Journal id: CPHP_A_284219

Corresponding author: Nicholas Bullot

Title: Keeping track of invisible individuals while exploring a spatial layout with partial cues: Location-based and deictic direction-based strategies

NO QUERIES

Philosophical Psychology
Vol. 21, No. 1, February 2007, pp. 1–31

Keeping track of invisible individuals while exploring a spatial layout with partial cues: Location-based and deictic direction-based strategies

Nicholas Bullot[1] and Jacques Droulez[2]

[1]*Nicholas Bullot is a Postdoctoral Fellow in the Department of Philosophy, University of Toronto, and* [2]*Jacques Droulez is Direct of Research at the Laboratoire de Physiologie de la Perception et de l'ActionCollège de France*

In contrast to Constructivist Views, which construe perceptual cognition as an essentially reconstructive process, this article recommends the Deictic View, which grounds perception in perceptual-demonstrative reference and the use of deictic tracking strategies for acquiring and updating knowledge about individuals. The view raises the problem of how sensory-motor tracking connects to epistemic and integrated forms of tracking. To study the strategies used to solve this problem, we report a study of the ability to track distal individuals when only their directions can be perceived and not their locations. We introduce a new experimental paradigm named the 'Modified Traveling Salesman Problem' (MTSP), which requires subjects to visit n invisible targets in a 2D display once each. Surprisingly, subjects are competent at this task for up to 10 targets. We consider two types of tracking strategies that subjects might use: 'location-based' strategies and 'deictic direction-based' strategies. A number of observations suggest that subjects used the latter, at least for larger numbers of targets. We hypothesize that subjects used perceptual-demonstrative reference and deictic strategies (i) to perform the sensory-motor tracking of directional segments, (ii) to bind the segments with their updated status in the task, and (iii) to perform the epistemic tracking of invisible targets by means of perception-based inferences.

Keywords: Demonstrative Reference; Deictic Strategies; Tracking; Location; Direction; Identification; Predicate; Perceptual Inference; Spatial Memory

Correspondence to: Nicholas Bullot, Bullot University of Toronto, Department of Philosophy, 3359 Mississauga Rd. N.N. Bldg., Room 227, Mississauga, ON L5L 1C6, Canada. Email: nicolas@objectcognition.net

ISSN 0951-5089 (print)/ISSN 1465-394X (online)/08/010001-31
DOI: 10.1080/09515080701840861

1. Outline

We begin in §2 by contrasting two views on the nature of perceptual cognition. The first—the *Constructivist View*—apprehends perceptual cognition as a process of 'reconstruction' in which an internal model of distal objects in the world is constructed. The second—the *Deictic View*—grounds perception in perceptual-demonstrative reference and deictic tracking strategies (which use sensory-motor and attentional systems as 'pointers' to available information), and leaves much of the world as a source of information that can be consulted as needed. In §3, to study the deictic strategies used for tracking individuals that cannot be *directly* identified and located, we report a study of human subjects' ability to track distal individuals using only directional, rather than locational, cues. This study introduces a new experimental paradigm named the 'Modified Traveling Salesman Problem' (MTSP), which requires subjects to visit *n* invisible targets in a 2D display once and only once. MTSP studies integrated and flexible forms of tracking that bind sensory-motor tracking with the epistemic tracking of target individuals. Alongside its relevance for the theory of tracking and reference, MTSP relates to recent research on memory[1] for object locations and perception-based inferences[2]. Finally, in §4, we present the findings and propose an analysis based on the distinction between two types of tracking strategies, location-based strategies and direction-based deictic strategies.

2. The Deictic View of Perceptual Cognition

2.1. The Constructivist View of Perceptual Cognition

Classical experiments in psychophysics investigate the abilities of perceptual systems to recover information from sensory signals according to a Constructivist View (CV).[3] This approach conceives the perception of a scene as the reconstruction of an internal representation, or model, of properties of the distal objects. According to this view, the human capacity to perceive objects is defined entirely in terms of an internal model or representation of the physical (i.e. intrinsic, spatial) properties of the distal object—for instance, its 3D details and parts. The approach comes in a variety of subtypes, which differ in the properties or relations thought of as being 'modeled' in perceptual representations. Canonical examples have been analyzed in research on vision (Ballard, Hayhoe, Pook, & Rao, 1997; Churchland, Ramachandran, & Sejnowski, 1994; Pylyshyn, 2000).

A basic constructivist problem in vision science has been the so-called 'inverse problem'—or the problem of reconstructing a 3D layout from the 2D retinal stimulus; see for example Palmer (1999, pp. 23–24). Marr's proposal on vision and recognition (Marr, 1982; Marr & Nishihara, 1978) belongs to this approach. As a consequence of the CV, many psychophysical paradigms focus on the relation between stimulus—or object-intrinsic—characteristics, and their internal representation. More generally, works in the constructivist tradition have not been primarily concerned with the relation between the perceiver's actions, strategies and inferences (e.g. goals, sub-goals, plans, and vision- or perception-based inferences) and perceptual processing. In a related example, the pictorialist view of mental imagery is a constructivist approach, in the sense that it postulates internal mental displays that preserve fundamental spatial or topological relationships between represented elements (Kosslyn, 1994; Pylyshyn, 2002).

The application of the CV to vision runs into several objections. Numerous authors have outlined empirical evidence against the view that the visual system constructs a persistent internal model of the environment (Hayhoe, Bensinger, & Ballard, 1998; Hayhoe, Shrivastava, Mruczek, & Pelz, 2003; Horowitz & Wolfe, 1998; Karn & Hayhoe, 2000; Pylyshyn, 2003; Rensink, O'Regan, & Clark, 1997; Simons & Rensink, 2005). Moreover, the objections directed against the analyses of 3D internal reconstructions apply (accepting for some modifications) to other versions of the constructivist approach that are based on 2D internal reconstructions (Clark, 2000, p. 100; Humphreys, 1999, pp. 175–176; O'Regan, 1992, pp. 464–471; Pylyshyn, 2002). On the theoretical standpoint, the idea of an elaborate scene model, has been challenged by the emergence of alternative approaches, called active (Findlay & Gilchrist, 2001, 2003), animate (Ballard, 1991), bounded (Gigerenzer, Todd, & ABCgroup, 1999; Todd & Gigerenzer, 2000), deictic (Agre, 1997; Ballard et al., 1997), dynamical (van Gelder, 1998), embodied (Clark, 1999), interactive (Churchland et al., 1994) or situated (Brooks, 1999; Pylyshyn, 2000) vision or cognition, that take advantage of the observer's actions and tracking abilities in order to minimize the number of object-intrinsic properties that must be encoded in the perceptual representation. In these works, there has been a lasting discussion of various shortcomings of constructivist views, in particular the need for a homunculus to monitor the 'internal screen' (O'Regan, 1992), 'Cartesian Theater' (Dennett, 1991, 2001), 'internal image' (Pylyshyn, 2003), or 'internal 3D model' (Ballard, 1996; Churchland et al., 1994).

2.2. The Deictic View of Perceptual Cognition: Grounding Perceptual Cognition in Perceptual-Demonstrative Reference and Deictic Tracking Strategies

The Deictic View (DV) of perceptual cognition can be thought of as an alternative approach to that championed by the constructivists.[4] It appeals to perceptual demonstratives and deictic strategies that allow a subject who is acting as a tracker to track, reason about, and perform actions on a limited number of distal individuals without constructing a detailed internal model of each of them. This view is *externalist* in the sense that the tracker's perceptual contents are essentially dependent on interactions with individuals external to the tracker's organism. Thus, it focuses on the active strategies that build the tracker's *cognitive access* to individuals instead of the internal duplication of the targets' properties. A theoretical account of perceptual processes belongs to the DV if it:

(i) includes a denial of the (aforementioned) principles of the CV; and
(ii) explains aspects of perceptual cognition by means of a theory of perceptual-demonstrative reference or/and deictic tracking strategies.

There are two central concepts in the second part of the definition: perceptual-demonstrative reference and deictic strategies. *Perceptual-demonstrative reference* is a direct way to refer (in experience, thought or linguistic communication) to an individual ***i***[5] (e.g., a particular cup, cat or person). This kind of reference occurs when the tracker perceptually tracks ***i*** and refers to it on the basis of the acquaintance provided by its current perception. Since early analyses (Kaplan, 1989; Russell, 1910–11, 1956b), demonstrative reference is usually contrasted with more *indirect* ways of referring to individuals (Bach, 1987; Clark, 2000, pp. 131–136; Evans, 1982, pp. 143–45; Ganea, Shutts, Spelke, & DeLoache, 2007; Peacocke, 2001; Pylyshyn, 2000, p. 199; 2001, pp. 129–130; Recanati, 1993, pp. 97–118). Such indirect ways include the description[6] of an individual ***i*** when the latter cannot be perceived or its description by a notational system (Goodman, 1968) such as a diagram or a map. For a linguistic example, consider these two sentences: "This [*pointing in the direction of perceived* **i**] is *P*."; and "The *D* is *P*.", wherein *D* is a description satisfied by ***i***. Only the former sentence is traditionally conceived of as a perceptual-demonstrative reference to ***i***.

Perceptual-demonstrative reference allows human trackers to perform acts of perceptual-demonstrative *identification*, which have been studied in terms of the following distinct aspects: (1) the maintenance of an 'informational link' with an individual, and (2) the binding of a predicate, or 'tag', to an individual which allows the tracker to categorize the target and draw perceptual inferences about it. The first aspect—the maintenance of an *informational link*—refers to the capacity to maintain cognitive *access* to an individual ***i*** by the means of the tracker's sensory-motor and attentional skills and in spite of ***i***'s property changes over time.[7] Recent psychological theories have investigated the condition of this *link* between cognitive states (e.g., content of working memory, object files, intention detection, and communication) and various sensory-motor primitives whereby perceptual-demonstrative identification can be achieved. These sensory-motor primitives include eye fixations[8], visual and attentional tracking[9], pointing gestures[10], hand grasping[11] or gestures that help the anchoring or learning of language[12] and reasoning[13].

As for the second aspect (predication), in order to contribute to the tracker's knowledge of individuals, perceptual-demonstrative identification relies on the binding of external targets to structured mental contents or epistemic predicates, which can then take part in conceptual and *inferential* processes. This condition is accepted in various forms. The study of perceptual-demonstrative identification is closely related to the study of predicative structures, particularly in the work of Strawson (1959; 1974), Miller & Johnson-Laird (1976), Evans (1982), Garnham (1989), Logan & Sadler (1996), Clark (2000; 2004a; 2004b), Campbell (2002), Pylyshyn (2003), Hurford (2003) and Matthen (2005). Perceptual-demonstrative identification relies on mechanisms that make it possible to represent facts, properties, and states of affairs about perceived individuals via property or predicate ascription. In this article, we will represent an *epistemic predicate* by the expression '$F(\boldsymbol{i}_1, \ldots \boldsymbol{i}_n)$'. In this notation, '$(\boldsymbol{i}_1, \ldots \boldsymbol{i}_n)$' is a set of n singular mental states (i.e., object files, assigned deictic variables or indexes linked to individuals) referring to the distal individuals $\boldsymbol{i}_1, \ldots \boldsymbol{i}_n$ and '*F*' is a conceptual predicate linked with each of these individuals. From these characteristics, one can conclude that a system performing perceptual-demonstrative identification has to solve computational problems related to the cognitive tracking of token individuals. Typically, the system has to keep the informational link with an individual target while dynamically and incrementally updating a *singular* representation *of* the target (Bullot & Rysiew, 2007; Dretske, 1969, 1981; Evans, 1982; Pylyshyn, 2001, pp. 129–131).

