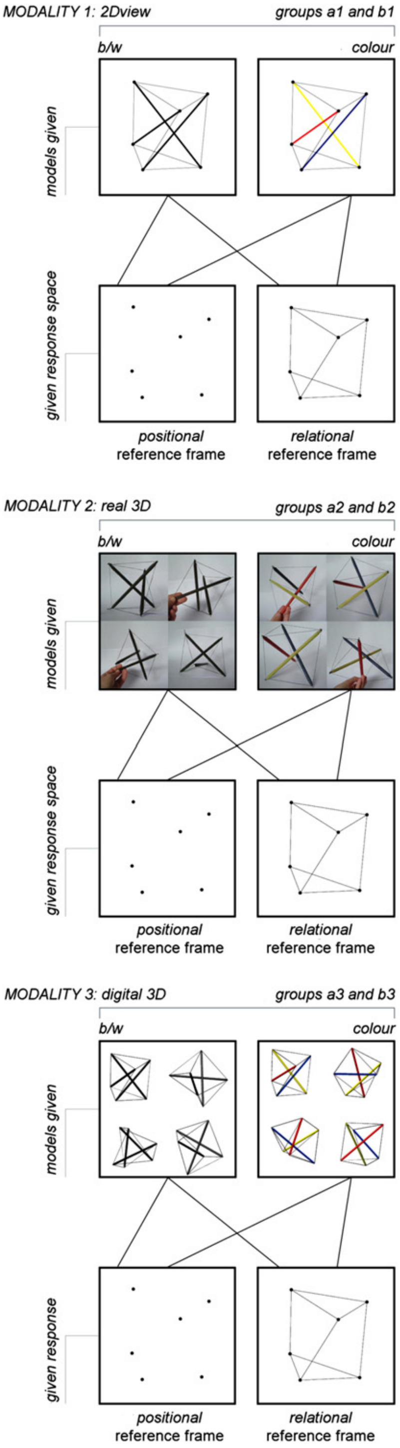

Figure 5 The models, viewing conditions and reference frames given for drawing from memory. Each of the two study populations (a and b) was divided into three independent groups of eight observers, one group for each representation modality (a1–a3, b1–b3). Each observer saw the structure twice (once in colour and once in b/w). Each of the two viewings was followed by two trials for drawing the three central bars of the structure from memory into the 2D reference frames. Click on pictures to see videos of exploration modalities 2 and 3

If selective attention strongly influences the study task proposed here, the response frame with 2D cues to structure (relational frame) is expected to enable better performances, because it makes structural cues directly available to visual attention. The response frame with positional information only (positional frame) does not make structural cues available to visual attention; they have to be retrieved from visual spatial working memory instead.

3.3 Observers

The observers who participated were 48 volunteers, divided in two study populations: '*population a*' is '*experts*' (high geometrical skills and familiarity whit numerical tools) and '*population b*' is '*non-experts*'. The 24 non-experts were post-graduate adults of homogenous cultural background, who all had successfully accomplished at least a Master Degree. We ensured that none of the non-experts was an experienced user of 3D image processing software tools, or regularly played computer or video games. The 24 expert subjects were high achiever post-graduates in architecture or engineering, of homogenous cultural background. All of them are regular users of computer tools for 3D shape representation, with a level of expertise corresponding to at least 5 years practice in modelling software (AutoCAD, 3D Studio, Catia, etc.). None of the subjects of either of the two study populations had ever seen a simplex before.

3.4 Procedure

Each population of 24 subjects (*a* and *b*) is divided into three groups of eight observers (*a1—a3* and *b1—b3*); the group number corresponds to the exploration modality (e.g. real object for *b2*). Subjects explored a monochrome (black/white) and a coloured version of the simplex, in counterbalanced order. They had the opportunity to inspect a given model for as long as they wanted. They were informed that they would have to draw certain parts of the objects from memory, but not told which parts. After each observation, when they felt ready, the model was taken away and they were asked to draw the three bars on two separate sheets featuring the two different reference frames (in counterbalanced order). In each trial, the time taken to draw the bars from memory was measured and the number of positional errors made was recorded.

4 Results

We computed average times (*T*), expressed in seconds, for drawing the three bars from memory, and their standard deviations. Errors (*E*) were counted for bars drawn in the wrong position: *E* varied between 0 (all three bars in the correct position) and 3 (all three bars in the wrong position) for every trial.

The effects of repeated measures within each study population (*a* and *b*) were analysed first. Means in *T* and *E* were compared as a function of the colour of the bars of the structure, and of the information contained in the two spatial reference frames given for drawing. The effects of the different exploration modality within and between the two study populations were analysed next.

Even if precedent studies seem to demonstrate the importance of gender differences in visual spatial processing (Crucian & Berenbaum, 1998; Voyer, Voyer, & Bryden, 1995), we found no significant differences in performance between female and male subjects. This finding is coherent with other studies that show that differences between genders in spatial tasks can depend on the specific conditions of a given test (Seurinck, Vingerhoets, de Lange, & Achten, 2004) or are absent in highly experienced subjects (Unterrainer, Wranek, Staffen, Gruber, & Ladurner, 2000). Therefore, gender differences will not be discussed further.

4.1 Colour and spatial reference frame

Whether observers saw the coloured structure first followed by the b/w one, or the other way round, produced no significant effect on T or E within each of the two study populations. The information provided in the two drawing spatial reference frames had no significant effect on T or E either, whether observers drew into the positional frame first and then into the relational frame, or the other way round. These results are presented in Tables 1 and 2.

The T values and their standard deviations reveal small, non-significant differences between trial conditions relative to bar colour and/or reference frame within a given population and small, non-significant variations in E.

Figure 6 shows the distributions of the individual data for T, in ascending order, in the two study populations. Non-parametric normality test – Kolmogoroff–Smirnoff Test (Kendall & Stuart, 1979) – was applied to calculate probabilities for the data in Figure 5 to obey the law of normal distributions. This test gave KS-statistics with $\alpha = 0.13$ (goodness of fit 87%) for the

Table 1. Average times T (medians of the distributions, standard deviations σ), on top and number of errors E, on bottom, are shown for the two study populations, irrespective of the viewing condition, as a function of the b/w–colour of the object

	Non-experts		*Experts*	
	B/w first	*Colour first*	*B/w first*	*Colour first*
Black/white (N = 48)				
$T(s)$	43.5	41.6	27.4	26.8
Median; σ	30; 30.6	20; 31.2	21; 26	20; 24.3
E	15	22	7	2
Colour (N = 48)				
$T(s)$	51.4	44.6	27.8	30.5
Median; σ	35; 40.8	31; 36.9	18; 26.7	24; 24.9
E	15	29	3	5

Table 2. Average times T (medians of the distributions, standard deviations σ), on top and number of errors E, on bottom, are shown for the two study populations, irrespective of the viewing condition and as a function of the reference frames given for drawing the solution

	Non-experts		*Experts*	
	Relat. first	*Posit. first*	*Relat. first*	*Posit. first*
Relational ($N = 48$)				
$T(s)$	44.2	46.6	31.3	24.5
Median; σ	26; 34.2	29; 39.1	24; 26.8	17; 21.9
E	18	19	2	6
Positional ($N = 48$)				
$T(s)$	47.2	49.0	29.0	30.1
Median; σ	31; 30.4	25; 37.0	20; 25.9	23; 28.2
E	21	23	4	5

T distribution of the non-experts, and $\alpha = 0.11$ (goodness of fit 89%) for the experts, signalling significant probabilities that these distributions are not different from normal distributions (probability > 0.87 for both distributions). As a consequence, parametric testing could be confidently applied to the T data.