In the definition of the Deictic View, we have also used the concept of *deictic tracking strategies*, which refers to the tracker's ability to use sensory-motor, perceptual and attentional systems as *pointing apparatuses* (or *pointers*) providing *cognitive access* to task-relevant information in the tracker's environment. The notion has been developed in a deictic theory introduced by Dana Ballard and co-workers (Ballard, 1997; Ballard, Hayhoe, Li, & Whitehead, 1992; Ballard et al., 1997; Ballard, Hayhoe, Salgian, & Shinoda, 2000; Hayhoe & Ballard, 2005). A crucial thesis of this theory is that human eyes are used deictically—the concept *deictic* is used to refer to the property of certain actions to serve as a pointer to information available in the organism's environment. Ballard et al. (1997, pp. 726–730) argue that the use of pointers is essential to the performance of cognitive and epistemic procedures. Eye fixation is conceived of as eye pointing directed at a referent (a source of information in the tracker's environment). Additionally, selection by covert attention is a neural pointer that interacts with eye pointing (Ballard et al., 1997, pp. 725–726). Eye pointing is known to be particularly important when vision interfaces with action controlled by epistemic states.[14] Moreover, a visual fixation presents this advantage to allow "the brain's internal representations to be implicitly referred to an external point" (Ballard et al., 1997, p. 724), which can serve in the control of actions directed at individuals. Thus, fixations are parts of more encompassing hierarchical structures termed *deictic strategies* or, 'do-it-where-I'm-looking' strategies (Ballard et al., 1997, p. 725), which are sequential combinations of routines that use discrete deictic pointers to solve epistemic problems and activate action sequences. Ballard et al. (1997, pp. 729–737) distinguish two basic routines combined in the integrated tracking of individuals: the *identification routine* (e.g., trying to identify the target of an eye pointing) and the *location routine* (e.g., trying to locate in the environment the target of a pointer in memory).

In a related deictic account, Zenon Pylyshyn (1989, 2000, 2001, 2003) has tried to unify the theory of deictic strategies and the philosophical theories of demonstrative reference. He argues that classical theories of object representation cannot give satisfactory explanations of the way representations connect with individuals in the world, because such theories do not clearly analyze the interactions between active trackers and their environment. According to his analysis, we need to explain how each tracker can connect sensory-motor information with conceptual knowledge about his or her environment. To do that, we must study the use of perceptual demonstratives which serve to directly connect token individuals with mental representations *or* with certain actions that may be performed on them.[15] Traditionally, in philosophy and semantics, a representation—such as a thought, a belief, or an utterance—is said to be *indexical* [16] if the knowledge of its referent depends on the knowledge of the context in which the representation occurs. Pylyshyn proposes to deploy this notion of indexicality in developing a theory of situated vision and visual attention. What he calls *indexical reference*, or *demonstrative reference*, therefore applies also to the visuo-motor capacity of trackers. It allows them to individuate and act upon objects by tracking them in the course of active and epistemic perception.

2.3. The Flexible Binding of Sensory-Motor Tracking with Epistemic Tracking: Deictic Strategies, Perception-Based Inferences and Integrated Tracking

With regard to experimental research on deictic tracking strategies, a number of studies have been published on visual tracking (e.g., Blaser, Pylyshyn, & Holcombe, 2000; Cavanagh & Alvarez, 2005; Keane & Pylyshyn, 2006; Pylyshyn & Storm, 1988; Scholl & Pylyshyn, 1999). However, understanding the perceptual cognition of individuals requires more than the study of perceptual tracking *in isolation* from other forms of tracking. While continuous sensory-motor and cross-modal tracking is essential for knowledge by perceptual acquaintance[17], tracking in perception cannot be the only method for tracing the continued identity of distal individuals. The reason is that, with regard to long-term periods (i.e. days, weeks or years) and scattered distributions in space, our faculty of perceptual observation is limited and has to operate on the basis of disconnected episodes. One implication of such limitations is that it is impossible to *continuously* track in perception each of the individuals to which our perceptions and thoughts refer. However, arguably, at least human beings possess a faculty of *integrated tracking* understood as the ability (i) to *bind* perceptual and conceptual forms of tracking so as to store historical information about the location and identity of some individuals (in a unified spatio-temporal conceptual frame), and (ii) to *update* this historical information in order to trace or trail them over space, time and change.

P. F. Strawson (1959, p. 32) expressed this point in connection with the limitations of our faculty of attention. According to Strawson, there can be "no question of continuous and comprehensive attention" to the preservation or change of spatial boundaries and relations "on the part of things mostly undergoing no, or only gradual, qualitative change." Strawson (1959, pp. 32–33) considers the limitations of our faculty of perceptual tracking as a fundamental challenge for epistemology. If one wants to avoid skepticism one has to accept the ontological notion of *continued identity* for an individual and the epistemological notion of (reliable) skills/methods for *tracking* and *re-identifying* an individual over distinct and unconnected episodes of perceptual tracking. Indeed, and against skepticism, in spite of our attentional limitation, our mental states can instantiate attitudes and acts of reference directed at unperceived physical individuals with continued identity. We can think about dead relatives, distant friends, lost loves, lost or recently manipulated artifacts. Although a subset of these examples can sometimes be reinstated to the status of perceived individuals, all of these are excluded from our direct perceptual knowledge.

The *integrated* tracking of individuals is therefore dependent on flexible interplay of sensory-motor and various *epistemic* forms of tracking and reference (Bullot, 2006; Bullot & Rysiew, 2007). *Epistemic tracking* refers here to cases in which the target individual cannot be perceived (hence, cannot be perceptually tracked because it has left the sensory fields) but can be identified and/or located by the tracker on the basis of indirect information gathered by such sources as memory, reasoning or communication.

As epistemic trackers in everyday deictic and epistemic actions (Ballard et al., 1997; Kirsh & Maglio, 1995), human agents frequently have to keep cognitive contact with several individuals around them, although they may not be able to see or perceive them—for instance when they have to remember recently seen or used artifacts (e.g. keys, wallets, identification credentials) or the place of objects in the dark (Graziano, Hu, & Gross, 1997). A first type of epistemic tracking may depend on the tracker's *spatial memory* for distal target location and identity—e.g., the tracking of distal targets by recall of their partially-specified locations and identity (Attneave & Farrar, 1977; Attneave & Pierce, 1978; Milner & Goodale, 1995, pp. pp. 88–91; Posma & De Haan, 1996; Shelton & McNamara, 2001a) or the capacity of automatic spatial updating of locations and directions (Farrell & Thomson, 1998, 1999; Rieser, 1989). Another type of epistemic tracking must depend on the use of communication and linguistic abilities (Ganea et al., 2007). Moreover, deictic tracking strategies may subserve epistemic tracking, because they may be associated with the capacity to draw *perceptual inferences* about the distal target individuals, on the basis of perceptual demonstratives and singular representations such as object/singular files.[18] The use of deictic pointers (e.g., visual demonstratives or indexes) may serve to link thoughts and memories with distal token individuals (Ballard et al., 1997; Pylyshyn, 2001, 2003).

In the present work, we experimentally study how human subjects can bind perceptual and epistemic tracking to trace individuals that are neither directly perceived nor directly located. More specifically, we address the following two questions. First, how do human subjects track target individuals when information about the individuals' location is partial or indeterminate? Second, how many stationary targets can one epistemically track while moving among them, when only the directions to the targets (and an indication of when the targets are actually encountered) are known? We studied these two questions using a new paradigm, the Modified Traveling Salesman Problem (MTSP).

3. Modified Traveling Salesman Problem (MTSP)

3.1. Principles and Structure of the MTSP Paradigm

3.1.1. Principles, goals, instructions

The experimental paradigm, named the Modified Traveling Salesman Problem (hereafter MTSP), serves to study the strategies that can be used to bind the perceptual and epistemic tracking of individuals during a simulated navigation.

More specifically, it is designed to study how subjects can perform the epistemic tracking of a set of *n* invisible target individuals knowing only their current direction with respect to the vehicle and whether they had been contacted in the course of the vehicle's travels (Figures 1 and 2). In the MTSP task, to perform the epistemic tracking of targets essentially means to store and update a cognitive record of the history of contacts with *distinct* invisible targets and/or with their location, or to retrieve such a history of contacts via perception-based inferences.

This task was inspired by the Traveling Salesman Problem (hereafter TSP)—i.e. find the shortest path for joining a set of *n* points (Michie, Fleming, & Oldfield, 1968; Verblunsky, 1951)—which has been studied in various fields including psychophysics (MacGregor, Ormerod, & Chronicle, 1996; MacGregor, Ormerod, & Chronicle, 1999; MacGregor, Ormerod, & Chronicle, 2000; Ormerod & Chronicle, 1999).

The goal of the MTSP task is to reach each target by controlling the displacements of a small circular vehicle displayed on a computer screen. A set of *n* stationary target individuals ($\boldsymbol{i}_1$, $\boldsymbol{i}_2, \ldots, \boldsymbol{i}_n$) have to be visited serially. Subjects are instructed to

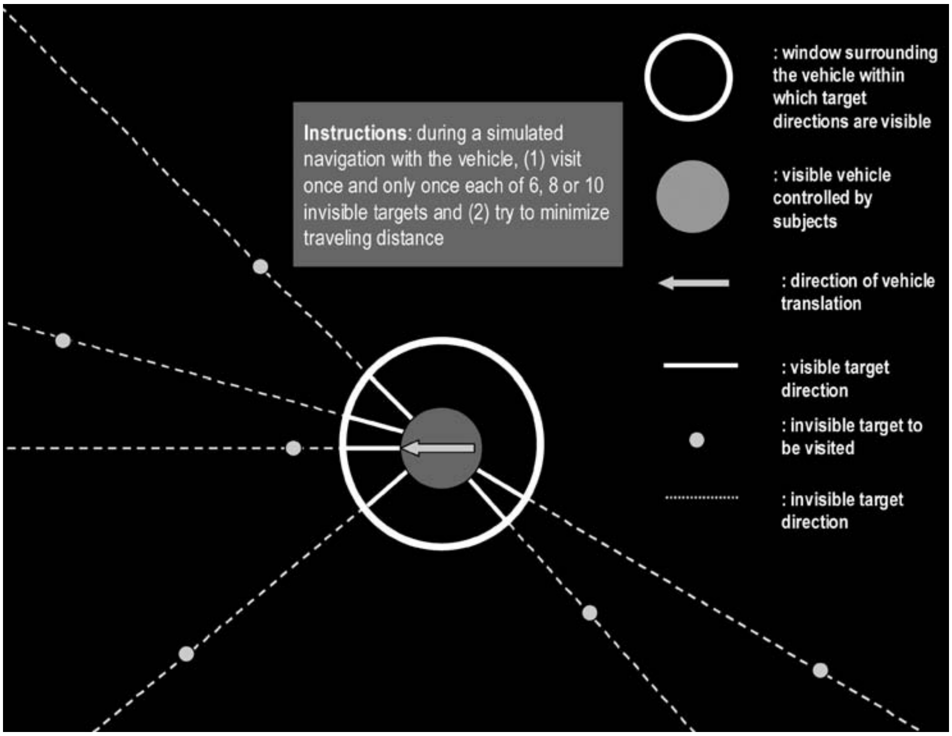

Figure 1. The standard MTSP (in *allocentric* condition according to the taxonomy explained in the text): Schematic presentation of the display shown to the subjects during the MTSP experiment. The arrow indicates the vehicle's frontal direction, which is its direction of translation. White directional segments point towards the invisible targets (6 in this case).

XML Template (2008) [10.1.2008–5:01pm] [1–31]
{TANDF_FPP}CPHP/CPHP_A_284219.3d (CPHP) [First Proof]

(1) move the vehicle so as to visit each invisible target *once and only once* (primary goal), and (2) try to minimize the distance traveled (secondary goal). The specific constraint in the MTSP task is that *only* the *directions* of targets from the current vehicle location are shown—indicated by directional segments displayed inside a circular window surrounding the vehicle (see Figures 1, 2 and 3). These directional segments are referred to in the article by the symbol 's_n' and will be referred to by particular names such as 's_1', 's_2' . . . 's_n'.[19] The directional segments are all identical in appearance. Each can be distinguished, or picked out, only on the basis of its orientation and motion, i.e., on the basis of its spatio-temporal trajectory. In order to

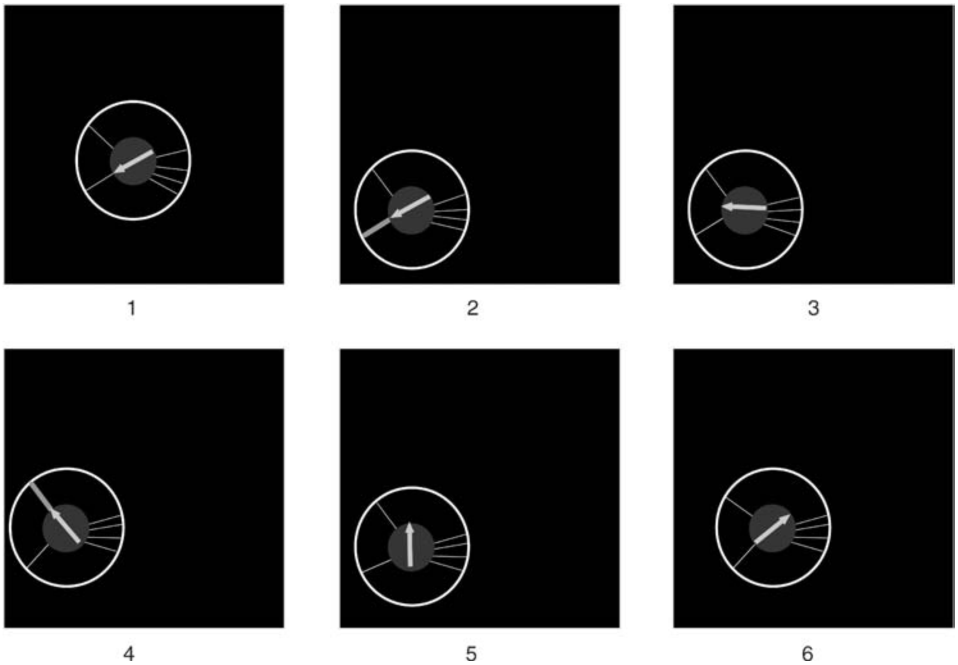

Figure 2. Representation of typical steps in the performance of MTSP (in "allocentric" condition): The images present the bottom left part of the screen in a sequence in which two targets out of 6 are being visited. The placement and orientation of each of the directional segments (pointing toward an associated target) are updated as a function of the vehicle's location. (2.1) The vehicle (the innermost grey circle) is surrounded by 6 directional segments that reveal the presence of 6 targets in the vicinity. The vehicle progressing in the direction indicated by the control arrow (white arrow in the figure), towards the invisible target in the bottom left of the screen. (2.2) The vehicle hits the left bottom target and the relevant segment changes its color. In images 2.2 and 2.4, contact with the target is represented by a bold grey segment instead of thin white line. In the actual MTSP display, the signal was a color change from white to red. (2.3) The control arrow is moved toward another target, above that which was formerly encountered in (2). (2.4) The second left target is hit. (2.5) The control arrow is rotated clockwise toward the locations of other targets (on the right). (2.6) After having visited the 2 left targets, the vehicle moves in direction of 4 targets located toward the right side of the screen.

Figure 3. Detail of a screenshot of the MTSP display. The image comes from a trial with six targets, in the allocentric condition (the arrow can move inside the circular blue disk). One can see the six white directional segments which point toward each target location.

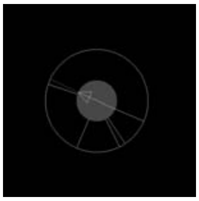

Figure 4. Detail of a screenshot of the MSTP display: when a target is hit, the segment that was followed changes its color from white to red. This image comes from a trial with 6 targets, in the allocentric condition.

reach a particular target, the subject typically follows a particular directional segment. When the target is hit, the segment being followed changes its color (see Figures 1, 2 and 4). Only this color change provides the crucial information that one has *reached* a target. In order to complete the task, subjects may draw different *perception-based inferences* by scrutinizing the motion of the directional segments, which co-vary with their motor commands. Our goal was to study the contribution of these perceptual inferences to epistemic tracking.

3.1.2. Hypothesis: deictic tracking strategies and current-status predicates in the MTSP task

The challenge of the MTSP task is to keep track of the *status* of each target (whether or not it has been visited) in a context where targets are invisible and directional segments are identical in appearance. Hence, subjects cannot directly pick out targets either by feature or by location. They must nonetheless *distinguish* each target individual from the other targets in order to avoid missed or multiple visits. Two questions are raised by this task. First, given the restriction that prevents the visual tracking of targets, and given subjects' memory limitations, how can subjects individuate and *keep track* of each target? Second, what strategies are available for solving the MTSP task? Our approach to these questions is derived from deictic theories (see §2.2). Below we suggest a way of experimentally studying possible strategies by examining performance in two different conditions of the MTSP task, which we refer to as the 'egocentric' and the 'allocentric' conditions.

Our hypothesis is that the MTSP task can be performed with deictic tracking strategies needed for (i) visually tracking directional segments (i.e., traceable cues relevant for the epistemic tracking), (ii) binding the segments with their updated status in the task and (iii) attempting the epistemic tracking of the targets via perception-based inferences. Deictic strategies in vision—in the sense defined in §2.2—use visual fixations and visual attention as pointers to information, that can rapidly move to different locations and be used in motor or cognitive procedures (aiming at providing the pragmatic and cognitive access to the targets).

In MTSP, because targets cannot be visually selected (since they remain invisible), directional segments themselves may be the targets of the pointers of visual tracking and attention (part (i) of the hypothesis). For instance, a subject could use perceptual-demonstrative references, such as the reference introduced by the demonstrative 'this' in the sentence "this segment [*eye fixation and visual attention directed at a token directional segment*] points toward the second target to be visited." As noted in §2.2, using such indexical representations requires the capacity to bind the tracked segments with a 'tag' or a current status predicate (part (ii) of the hypothesis). This point connects with the fact that, by hypothesis, perceptual-demonstrative references can be used to bind their referent(s) to epistemic predicates (§2.2.). Presumably, once a system has a set of indexed deictic pointers, it can predicate or 'tag' attributes to them. In the MTSP task, counting visits and avoiding revisits may involve the ascription of two types of current-status predicates. The first type involves predicates of ordering and counting, such as FIRST(***s***), SECOND(***s***), THIRD(***s***), N(***s***). The second type involves status predicates associated with the navigation goal, such as VISITED(***s***) and NOT-VISITED(***s***).

Both types of predicates convey information about context-dependant states of affairs that cannot be *directly* visible in the display. When a subject ascribes a current status predicate to a particular directional segment, he or she focuses on the *relation between the selected segment and the current step of his or her action*. Such an ascription is therefore context-dependant at several levels. It depends on the selected token individual, for which it is correct or incorrect (with respect to the subject's ongoing action). Moreover, its value is related to a specific time of performance: the ascription of the current-status predicate can be correct at a time t_1 but incorrect at a time t_2.

3.1.3. Perception-based inferences based on location-based and deictic direction-based strategies

The MTSP task allows investigation of the strategies subjects use in order to keep track of targets in a spatial layout with partial cues. Strategies are linked with specific types of visual information and perception-based inferences bound to demonstrative references. Generally speaking, in the absence of featural information, two types of strategies might be used for solving the MTSP task: *location-based* strategies and *direction-based* strategies. Each type of strategy might work in a global or grouping variety, meaning that there are four logically possible strategies. Each of these four logically possible strategies is described below, in Table 1, as a function of their respective frame of reference, relevant cues and memory requirements.

XML Template (2008) [10.1.2008–5:01pm] [1–31]
{TANDF_FPP}CPHP/CPHP_A_284219.3d (CPHP) [First Proof]

Table 1. Taxonomy of possible tracking strategies in the MTSP task: location-based and direction-based strategies.

Types of strategies		Frames of reference	Relevant cues for perception-based inferences	Memory requirements
Location-based	Global	*Allocentric* (i.e. indexed on environmental landmarks—e.g., screen edges)	(i) Vehicle location when a change of color occurs (target location). (ii) Segments' angular velocity (inferences about relative distance).	Recall of all *n* target locations in the allocentric framework.
	Grouping	*Allocentric*	(i) Vehicle location when a change of color occurs. (ii) Segments' angular velocity. (iii) Spatial proximity of subgroups of targets.	'Chunking' recall of a subset of the *n* target locations in the allocentric framework.
Direction-based	Global	*Deictic* (i.e. vehicle and gaze-fixation centered— relative to the current focus of attention)	(i) Segments' changes of color. (ii) Spatio-temporal continuity of segments' motion for visual tracking	Recall of the current-status of each of all *n* segments in the deictic frame of reference.
	Grouping	*Deictic*	(i) Segments' changes of color. (ii) Spatio-temporal continuity of segments' motion for tracking (esp. when crossing over). (iii) Spatial proximity of subgroups of directional segments.	'Chunking' recall of the current-status of a subset of the *n* segments in the deictic frame of reference.