Analysis of variance (ANOVA) for repeated measures of T in a $24 \times 2 \times 2 \times 2$ design (24 observers × two study populations × two colour conditions × two repeated measures with different spatial reference frames) revealed, as expected, a statistically significant effect of study population ($F\,(1, 16) = 6.03$; $p < 0.01$; i.e. a statistical significant difference between global variance and variance of the 'study population' factor), and statistically non-significant effects of colour ($F\,(1, 16) = 0.43$; no statistically significant difference between

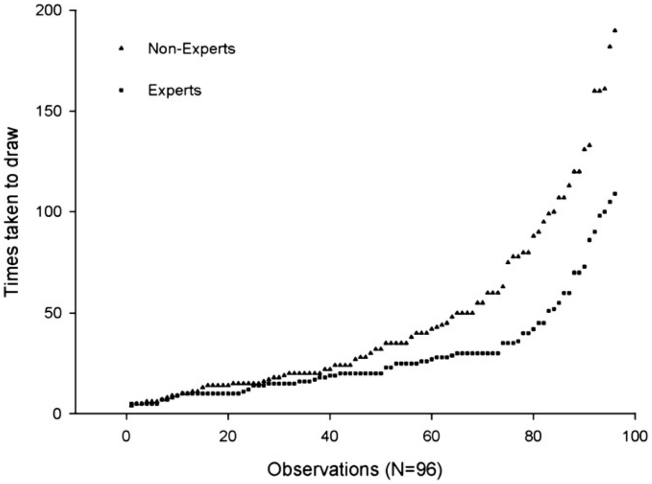

Figure 6 Distributions of times taken to draw from memory for the two study populations

global variance and variance of the 'colour' factor) and spatial reference frame ($F(1, 16) = 0.328$; no statistically significant difference between global variance and variance of the 'reference frame' factor).

Parametric testing for differences in E is not appropriate, because the number of possible errors was limited to three at any given trial. From the data in Tables 1 and 2 we may conclude that the experts made, as expected, fewer positioning errors when drawing the bars than the non-experts. However, no systematic differences in E as a function of either the colour or the reference frame are observed.

Given the non-significant variations in T and E as a function of the colour of the object and the type of reference frame (Tables 1 and 2), we calculated an average performance measure based on the four repeated measures/trials of a given observer ($T/4$ and $E/4$). All further analyses were performed using these average measures.

Noticeable differences in both T and E are found between the two study populations, with consistently shorter T (29 s *versus* 47 s) and fewer E (17 *versus* 81) in the expert population, as would be expected.

4.2 Exploration modality

Both experts and non-experts performed equally well in the 2D single view modality, which could be expected. The two groups which revealed the greatest differences in performance between the two study populations are those who explored the structure through virtual 3D viewing. The best performances, in terms of the shortest T values and the fewest E, were produced in the exploration modality of 3D digital viewing by the corresponding group in the expert population. The poorest performances were delivered in the real 3D object exploration by the corresponding group in the non-expert population, as summarised in Table 3.

The data show that the benefit of digital 3D viewing on the times taken to draw by the experts, compared with visual and tactile exploration of the simplex, is approximately 50%, which is considerable. The average number of errors reveals a noticeable benefit on accuracy that goes in the same direction. The performances of the non-experts with the real 3D object and the digital 3D quite clearly show that they do not, as might have been expected, draw any advantage from the multi-sensory exploration of the real object, compared with multiple 3D viewing on a screen. In fact, performances tend to be slightly slower and less accurate with the real 3D for the non-experts. The effect goes in the same direction as the one observed for the experts, but without the same noticeable advantage of digital viewing.

Table 3. Average values for *T* (with minima, maxima, medians of the distributions and standard deviations) and *E*, as a function of study population and exploration modality

	Non-experts	*Experts*
2D single view		
T(*s*)	19.0	22.7
Median; σ	5; 6.9	15; 18.6
Minima; maxima	4; 13	5; 35
E	2.25	0
Real 3D (vision and touch)		
T(*s*)	59.1	43.3
Median; σ	43; 51.4	20; 27.3
Minima; maxima	15; 160	7; 109
E	11.25	3.5
Digital 3D		
T(*s*)	57.0	18.6
Median; σ	35; 43.2	14; 11.5
Minima; maxima	14; 190	9; 50
E	7.75	1

ANOVA for a 16 × 3 design (16 observers × three groups/exploration modality) revealed a statistically significant effect on T of the exploration modality ($F(2, 13) = 4.53$, $p < 0.01$; i.e. a statistical significant difference between global variance and variance of the ‘exploration modality’ factor).