A central characteristic for this taxonomy is that location-based and direction-based strategies use distinct frames of reference and distinct cues. *Location-based* strategies work to infer the allocentric locations of the set of targets and to compute a mental map of the *n* target locations. Thus, these tracking strategies rely on an allocentric frame of reference, with respect to which target locations can be inferred and memorized. For instance, by controlling the movement of the vehicle, subjects can gather information and draw inferences about spatial properties related to directional segments and targets. These inferences primarily include the segments' orientation and angular velocity as a function of the vehicle position and movement. Subject can also estimate the relative distance, and perhaps even the time-to-contact of a particular target $\boldsymbol{i}$, on the basis of the angular velocity of the directional segment pointing toward $\boldsymbol{i}$. The inference rule might be something like:

> If this segment [*selection of* $\boldsymbol{s}_1$] which points towards the first target [*i.e., target* $\boldsymbol{i}_1$] moves faster than the others [*i.e.,* $\boldsymbol{s}_2, \boldsymbol{s}_3, \ldots \boldsymbol{s}_6$], then the vehicle is closer to the first target than to the others [*i.e.,* $\boldsymbol{i}_2, \boldsymbol{i}_3, \ldots, \boldsymbol{i}_6$].[20]

In addition, target locations are known precisely when they are encountered, since they are located where the vehicle is standing at each time the color of a directional segment changes. As a consequence, in location-based strategies, the subject has to memorize the vehicle's allocentric location when this change occurs (relevant cue (i) for the location-based strategies). Since they rely on the recall of *n* target locations, location-based strategies put strain on visuo-spatial memory capacities, and do not use the vehicle as a crucial target for securing a frame of reference.

In contrast, the principle of direction-based strategies consists in the visual tracking of *n* directional segments, of the ones that represent to-be-visited or already-visited targets, so as to dynamically bind each directional segment with its correct current-status predicate. In contrast with location-based strategies, direction-based strategies are *deictic* in the sense that they use visual tracking and attention as a pointer to relevant information (see §2.2), and use deictic frames of reference centered on the gaze fixation point or relative to the current focus of attention (Ballard et al., 1997). In the MTSP task, direction-based strategies are deictic tracking strategies since the focus of attention has to split exclusively over the *directional segments* themselves, without keeping track of the target locations. In contrast to the global location-based strategy, one can carry out this strategy without keeping target locations in memory because one only needs to bind each selected segment with its current-status predicate. Typically, direction-based strategies aim first at a color change of a token directional segment (target contact), then update the current-status predicate (from NOT-VISITED($\boldsymbol{i}$) to VISITED($\boldsymbol{i}$)), and subsequently shift the focus of attention to another segment, until all segments' colors have been changed once.

Both location-based and direction-based *grouping* strategies rely on the same general principle: Grouping or 'chunking' either locations or segments, in order to compensate for memory limits. These strategies are useful for conditions with high number of targets (e.g., with $n = 10$ or higher value), in which the complete set or locations/segments cannot be memorized with their correct current status.

3.1.4. Available strategies in 'allocentric' and 'egocentric' conditions

In order to investigate which strategy was chosen by subjects, we introduced a new condition to the standard display, which we refer to as the *egocentric* condition (the terms 'allocentric condition' and 'egocentric condition' are simply a descriptor of the experimental setup).

In the standard *allocentric condition* (see Figure 5), targets are fixed within the screen reference-system. The vehicle icon and arrow move (according to the subject's commands) both in translation and rotation in the screen reference-system. In this condition, subjects may be able to use the location-based strategies for three reasons. Firstly, they can use the screen borders (and other stationary features of the visible environment) as the relevant allocentric landmarks for memorizing target location, since in this condition the target locations and inter-target spatial relations are fixed (stationary) in screen coordinates. Secondly, they can infer target positions with respect to the screen from the cues provided (directional segments motion and vehicle successive locations). Finally, they could construct in memory a spatial representation of target locations with respect to the screen frame of reference. However, location-based strategies may not be used in the egocentric condition.

In the *egocentric condition*, the vehicle icon and arrow remain fixed at the center of the screen (frontal direction upward). Only directional segments move according to

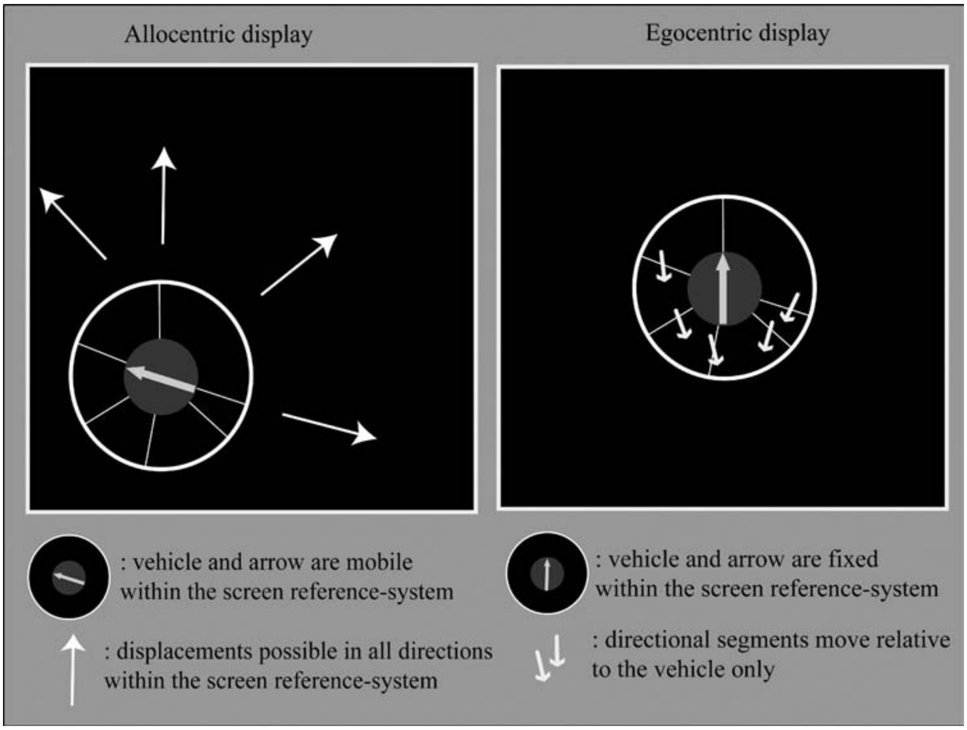

Figure 5. The difference between MTSP in the *allocentric* and *egocentric* conditions.

subject's commands, as if subjects were directly moving the display beneath their apparently stationary vehicle. As a consequence, target locations on the screen are no longer fixed, although target-target spatial relations are stable and coherent within an invisible rotating and translating frame of reference (which changes relative to the vehicle orientation and movements). Since the relevant frame of reference for locating targets moves and remains invisible, it is unlikely that subjects use it in a location-based strategy. As a consequence, one would expect the direction-based strategy to be chosen in the egocentric condition.

According to our experience explaining the goal of the MTSP task to subjects, the most natural and ecological strategy is the global location-based strategy. If this were true, the difficulty of the MTSP task should be different in the two conditions. Because the egocentric condition may not allow recovery of target locations within an allocentric frame of reference, this condition should be difficult, if not impossible, if subjects were not spontaneously using a direction-based deictic strategy.

3.2. Method

Subjects controlled the movement of a vehicle displayed on a computer screen, and were told to visit each of a set of *n* targets. Only the target directions with respect to the controlled vehicle were displayed, so subjects had no *direct* way to evaluate target distances or locations. They were instructed to move the vehicle to the vicinity of each target in order to: (1) visit each target once and only once (highest priority constraint); and (2) attempt to minimize the travel distance (lower priority constraint).

This task was performed under 6 experimental conditions. There were three values of the number of targets (6, 8 or 10), and two types of displays (allocentric and egocentric). In the *allocentric* condition, targets had fixed positions with respect to the screen frame of reference, while the displayed vehicle changed position and orientation under subject control.

In the *egocentric* condition, the vehicle position and orientation was fixed with respect to the screen, but target positions and directions with respect to the vehicle changed according to the displacements of the vehicle. Thus, in the egocentric condition, target locations, including their angular orientations relative to the screen, changed as the vehicle moved and rotated. We hypothesized that it would be very difficult to keep track of (invisible) target locations in this moving frame of reference.

Subjects were 9 adult volunteers, 5 males and 4 females. All of them were either students or colleagues of the laboratory, and were naïve with respect to the hypotheses under examination. They were seated in front of a 17'' computer screen at their usual working distance. They were given instructions prior to training with four simple examples (in the training trials the number of targets was reduced to $n = 3$ and $n = 4$). They were familiarized with the vehicle controls using the computer keyboard. The task was explained verbally through metaphorical descriptions such as "imagine that you are sitting in a moving robot and that you can control its movements." The Experiment was carried out in two sessions, with 3 blocks of 15 trials. During the first session, subjects performed the allocentric condition in three successive blocks, with 6, 8 and 10 targets. During the second session, subjects performed the egocentric condition, in three additional blocks with 6, 8, and 10 targets. The total duration of each block was approximately between half an hour and one hour and a half, but subjects could rest between trials and blocks. Subjects decided when each trial was finished, i.e. when *they* thought all targets had been reached. Within each block, target locations as well as the initial vehicle position and orientation were randomized. The average distances between closest locations on the screen (either targets or initial vehicle position) were about 197 pixels (for $n = 6$ targets), 167 pixels (for $n = 8$) and 144 pixels (for $n = 10$).

The visual display (resolution 1280 * 1024 pixels) was refreshed at 60 Hz. It consisted of a circular window (radius 64 pixels), in which the vehicle direction was indicated by an arrow, while target directions with respect to the vehicle were depicted by straight radial segments (see Figure 1). By pressing special keys in the keyboard, subjects could independently control the vehicle orientation and displacement. Vehicle displacement was only along the straight-ahead (forward) direction indicated by the arrow. The vehicle could not rotate while traveling. Subjects had to stop the vehicle before rotating it (i.e. rotate the directional arrow). The angular velocity of the rotations was 30 deg/sec—the vehicle needed 12 sec to complete an entire revolution. The default linear velocity of the vehicle was 75 pixels/sec—the vehicle needed 13 sec to cross the width (1000 pixels) of the display on a straight line. Subjects could decrease the vehicle velocity up to 20% of the default velocity by pressing the control key or the middle mouse button. (Subjects usually use this option when they are approaching the targets' vicinity in order to secure contact with the target.). The transitions from halt to default velocity, or from reduced to normal velocity (and conversely) occurred 'instantaneously', i.e. at the very next video frame (refresh rate was 60 Hz). Instantaneous vehicle position and orientation were recorded in a data file for further processing.

Subjects' performance was evaluated by two criteria corresponding to the two goals of the task (which had different priorities). The first criterion was the *percentage of correct trials*. An individual trial was classified as correct if all targets had been visited once and only once, regardless of the travel time or distance. The second criterion, only applied to correctly performed tests, was the *normalized travel distance*, defined as the ratio of the actual travel distance to a normalization distance. The normalization distance was computed for each individual trial as the sum of the distance between initial location and the closest target, plus the distance between the first target to the closest remaining target, and so on to the last target. The normalization distance is slightly greater than the optimal solution of the traveling salesman problem, i.e. the shortest path starting from the initial point and passing through all targets. Percentage of correct trials and normalized travel distance were used to evaluate subject performance. We considered scores of 50% correct trials or better to indicate a capacity to solve the MTSP task. This threshold of 50% is conservative, as random choice of targets would yield a percentage of correct trials equal to about 4% for 6 targets and less than 1% for 8 and 10 targets (that is, equal to 100 $(n-2)!/(n-1)^{n-2}$).