5 Discussion

This experiment with the simplex structure shows that observers who do not have any particularly trained visual spatial skills are perfectly able to learn the spatial configuration of a novel, complex and abstract object in a single trial by exploring it on the screen of a computer or through the manipulation of a scale model. Visual spatial working memory was not solicited at its capacity limits, given that the number of errors made by the non-experts in the different exploration modalities was far from attaining the maximum possible. Both times taken for drawing from memory and the number of errors made were consistent within a given study population. Whether the object had colour or not, or whether the reference frame given for drawing contained more or less 2D structural information had no significant influence on the performances of either population. Reference frame effects were found previously in the recognition of unusual novel shapes, explored previously under different conditions of aperture viewing (Krüliczak, Goodale, & Humphrey, 2003). The absence of any such effect or any effect of colour in our results suggests the hypothesis of a central processing of complex spatial information: in this case, the importance of strictly perceptual cues, like colour or frame of reference, is not relevant.

The absence of any noticeable effect of repeated measures on performances in either study population further sustains the idea that the proposed study task involved complex memory matching operations well-beyond perception. This latter together with the fact that there were no measurable effects of visual attention confirm earlier intuitions that the perceptual processes necessary for learning abstract visual and spatial representations are central and involve cognitive operations beyond the sensory level (Gibson, 1963). Our data do not reproduce any haptic advantage effect, found previously by others in the visual recognition of familiar objects (Norman et al., 2004). This is readily explained by the fact that the perceptual learning process underlying the formation of representations of novel and abstract objects is obviously controlled by a process that disregards whether structure is explored, and thereby learnt, visually, manually, or both.

The most striking finding in the results is the considerable comparative advantage of digital 3D viewing on times taken to draw from memory and errors made by the experts, compared with exploration of the real 3D object. The transformation of real 3D samples into 2D sketches is a mental operation that design experts are used to perform; it constitutes one of the most specifics skills of designer's education end expertise. It should, therefore, not incur any cost in performance compared with digital viewing, where such a transformation is not necessary. The fact that this additional transformation incurs a measurable and quite considerable cost on the performances of highly experienced specialists can only be explained by the difference in nature of the eligibility traces made available in the two different exploration modalities. With the fact that selective and more specifically visual effects are absent from data, this confirms the hypothesis that the mental transformations solicited by the experimental task of this study entirely rely on high-level cognitive operations, like the matching operations. In the results, digital 3D exploration appears to be the most 'efficient' representation modality to perform the task.

We can explain this result by the idea of virtual reality as an augmented reality: for any given, limited amount of time, a larger and potentially richer sample of different views can be generated virtually compared with active exploration of the real object. We could suppose that through a sequence of 2D views, like on the computer screen, it might be possible to transmit more efficiently than with a real model the eligibility traces necessary to process spatial configurations. This finding raises deep questions about the nature of 3D object representation in the human brain. Our observations could be consistent with certain theories (Sinha & Poggio, 1996) which claim that representations of 3D visual perceptual structure, especially when the latter is abstract and not supported by familiarity cues, are learnt and stored in the brain through cognitive processes that rely on 2D spatial information. More realistically, we must look deeper to the nature of the task proposed to subjects (drawing on a 2D sheet of paper): the eligibility traces that are needed to perform this operation are necessarily 2D. In this case, the intrinsically 2D nature of the eligibility traces that are being sampled while exploring the virtual object through multiple 2D