3.3. Results

Eight of the nine subjects were able to perform the task, even with 10 targets in the egocentric condition. An example of a trajectory in the 10 target condition is shown in figure Figure 6A. Only one subject systemically failed to perform the task in either of the two conditions. The trajectories produced by this subject (Figure 6B) were the convex hulls of the sets of targets. Thus, we excluded this subject.[21]

Percent correct trials (PCT) tends to decrease with the number of targets, although this effect was significant only in egocentric condition between 8 and 10 targets ($t = 3.33$, $p = 0.012$) and when egocentric and allocentric conditions were pooled between 6 and 10 targets ($PCT_6 = 83.3\%$ ($SD = 15.2\%$), $PCT_{10} = 66.7\%$ ($SD = 13.7\%$), $t = 2.65$, $p = 0.01$). No statistical differences were found between egocentric and allocentric conditions, regardless of the number of targets. In fact, the difference always tended to be in favor of egocentric conditions (see Table 2). That decrease of PCT between 8 and 10 targets is consistent with a constant probability of assignment error of about 3%,[22] and therefore, cannot be interpreted as an abrupt fall-off of performance reflecting subjects' limited resources (see also section 3).

No clear statistical difference between conditions was found for the normalized travel distance (NTD) score. Only a marginal increase of travel distance from allocentric to egocentric conditions was found, by pooling NTD for all target numbers ($NTD_{allocentric} = 1.26$ ($SD = 0.087$), $NTD_{egocentric} = 1.33$ ($SD = 0.16$), $t = 1.83$, $p = 0.07$). The analysis of trajectories showed that subjects changed their choice of the next target to visit from time to time (see Figure 6A for an example). This suggests that they were not able to accurately estimate or compare the distance to unvisited targets at the time they chose which target to visit next. The normalized traveled distance was clearly sub-optimal and was found to increase with the number of targets. The traveled distance was 30% to 40% higher than the performance predicted by the nearest neighbor heuristic, which itself generates a distance about 10% above the optimal solution. As expected, these results differ from those of standard TSP experiments.[23] In contrast to PCT, NTD was higher in the egocentric condition compared with the allocentric condition. The allocentric display, which showed the vehicle displacement in display coordinates, seemed to improve the ability to evaluate target distances, but did not reduce the target assignment error.

4. Discussion

4.1. Discrepancies with Previous Findings

We found that subjects were able to carry out the MTSP task correctly, visiting all targets in up to 70% of the trials, in both allocentric and egocentric conditions. This finding is surprising for several reasons.

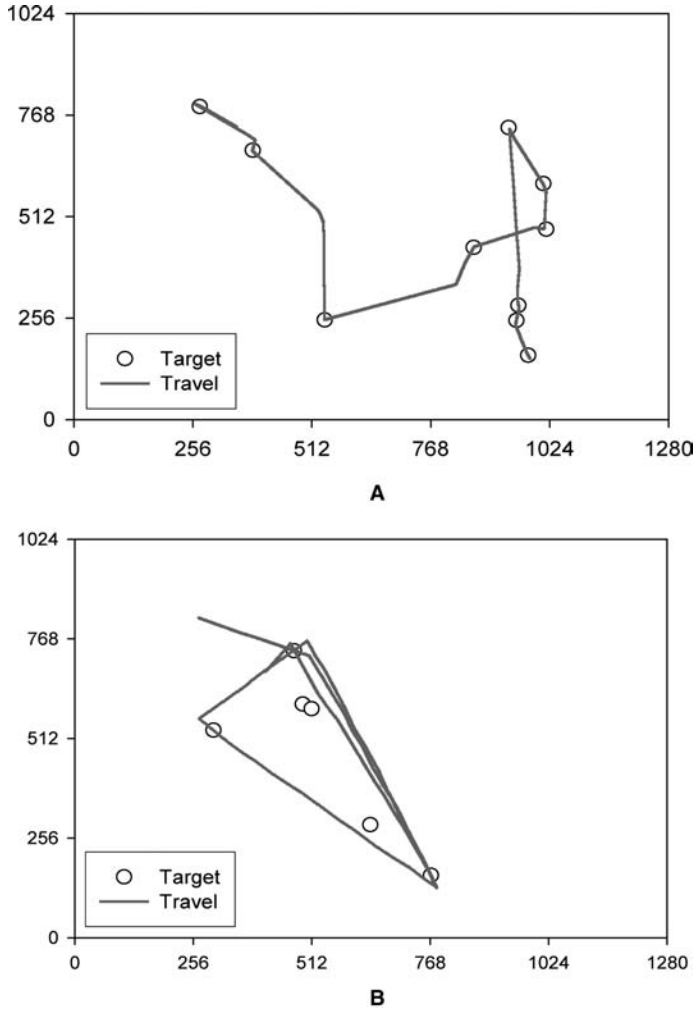

Figure 6. Examples of trajectories (solid lines) reconstructed from two different subjects. Targets' locations are shown by small circles. X/Y coordinates are expressed in pixels. **A**: 10 targets in egocentric condition. **B**: 6 targets in allocentric condition. The subject in B fails to visit the three inner targets while visiting twice or more the outer ones.

Table 2. Percentage of correct trials (PCT) in each condition.

	6 targets	8 targets	10 targets
PCT Allocentric	82.5% (SD = 14.9%)	72.5% (SD = 14.4%)	64% (SD = 18.0%)
PCT Egocentric	86.7% (SD = 16.3%)	84% (SD = 5.96%)	70% (SD = 6.7%)

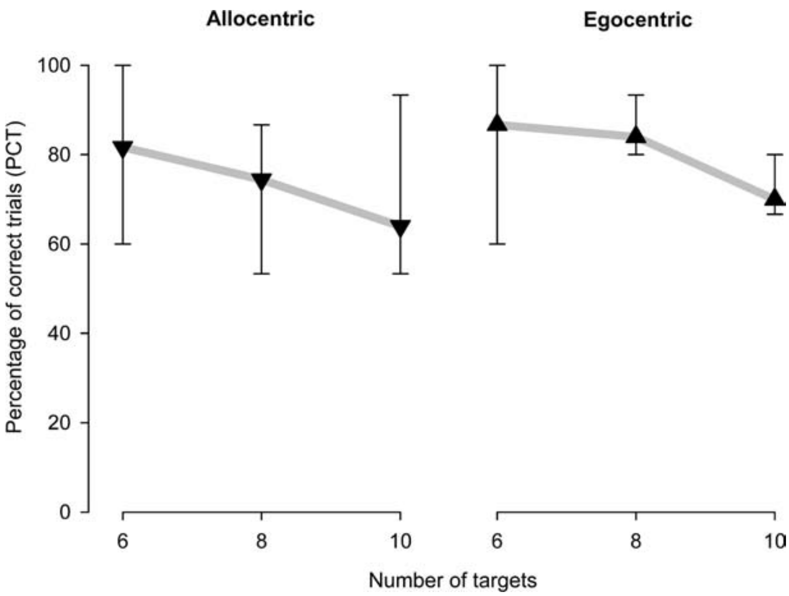

Figure 7. Percentage of correct trials (PCT), averaged for 8 subjects, with 6, 8 and 10 targets in both allocentric and egocentric conditions. Taking into account the large variability between subjects, there is no clear increase of mistakes in egocentric condition as compared to allocentric one. Error bars represent minimum and maximum performance for each condition.

First, the result dispels our concern that the MTSP task with 8 and 10 targets might be beyond subjects' capacity. Our finding also contrasts with at least two kinds of results in the literature. For example, with Multiple Object Tracking (MOT) methodology (Pylyshyn & Storm, 1988) it was found that that people can track only about four or five target objects.[24] Our finding suggests, therefore, a difference between the visual tracking of *n* independently-moving targets and the epistemic tracking of *n* static targets on the basis of their direction—as in the MTSP task.

In addition, human short-term or working memory is known to have storage capacity limits, sometimes presented (controversially) as a ''single, central capacity limit averaging about four chunks'' (Cowan, 2000). It has also been suggested that visual working memory has a limit of about 4 items (Luck & Vogel, 1997). Our findings suggest a surprisingly higher limit: The number of 8 or 10 targets far exceeds the most common estimations of the limits of short-term memory, visual working memory or visual tracking.

Our findings do not support our expectation that the global location-based strategy would be the most natural in the MSPT task (see §3.1.4). Choice of this unique strategy would have led to poorer performance—the global location-based strategy requires a high memory load in order to retain locations of the 8–10 targets, as well as the current status predicates/tags VISITED(***i***) or NOT-VISITED(***i***) associated with them.

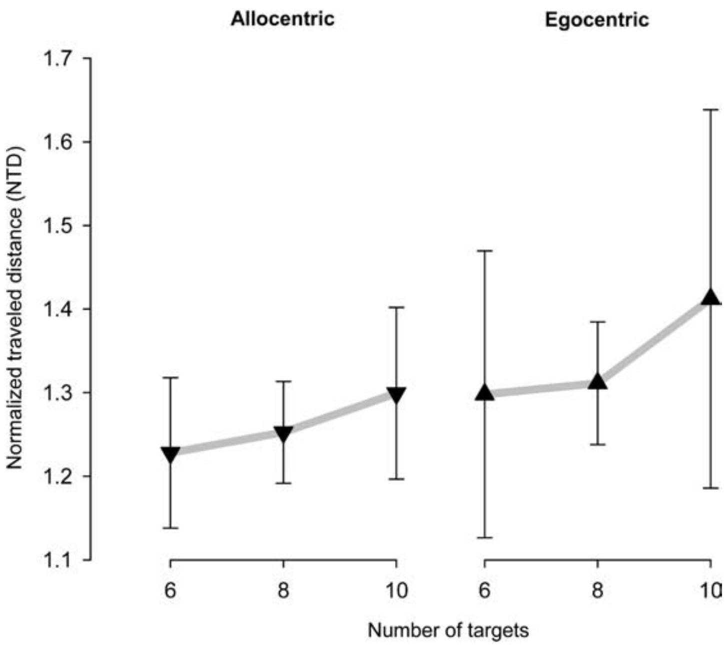

Figure 8. Normalized traveled distance (NTD), averaged for 8 subjects, with 6, 8 and 10 targets in both allocentric and egocentric conditions. There is a slight increase as a function of targets number. Traveled distance is also slightly increased in egocentric condition as compared to allocentric one (note also the increase of variability). Error bars represent standard deviation.

4.2. The Results Explained by the Use of Different Strategies

In order to account for the discrepancies between our findings and those of previous studies, we focus on the hypothesis that subjects used different tracking strategies to complete the MTSP task (see Table 1) and that the use of segment strategies allowed them to succeed in the egocentric condition (with any number of targets) and also in allocentric condition trials with large numbers of targets. Since direction-based tracking strategies are deictic (see §2.2), this hypothesis emphasizes the use of deictic tracking strategies as a crucial means for 'overcoming' resource limitations, which is consistent with former findings associated with the Deictic View in vision (Ballard et al., 1997; Pylyshyn, 2000). If this hypothesis is true, then our findings provide, to the best of our knowledge, the first empirical evidence of the systematic use of a deictic tracking strategy in a task that (only) at first glance has to be solved by location-based memory.

The global location-based strategy relies on memory of *all* target locations within a stationary allocentric frame of reference, and on the capacity to infer target locations on the basis of spatial cues. Hence, limitations of visuo-spatial working memory are its relevant constraints. Given the most common estimations of the limits of visuo-spatial working memory (Baddeley, 2000, 2003; Cowan, 2000; Logie, 1995;

Luck & Vogel, 1997), it is unlikely that subjects could recall all target locations with 8 or 10 targets. This observation is consistent with explicit reports about performing the MTSP task with 8 or 10 targets in allocentric and egocentric conditions, according to which it seems impossible to memorize all locations. However, one cannot exclude the possibility that subjects in the allocentric condition used the *grouping* location-based strategy in order to perform well with 8 or 10 targets.

The most interesting and surprising result remains, however, the performance on the egocentric condition, because location-based strategies are unlikely to be effective to explain the high performance in this condition. Several converging arguments support the view that only the *segment* strategies might explain performance in the egocentric condition. Firstly, visual inferences based on the *global location-based* strategy do not seem sufficient to explain performance for the egocentric condition, even with the smallest number of targets. The global location-based strategy is unlikely to be available in the egocentric condition since the location of targets in the screen frame of reference is constantly rotating and translating (§3.1.4). As a consequence, subjects cannot use the environmental stationary features (screen borders and other salient landmarks) as a frame of reference, because target locations are not fixed in relation to the screen and even move off the area of the screen. Secondly, location-based strategies in this condition would require path and angular integration of all past vehicle displacements. Given that path and angular integrations are corrupted by uncertainty and noise accumulation (e.g., NcNaughton, Battaglia, Jensen, Moser, & Moser, 2006), one could expect an important fall-off of subjects' PCT in the egocentric condition as compared to allocentric condition, in which these kinds of integration are not required. However, there is no evidence of this fall-off in the results. Thirdly, when questioned about the location of target in the egocentric condition, subjects reported that it was impossible for them to figure out where the targets were, and usually described the choice of segment strategies.

While subjects in the egocentric condition could not have used the *global* location-based strategy, they may have used the *grouping* location-based strategy. For example subjects may have encoded the relative location of the closest targets as they moved through the space. In order to address this possibility, a small number of additional subjects were run on the experiment, with only 6 targets in the allocentric and egocentric conditions. They were instructed to visit each target once and only once, and to keep track of target locations. At the end of some trials, we debriefed these subjects concerning their memory for locations. If subjects used spatial memory in the egocentric condition, they might be able to recall, at least, the relative *distance* between each target and the vehicle, and therefore detect when targets moved off the screen borders (an event which happens frequently during most trials in the egocentric condition). We found, however, that *none* of the subjects noticed that this happened. In addition, every subject in the egocentric condition in this exploratory test reported that they were unable to recall where the targets had been—they had no idea where the targets were, except that they could see from the segments at the end of the trial which direction they were in. Thus, the most plausible explanation for how the task was carried out in the egocentric condition (and perhaps in the allocentric with 8 or 10 targets) is that subjects used direction-based deictic strategies (i.e., subjects combined visual tracking of directional segments with vision-based inferences about the targets). Given the limitation of visual working memory; the global location-based strategy might explain performance only for the smallest number of targets ($n = 6$) in allocentric condition.

4.3. The Deictic Direction-Based Strategies: Applicability and Cost

In contrast to the global location-based strategy, direction-based tracking strategies may explain our findings. When using direction-based strategies, subjects do not compute target locations, but rather *track local or proximal cues*, i.e. subsets of directional segments in MTSP. The primary focus of direction-based strategies is to track and focus on the directional segments, and so this strategy is less sensitive to visuo-spatial memory limitations. Nonetheless, the relevant constraint in this strategy is the capacity of visual tracking. As previously mentioned, human subjects can track up to four or five visual objects in parallel, and this constraint should apply when using segment strategies because directional segments in the MTSP task have similar characteristics to visual objects in classical (MOT) tracking experiments (i.e., they have identical features and the historical continuity of each segment's motion is the only available clue to specify or access its status as a distinct spatio-temporal individual).

What differences between the MOT and MTSP paradigms may be the most relevant for explaining our findings with the MTSP task? First, part of the explanation could be that targets in MTSP do not follow arbitrary motion (in contrast to MOT), but change in accordance with the voluntary movement of the vehicle through a stationary configuration of targets in two dimensions. This coherence between the movements could facilitate tracking and reduce the errors of current-status predicate ascription.

Second, in the MOT task subjects observe the motion of targets passively, whereas MTSP is an (active) subject-performed task in which the subject chooses movements in order to visit the next target. This might help memory performance since memory is improved in subject-performed tasks (Senkfor, Van Petten, & Kutas, 2002; Zimmer et al., 2001), even in the case of navigation in virtual environments (Brooks, Attree, Rose, Clifford, & Leadbetter, 1999). Also visual memory for object location benefits from active interactions with layout (Shelton & McNamara, 2001b).

Finally, in cases where the use of the global direction-based strategy does not circumvent memory limitations, grouping (or 'chunking') subsets of segments may allow performance to exceed the visual memory limits—grouping or 'chunking' strategies are often used to increase memory (Cowan, 2000, pp. 87–93), and to allow the enumeration of larger numbers of items (Klahr, 1973; Mandler & Shebo, 1982; Van Oeffelen & Vos, 1982). Moreover, effects of regularities in the point layout (esp. point clusters) have been found in performance in standard Traveling Salesman Problem (MacGregor et al., 1999; Ormerod & Chronicle, 1999). The addition of grouping strategies to segment tracking might allow subjects to make use of statistical clusters of segments to improve performance. Because the vehicle motion is under subjects' control it may be possible to switch between the chunks and the individual segments. For example, with 8 target individuals ($\boldsymbol{i}_1$ to $\boldsymbol{i}_8$) and 8 segments ($\boldsymbol{s}_1$ to $\boldsymbol{s}_8$), a subject might first group 4 adjacent segments on one side, or segments that happen to be close together into one chunk (set A, $\boldsymbol{s}_1$ to $\boldsymbol{s}_4$), and track the 4 others (set B, $\boldsymbol{s}_5$ to $\boldsymbol{s}_8$). Subjects might then switch to tracking the individuals in set A, treating set B ($\boldsymbol{s}_5$ to $\boldsymbol{s}_8$) as a group. By switching between groups and individual targets subjects might be able to use a time-sharing strategy. The operation of grouping allows subjects to *draw deictic inferences* such as: "Given that all targets in this group g_1 have been visited targets; and that this segment belongs to g_1, therefore, $\boldsymbol{s}_1$ and $\boldsymbol{i}_1$ must have already been visited." In this manner, the subject can dynamically recover (or retrieve) the correct current-status predicate on any member of the group if he or she has correctly tracked the group and avoided swaps between group members and other tracked targets. To avoid losing track of segments that cross over one another, subjects would have to switch between groups and individuals sufficiently quickly to anticipate that a pair of segments were about to cross and select those segments to track.

5. Concluding Remarks

We have argued that the Deictic View is a viable alternative to the Constructivist View of perceptual cognition. In the framework of research on perceptual-demonstrative reference and deictic tracking strategies, we have presented an experimental study of the capacity to bind the perceptual and epistemic tracking of several invisible target individuals in an informationally impoverished environment: the MTSP paradigm. We found that subjects were able to carry out the MTSP task with an accuracy approaching 70% of the trials, in both allocentric and egocentric conditions, with 6, 8 and 10 targets. Perception-based inferences based on global location-based strategy and grouping location-based strategy might explain performance for the smallest number of targets ($n = 6$), but do not explain performance for high number of targets ($n = 8, 10$) and for the egocentric condition. Perception-based inferences based on the direction-based and deictic tracking strategies might explain performance for high number of targets ($n = 8, 10$) in the large-*n allocentric condition* and in the *egocentric condition*, although it would require a chunking and time-sharing process that has not been modeled in detail. In conclusion, further experiments and theoretical investigations remain to be carried out on how deictic strategies and perceptual inferences contribute to the integrated tracking of distal individuals and on the kind of task that would make the use of a location-based strategy mandatory.

Notes

[1] On the memory of object locations, see, namely, Burgess, Jeffery and O'Keefe (1999), Morris, Nunn, Abrahams, Feigenbaum & Recce (1999), Posma, & De Haan (1996), Shelton & McNamara (2001a).

[2] Works on perception-based inferences have been developed with regard to, namely, the visual routines needed to analyze a perceived scene (Cavanagh, 2004; Ullman, 1984), the role of graphical representations in reasoning and logic (Allwein & Barwise, 1996; Stenning, 2002; Stenning & Oberlander, 1995), the relations between perceptual-demonstrative reference and reasoning (Campbell, 2002; Evans, 1982; Garber & Goldin-Meadow, 2002; Goldin-Meadow, 2003), spatial thinking (Eilan, McCarthy, & Brewer, 1993; Gattis & Dupeyrat, 2000) or heuristics and bounded rationality (Gigerenzer et al., 1999; Todd & Gigerenzer, 2000).

[3] An anonymous referee of this article suggested that the Constructivist View may be reducible to a Representationalist View of cognition. We remain, however, reluctant to identify the two views. One of the reasons for keeping them distinct is that, at least with regard to perceptual cognition, a theory can belong to the Deictic View and confront the Constructivist View without abandoning the concepts of representation or intentional content.

[4] Based on the criteria discussed in this section, we consider that crucial principles of the Deictic View are shared, *mutatis mutandis*, by a number of philosophers who study perceptual reference and vision scientists who study tracking. The former include works by Campbell (2002), Andy Clark (1998; 1999; 2001), Austen Clark (2000; 2004a; 2004b), Mohan Matthen (2004; 2005). The latter include works by Pylyshyn et al. (Pylyshyn, 1989, 2000; Pylyshyn & Storm, 1988), Ballard et al. (Ballard, 1991; Ballard et al., 1997; Hayhoe & Ballard, 2005; Triesch, Ballard, Hayhoe, & Sullivan, 2003) and Land et al. (Land & Hayhoe, 2001; Land, Mennie, & Rusted, 1999).

[5] In the present article, each single character in italicized and bold fonts, such as '***i***' or '***o***', stands for a proper name referring to a single individual. This notation is introduced in Bullot (2006) and Bullot and Rysiew (2007) for discussing the problem of perceptually or epistemically tracking an individual as being the same individual over time (i.e., the Problem of Singular Cognition).

[6] In philosophy, the theory of the role of descriptions in reference originates mainly in Russell's theory of descriptions (Russell, 1905)—see Neale (1990). The contrast we draw in the article between perceptual-demonstrative tracking/identification and epistemic tracking is reminiscent of Russell's distinction between knowledge by acquaintance and knowledge by description (Russell, 1910-11).