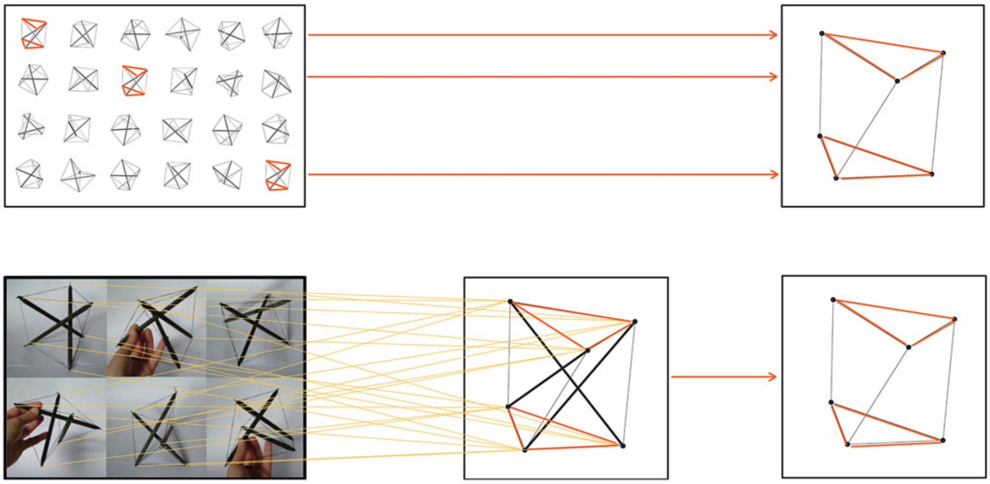

Figure 7 Multiple digital viewing on the computer 2D screen allows a direct one-to-one matching operation between some representations seen on the computer screen and the 2D reference frame. Multimodal (visual and tactile) exploration of the real 3D object implies a many-to-one matching, which requires the elaboration, in the short-term memory visual sketchpad, of a temporary 2D representation to enable the matching with the 2D reference frame

viewing on the screen enables many-to-one matching with the 2D reference frame directly, without any further transformation. Instead, the working memory samples from the real 3D object necessarily need to undergo an additional transformation (Figure 7), consisting in the elaboration of a temporary 2D sketch in the visual spatial working memory sketchpad; this elaboration requires longer times of processing, as clearly appears in our results. Therefore, we can suggest that numerical models constitute an augmented reality specifically in the transmission of 2D data (figurative and visual qualities of objects). Therefore, the efficiency in the transmission of real 3D spatial qualities of numerical models cannot be directly concerned by the results of this study; it constitutes instead one of the possible fields of development of the methods and results of this research.

6 Conclusion

The principal objective of this study is to propose a framework to better understand the central processing of perceptual and cognitive elaboration of complex forms in relation to expertise and to representation modality, with a particular attention to expertise and representation tools referring to the architectural design domain. Our results seem support the main hypothesis at the base of this study. Representation tools do have an incidence on perceptual and cognitive processing and by consequence tools that are used during architectural design do have an influence on both process and results. Research in representation and design process, especially on the subject of the mechanisms of perceptual and cognitive elaboration of complex spatial information, appears necessary in the context of contemporary architectural design, which relies extensively (or even exclusively, as in the case of 'Digital' or 'Non-Standard' architecture) on numerical modelling and representation. The experiment performed with the simplex aims to be the first step on the way of an operational research on cognitive elaboration of complex spatial configurations, with the objective to exploit the results to directly improve design process and methods.

Acknowledgements

Financial support from the I.U.A.V. in Venice to C. Silvestri is gratefully acknowledged.

Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.destud.2010.03.001.

References

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.

Borgart, A., & Kocaturk, T. (2007). Free-form design as the digital "Zeitgeist". *Journal of the International Association for Shell and Spatial Structures, 48*(4), 3–10.

Carpenter, G. A., & Grossberg, S. (1991). *Pattern recognition by self-organizing neural networks*. Cambridge, MA: MIT Press.

Christou, C. G., & Bulthoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory and Cognition, 27*, 996–1007.

Crucian, G. P., & Berenbaum, S. A. (1998). Sex differences in right hemisphere tasks. *Brain & Cognition, 36*, 377–389.

Darken, R. P., & Sibert, J. L. (1996). Wayfinding strategies and behaviours in large virtual worlds. In *Symposium Proceedings 'Human Factors in Computing Systems'* (pp. 142–149). London: ACM Press.

Dehaene, S., Izard, V., Pica, P., & Spelke, E. (2006). Core knowledge of geometry in an Amazonian indigene group. *Science, 311*, 381–384.

Dresp, B., Silvestri, C., & Motro, R. (2007). Which geometric model for the curvature of 2-D shape contours? *Spatial Vision, 20*, 219–264.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review, 69*, 477–491.