[7] Traditional philosophical theories of demonstrative identification (Evans, 1982; Kaplan, 1989; Strawson, 1959) emphasize the epistemic importance of linking sensory-motor tracking and perceptual experience with conceptual representation. In such theories, it is generally assumed that perceptual-demonstrative reference is realized by the capacity to make relevant sensory-motor discriminations to perceive the referent—see Clark (2000, 130–163) for a philosophical overview. Following Strawson (1959: 18–20), Evans (1982: 121, 146), and the psychology of attention (Kahneman, Treisman, & Gibbs, 1992; Treisman, 1992), Campbell (2002; 2004) makes the further claim that selective attention might be a crucial mechanism for accomplishing perceptual-deictic reference.

[8] Many recent experimental works have studied eye fixations in frameworks consistent with the Deictic View (see for instance Ballard et al., 1992; Ballard et al., 1997; Findlay & Gilchrist, 2003; Hayhoe, 2000; Hayhoe et al., 2003; Henderson, 2003; Land & Furneaux, 1997; Land & Hayhoe, 2001; Land et al., 1999; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1996; Triesch et al., 2003).

[9] On visual tracking, see for example the works of Pylyshyn et al. (e.g., Pylyshyn, 2003; Pylyshyn & Storm, 1988), Logan (1995), Yantis (1992), Cavanagh & Alvarez (2005). On the role of attention in epistemic perception and the knowledge directed at individuals, see Bullot & Rysiew (2007) and Bullot (in press).

[10] Pointing gestures have been studied in the literature on joint attention, language development and communication (Baldwin, 1993; Bruner, 1983; Butterworth & Grover, 1990; Clark, 2003; Gómez, 2007; Kita, 2003; Tomasello, Carpenter, & Liszkowski, 2007).

[11] See, e.g., Jeannerod (1988) and Milner & Goodale (1995).

[12] See, e.g., the review of Baldwin (1993).

[13] See, e.g., the works of Goldin-Meadow et al. (Garber & Goldin-Meadow, 2002; Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001) and Kirsh (Kirsh, 1995; Kirsh & Maglio, 1995).

[14] As discussed by Milner and Goodale (1995), Land et al. (1999) and Findlay & Gilchrist (2003), the visual control of non-cognitive (or non-epistemic) aspects of action does not necessarily need *eye pointing* (i.e., the deictic use of the visual system). This suggests a link between the deictic use of eyes as pointers and the performance of epistemic actions.

[15] Pylyshyn (1989, 2001) has suggested that deictic reference in the visual system is closely related to a mechanism, called a *visual index* or *FINST* (from FINgers of INSTantiation), which links individual object tokens with information upon which beliefs and actions can be based. A visual index serves like a pointer from a (conceptual) representation of an object to an actual object in the scene.

[16] The concept of *indexical* representation has been traditionally used in the philosophy of language and semantics in order to refer to terms like pronouns 'I', 'my', 'you', demonstrative pronouns like 'that', 'this', or adverbs like 'here', 'now', 'tomorrow'. The classical observation relative to the use of these terms is that: (1) the referent depends on the context of use; (2) their meaning provides a rule which determines the referent in terms of certain aspects of the context (Kaplan, 1989, p. 490); (3) 'true' demonstratives—as David Kaplan called them in his seminal article *Demonstratives* (Kaplan, 1989, p. 490)—need an associated demonstration (typically: a pointing) in order to determine their referent (this means that the linguistic rules which govern the use of the Kaplanian true demonstratives are not sufficient to determine their referent in all contexts). Since then, indexical representations have been studied in many domains, from semantics and philosophy of language (e.g., Dokic, 1996; Lewis, 1979; Lyons, 1977; Perry, 2000, 2001; Recanati, 1993) to the modeling of action and planning in artificial intelligence (e.g., Agre, 1997, pp. 230–234; Lespérance & Levesque, 1995).

[17] See the notion of knowledge by acquaintance in Russell's sense (Russell, 1910, 1956a, 1984 [1913]), and its development as *Russell's principle* (Bullot & Rysiew, 2007; Evans, 1982; McDowell, 1990; Peacocke, 1983).

[18] About object-based perceptual inferences, see for instance Ullman (1984), Ballard et al. (1997), Campbell (2002, pp. 84–113), Hodgson et al (2000), Pylyshyn (2003), Stenning (2002, pp. 54–92); on the relevance of the theory of object or singular files for the Problem of singular Cognition, see Bullot & Rysiew (2007).

[19] Given the notation adopted in this article (see note 5), a symbol such as "s_1" is a proper name for a particular directional segment. The name "s_1" refers to one and only one token individual (in this case, the directional segment s_1 on the display). This notation is used only for convenience. Of course, subjects in MTSP do not possess any proper name for identifying particular directional segments. According to our hypothesis, instead of proper names, they use deictic pointers (or, perceptual demonstratives) for tracking directional segments and attempting the epistemic tracking of invisible targets.

[20] In MTSP, if Θ is the angle between straight-ahead and the direction of the segment pointing to some target, and if the vehicle is moving at a constant velocity v, then the closer the target object is the faster will the angle be changing. Although this does not tell the observer how

far the vehicle is from the target, it does provide some guidance about the relative distance to various targets (for small non-zero Θ).

[21] This exclusion is justified by three reasons. First, her results were atypical (PCTallo 6 targets = 26.7%, 13.3% in PCTego 6 targets, and 0% in all other conditions). Second, it was impossible to compute NTD on all trials with her scores (NTD can be computed only for correct trials). Third, she apparently did not understand the nature of the task or did not understand why her strategy was not reliable for performing the task.

[22] A constant assignment error probability for each target yields a PCT roughly proportional to the target number—more precisely: P(correct trial) = [1− P(target assignment error)]$^n \approx 1 - n$P(target assignment error).

[23] MacGregor and co-workers (1996) have measured human performance in solving the standard traveling salesman problem for 10 and 20 points. They found that, on average, subjects solutions lie somewhere between the optimal and the nearest neighbor solutions (about 4% above optimal length for 10 points). Clearly a number of experimental conditions may account for this difference, the most likely among them is that in MTSP, contrary to the standard TSP, the exact location of points and therefore the in-between distances are not directly given to subjects. Moreover, our data are computed from the actual travel distance including detours, while Euclidean distance between connected points are used in standard TSP.

[24] This experimental paradigm was initially used by (Pylyshyn & Storm, 1988), and has now been developed in several experimental contexts (Pylyshyn, 2003; Sears & Pylyshyn, 2000). Pylyshyn and colleagues hypothesized that the human visual system uses visual indexes, or that humans could perceptually 'split' their selective attention over multiple visual objects. In the paradigm presented by Pylyshyn & Storm (1988), participants visually track a specified subset of identical, randomly moving objects in a display. The members of the subset to be tracked (the targets) were identified by briefly flashing them, prior to the onset of movement. According to the FINST model (Pylyshyn, 1989), targets designated in this way are automatically indexed. During the tracking task, targets were indistinguishable from the nontarget distractors, which made the historical continuity of each target's motion the only available clue as to the targets' identity. Visual indexes provided a deictic link with targets which enabled them to be tracked. Participants tracked the target objects for 5 to 10 seconds, after which either a target or a distractor was indicated by superimposing a bright square over it. The participants' task was to determine whether this indicated object was a target or a distractor. The authors found that performance in this multiple object tracking task was particularly high for subsets of up to five elements: subjects could track simultaneously up to five target objects at an accuracy approaching 90%.

References

Agre, P. E. (1997). *Computation and human experience*. Cambridge, MA: Cambridge University Press.

Allwein, G., & Barwise, J. (Eds.). (1996). *Logical reasoning with diagrams*. New York: Oxford University Press.

Attneave, F., & Farrar, P. (1977). The visual world behind the head. *American Journal of Psychology*, *90*(4), 549–563.

Attneave, F., & Pierce, C. R. (1978). Accuracy of extrapolating a pointer into perceived and imagined space. *American Journal of Psychology*, *91*(3), 371–387.

Bach, K. (1987). *Thought and reference*. Oxford: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

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

Baldwin, D. A. (1993). Infant contributions to the achievement of joint reference. In P. Bloom (Ed.), *Language Acquisition, Core Readings*. Cambridge, MA: MIT Press.

Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, *48*, 57–86.

Ballard, D. H. (1996). On the function of visual representation. In K. Akins (Ed.), *Problems in Perception. Proceedings, Simon Fraser Conference on Cognitive Science, February 1992*. Oxford: Oxford University Press.

Ballard, D. H. (1997). *An introduction to natural computation*. Cambridge, MA: MIT Press.

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions: Biological Sciences*, *337*(1281), 331–338.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. N. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences*, *20*(4), 723–767.

Ballard, D. H., Hayhoe, M. M., Salgian, G., & Shinoda, H. (2000). Spatio-temporal organization of behavior. *Spatial Vision*, *13*(2,3), 321–333.

Blaser, E., Pylyshyn, Z. W., & Holcombe, A. O. (2000). Tracking an object through feature space. *Nature*, *408*, 196–199.

Brooks, B. M., Attree, E. A., Rose, F. D., Clifford, B. R., & Leadbetter, A. G. (1999). The specificity of memory enhancement during interaction with a virtual environment. *Memory*, *7*(1), 65–78.

Brooks, R. A. (1999). Intelligence without reason. *Cambrian Intelligence: The Early History of the New IA* (pp. 133–185). Cambridge, MA: MIT Press.

Bruner, J. S. (1983). *Child's Talk: Learning to Use the Language*. Oxford: Oxford University Press.

Bullot, N. J. (2006). The principle of ontological commitment in pre- and postmortem multiple agent tracking. *Behavioral and Brain Sciences*, *29*, 466–468.

Bullot, N. J. (in press). Attention, information and epistemic perception. In G. Terzis & R. Arp (Eds.), *Information and the biological sciences*. Cambridge, MA: MIT Press.

Bullot, N. J., & Rysiew, P. (2007). A study in the cognition of individuals' identity: Solving the problem of singular cognition in object and agent tracking. *Consciousness and Cognition*, *16*, 276–293.

Burgess, N., Jeffery, K. J., & O'Keefe, J. (Eds.). (1999). *The hippocampal and parietal foundations of spatial cognition*. Oxford: Oxford University Press.

Butterworth, G., & Grover, L. (1990). Joint visual attention, manual pointing, and preverbal communication in human infancy. In M. Jeannerod (Ed.), *Attention and performance XIII: Motor representation and control* (pp. 605–624). Hillsdale, NJ: Lawrence Erlbaum Associates.

Campbell, J. (2002). *Reference and consciousness*. Oxford: Clarendon Press.

Campbell, J. (2004). Reference as attention. *Philosophical Studies*, *120*, 265–276.

Cavanagh, P. (2004). Attention routines and the architecture of selection. In M. I. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 13–28). New York: Guilford Press.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Science*, *9*(7), 349–354.

Churchland, P. S., Ramachandran, V. S., & Sejnowski, T. J. (1994). A critique of pure vision. In C. Koch & J. L. Davis (Eds.), *Large scale neuronal theories of the brain* (pp. 23–60). Cambridge, MA: MIT Press.

Clark, A. (2000). *A theory of sentience*. Oxford: Clarendon Press.

Clark, A. (2004a). Feature-placing and proto-objects. *Philosophical Psychology*, *17*(4), 443–469.

Clark, A. (2004b). Sensing, objects, and awareness: Reply to commentators. *Philosophical Psychology*, *17*(4), 553–579.

Clark, A. J. (1998). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.

Clark, A. J. (1999). An embodied cognitive science? *Trends in Cognitive Sciences*, *3*, 345–351.

Clark, A. J. (2001). Visual experience and motor action: Are the bonds too tight? *The Philosophical Review*, *110*(4), 495–519.

Clark, H. H. (2003). Pointing and placing. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 243–268). Hillsdale, NJ: Erlbaum.

Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Dennett, D. C. (1991). Multiple drafts versus the cartesian theater. *Consciousness explained* (pp. 101–138). Boston: Little, Brown and Company.

Dennett, D. C. (2001). Are we explaining consciousness yet? *Cognition*, *79*, 221–237.

Dokic, J. (1996). The dynamics of deictic thoughts. *Philosophical Studies*, *2*(2), 179–204.

Dretske, F. I. (1969). *Seeing and knowing*. Chicago: The University of Chicago Press.

Dretske, F. I. (1981). The objects of perception. *Knowledge and the flow of information* (pp. 153–168). Cambridge, MA: MIT Press.

Eilan, N., McCarthy, R., & Brewer, B. (Eds.). (1993). *Spatial representation*. Oxford: Blackwell.

Evans, G. (1982). *The varieties of reference*. Oxford: Oxford University Press.

Farrell, M. J., & Thomson, J. A. (1998). Automatic spatial updating during locomotion without vision. *The Quarterly Journal of Experimental Psychology*, *51A*(3), 637–654.

Farrell, M. J., & Thomson, J. A. (1999). On-line updating of spatial information during locomotion without vision. *Journal of Motor Behavior*, *31*(1), 39–53.

Findlay, J. M., & Gilchrist, I. D. (2001). Visual attention: The active vision perspective. In M. Jenkin & L. Harris (Eds.), *Vision and attention* (pp. 83–103). New York: Springer-Verlag.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active Vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Ganea, P. A., Shutts, K., Spelke, E. S., & DeLoache, J. S. (2007). Thinking of things unseen. *Psychological Science*, *18*(8), 734–739.

Garber, P., & Goldin-Meadow, S. (2002). Gesture offers insight into problem-solving in adults and children. *Cognitive Science*, *26*, 817–831.

Garnham, A. (1989). A unified theory of some spatial relational terms. *Cognition*, *31*, 45–60.

Gattis, M., & Dupeyrat, C. (2000). Spatial strategies in teasoning. In W. Schaeken, G. De Vooght, A. Vandierendonck, & G. D'Ydewalle (Eds.), *Deductive reasoning and strategies* (pp. 153–175). Mahwah: New Jersey.

Gigerenzer, G., Todd, P. M., & ABCgroup (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Goldin–Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516–522.

Gómez, J.-C. (2007). Pointing behaviors in apes and human infants: A balanced interpretation. *Child Development*, *78*(3), 729–734.

Goodman, N. (1968). *The languages of art*. Oxford: Oxford University Press.

Graziano, M. S. A., Hu, X. T., & Gross, C. G. (1997). Coding the locations of objects in the dark. *Science*, *277*, 239–241.

Hayhoe, M. M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*(1/2/3), 43–64.

Hayhoe, M. M., & Ballard, D. H. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 189–194.

Hayhoe, M. M., Bensinger, D. G., & Ballard, D. H. (1998). Task constraints in visual working memory. *Vision Research*, *38*(1), 125–137.

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*, 49–63.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504.

Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2000). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neuroscience*, *12*(5), 894–907.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*, 575–577.

Humphreys, G. W. (1999). Neural representation of objects in space: A dual coding account. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience* (pp. 165–182). Oxford: Oxford University Press.

Hurford, J. R. (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences*, *26*(3), 261–316.

Jeannerod, M. (1988). *The neural and behavioural organization of goal-directed movements*. Oxford: Clarendon Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. K. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.

Karn, K. S., & Hayhoe, M. M. (2000). Memory representations guide targeting eye movements in a natural task. *Visual Cognition*, *7*(6), 673–703.

Keane, B. P., & Pylyshyn, Z. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*, 346–368.

Kirsh, D. (1995). The intelligent use of space. *Artificial Intelligence*, *73*(1–2), 31–68.

Kirsh, D., & Maglio, P. (1995). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.

Kita, S. (Ed.). (2003). *Pointing: Where language, culture, and cognition Meet*. Hillsdale, NJ: Erlbaum.

Klahr, D. (1973). A production system for counting, subitizing and adding. In W. G. Chase (Ed.), *Visual information processing* (pp. 527–546). New York: Academic Press.

Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions: Biological Sciences*, *352*, 1358.

Land, M. F., & Hayhoe, M. M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*, 3559–3565.

Land, M. F., Mennie, N., & Rusted, J. (1999). The role of vision and eye movements in the control of activities of daily living. *Perception*, *28*, 1311–1328.

Lespérance, Y., & Levesque, H. J. (1995). Indexical knowledge and robot action: A logical account. *Artificial Intelligence*, *73*, 69–115.

Lewis, D. (1979). Attitudes de dicto and de se. *Philosophical Review*, *88*(4), 513–543.

Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, *28*, 103–174.

Logan, G. D., & Sadler, D. D. (1996). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 493–529). Cambridge: MIT Press.

Logie, R. H. (1995). *Visuo-spatial working memory*. Hove, Hillsdale: Lawrence Erlbaum Associates.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Lyons, J. (1977). Deixis, space and time. *Semantics, (Volume 2)* (pp. 636–724). Cambridge, MA: Cambridge University Press.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (1996). Human performance on the traveling salesman problem. *Perception & Psychophysics*, *58*, 527–539.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (1999). Spatial and contextual factors in human performance on the traveling salesperson problem. *Perception*, *28*, 1417–1427.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition*, *28*(7), 1183–1190.

Mandler, G., & Shebo, B. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*, 1–22.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B200*, 269–294.

Matthen, M. (2004). Features, places, and things: Reflections on Austen Clark's theory of sentience. *Philosophical Psychology*, *17*(4), 497–518.

Matthen, M. (2005). *Seeing, doing, and knowing: A philosophical theory of sense perception*. Oxford: Oxford University Press.

McDowell, J. (1990). Peacocke and Evans on demonstrative content. *Mind*, *99*(394), 255–266.

Michie, D., Fleming, J. G., & Oldfield, J. V. (1968). A comparison of heuristic, interactive and unaided methods of solving a shortest-route problem. *Machine Intelligence*, *2*, 245–255.

Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Morris, R. G., Nunn, J. A., Abrahams, S., Feigenbaum, J. D., & Recce, M. (1999). The hippocampus and spatial memory in humans. In N. Burgess, K. J. Jeffery, & J. O'Keefe (Eds.), *The hippocampal and parietal foundations of spatial cognition* (pp. 259–289). Oxford: Oxford University Press.

NcNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nature Reviews Neuroscience*, *7*, 663–678.

Neale, S. (1990). *Descriptions*. Cambridge, MA: MIT Press.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, *46*(3), 461–488.

Ormerod, T. C., & Chronicle, E. P. (1999). Global perceptual processing in problem solving: The case of the traveling salesperson. *Perception & Psychophysics*, *61*(6), 1227–1238.

Palmer, S. E. (1999). *Vision science, photon to phenomenology*. Cambridge, MA: MIT Press.

Peacocke, C. (1983). *Sense and content: Experience, thought, and their relations*. Oxford: Oxford University Press.

Peacocke, C. (2001). Does perception have a nonconceptual content? *The Journal of Philosophy*, *98*(5), 239–264.

Perry, J. (2000). *The problem of the essential indexical and other essays, expanded edition*. Stanford, CA: CSLI Publications.

Perry, J. (2001). *Reference and reflexivity*. Stanford: CSLI Publications.

Posma, A., & De Haan, E. H. F. (1996). What was where? Memory for object locations. *The Quarterly Journal of Experimental Psychology*, *49A*(1), 178–199.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, *32*(1), 65–97.

Pylyshyn, Z. W. (2000). Situating vision in the world. *Trends in Cognitive Sciences*, *4*(5), 197–207.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*, 127–158.

Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*, 157–238.

Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independents targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*(3), 179–197.

Recanati, F. (1993). *Direct reference: from language to thought*. Oxford: Blackwell Publishers.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive change in scenes. *Psychological Science*, *8*(5), 368–373.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1157–1165.

Russell, B. (1905). On denoting. *Mind*, *14*(56), 479–493.

Russell, B. (1910). Knowledge by acquaintance and knowledge by description. *Proceedings of the Aristotelian Society*, *11*, 108–128.

Russell, B. (1956a). In R. C. Marsh (Ed.), *Logic and knowledge, Essays 1901–1950*. London: George Allen & Unwin.

Russell, B. (1956b). On the nature of acquaintance. In R. C. Marsh (Ed.), *Logic and knowledge, Essays 1901–1950*. London: George Allen & Unwin.

Russell, B. (1984 [1913]). *Theory of knowledge: The 1913 manuscript*. London: Georges Allen & Unwin.

Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259–290.

Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processes. *Canadian Journal of Psychology*, *54*, 1–14.

Senkfor, A., Van Petten, C., & Kutas, M. (2002). Episodic action memory for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neuroscience*, *14*(3), 402–419.

Shelton, A. L., & McNamara, T. P. (2001a). Systems of spatial reference in human memory. *Cognitive Psychology*, *43*, 274–310.

Shelton, A. L., & McNamara, T. P. (2001b). Visual memories from nonvisual experiences. *Psychological Science*, *12*(4), 343–347.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*(1), 16–20.

Stenning, K. (2002). *Seeing reason, image and language in learning to think*. Oxford: Oxford University Press.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, *19*, 97–140.

Strawson, P. F. (1959). *Individuals: An essay in descriptive metaphysics*. London: Methuen.

Strawson, P. F. (1974). *Subject and predicate in logic and grammar*. London: Methuen.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI* (pp. 457–478). Cambridge, MA: MIT Press.

Todd, P. M., & Gigerenzer, G. (2000). Précis of Simple heuristics that make us smart. *Behavioral and Brain Sciences*, *23*, 727–780.

Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). A new look at infant pointing. *Child Development*, *78*(3), 705–722.

Treisman, A. (1992). Perceiving and re-perceiving objects. *American Psychologist*, *47*, 862–875.

Triesch, J., Ballard, D. H., Hayhoe, M. M., & Sullivan, B. T. (2003). What you see is what you need. *Journal of Vision*, *3*, 86–94.

Ullman, S. (1984). Visual routines. *Cognition*, *18*, 97–159.

Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*, 615–665.

Van Oeffelen, M., & Vos, P. (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory and Cognition*, *10*, 396–404.

Verblunsky, S. (1951). On the shortest path through a number of points. *Proceedings of the American Mathematical Society*, *2*(6), 904–913.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*, 295–340.

Zimmer, H. D., Cohen, R. L., Guynn, M. J., Engelkamp, J., Kormi-Nouri, R., & Foley, M. A. (2001). *Memory for action: A distinct form of episodic memory?* Oxford: Oxford University Press.