Gibson, J. J. (1963). The useful dimensions of sensitivity. *American Psychologist, 18*, 1–5.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Harrington, G. S., Farias, D., Davis, C. H., & Buonocore, M. H. (2006). Comparison of the neural basis for imagined writing and drawing. *Human Brain Mapping, 28*, 450–459.

Ho, C.-H., & Eastman, C. (2006). An investigation of 2D and 3D spatial and mathematical abilities. *Design Studies, 27*, 505–524.

Kanizsa, G. (1980). *Grammatica del vedere*. Bologna: Il Mulino.

Katz, M. G., Kripalani, S., & Weiss, B. D. (2006). Use of pictorial aids in medication instructions: a review of the literature. *American Journal of Health System Pharmacy, 63*, 2391–2397.

Kendall, M., & Stuart, A. (1979). *The advanced theory of statistics*. New York: Oxford University Press.

Kolarevic, B. (2003). *Architecture in the digital age: Design and manufacturing*. NY/London: Taylor and Francis.

Krüliczak, G., Goodale, M. A., & Humphrey, G. K. (2003). The effects of different aperture-viewing conditions on the recognition of novel objects. *Perception, 32*, 1169–1179.

Lacey, S., Peters, A., & Sathian, K. (2007). Cross-modal object recognition is viewpoint independent. *PLoS One, 12*, e890.

Matheis, R. J., Schultheis, M. T., Tiersky, L. A., DeLuca, J., Mills, S. R., & Rizzo, A. (2007). Is learning and memory different in a virtual environment? *Clinical Neuropsychology, 21*, 146–161.

Miller, G. A. (1956). The magic number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Motro, R. (2003). *Tensegrity: Structural systems for the future*. London: Kogan Page Science.

Newell, F. N., Ernst, M. O., Tjan, B. S., & Bülthoff, H. H. (2001). Viewpoint dependence in visual and haptic object recognition. *Psychological Science, 12*, 37–42.

Norman, J. F., Norman, H. F., & Clayton, A. M. (2004). The visual and haptic perception of natural object shape. *Perception & Psychophysics, 66*, 342–351.

Oberly, H. S. (1928). A comparison of the spans of attention and memory. *American Journal of Psychology, 40*, 295–302.

Oxman, R. (2006). Theory and design in the first digital age. *Design Studies, 27*, 229–265.

Parkin, A. J. (1999). Human memory. *Current Biology, 9*, 582–585.

Potter, M. C. (1993). Short-term conceptual memory. *Memory & Cognition, 21*, 156–161.

Seurinck, R., Vingerhoets, G., de Lange, F. P., & Achten, E. (2004). Does egocentric mental rotation elicit sex differences? *Neuroimage, 23*, 1440–1449.

Singh, S. P., & Sutton, R. S. (1996). Reinforcement learning with replacing eligibility traces. *Machine Learning, 22*, 123–158.

Sinha, P., & Poggio, T. (1996). Role of learning in three-dimensional form perception. *Nature, 384*, 460–463.

Snelson, K. D. (1965) Continuous tension, discontinuous compression structures. *US Patent Number 3, 169*, 611.

Stelzer, E. M., & Wickens, C. D. (2006). Pilots strategically compensate for display enlargements in surveillance and flight control tasks. *Human Factors, 48*, 166–181.

Unterrainer, J., Wranek, U., Staffen, W., Gruber, T., & Ladurner, G. (2000). Lateralized cognitive visuospatial processing: is it primarily gender-related or due to quality of performance? *Neuropsychobiology, 41*, 95–101.

Vetter, T., Poggio, T., & Bülthoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. *Current Biology, 4*, 18–23.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 92–114.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex-differences in spatial abilities – a meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.

Whitehead, S. D., & Ballard, D. H. (1991). Learning to perceive and act by trial and error. *Machine Learning, 7*, 45–83.

Yantis, S., & Jones, E. (1991). Mechanisms of attentional selection: temporally modulated priority tags. *Perception & Psychophysics, 50*, 166–178.