Spatial Cognition Adapts to Social Context

Nicholas D. Duran (nduran@memphis.edu) **Rick Dale (radale@memphis.edu)**
Department of Psychology, The University of Memphis, Memphis, TN, 38152

Abstract

We present a dynamical systems model that captures human perspective-taking behavior in a visuospatial mental rotation task. The task requires participants to interpret an ambiguous request for a visual referent, either by taking their own perspective (egocentric; "the referent is on *my* right") or the perspective of the requester (other-centric; "the referent is on *your* right"). Our primary interest lies in how perspective-taking behavior and spatial cognition adapt to socially-driven information. To do so, we manipulate whether the participant shares the same social status as their assumed interaction partner. We report critical influences of social role on response choice and on the processing demands required to enact the response. Furthermore, we discuss these results in the context of our computational model, showing that simple socially-induced constraints can produce rich behavioral patterns.

Keywords: social perspective-taking; visuospatial processing; social roles; dynamical systems model

Introduction

In this paper we present a human experiment and computational model that demonstrate a simple message: Social and spatial cognition may interact in deep and interesting ways. It is already known that spatial cognition is used on a regular basis during linguistic interaction. For example, if seated across from another person at a dinner table, instructions to pass the salt may require mindful preparation, noting the relative position of sender and receiver in order to disambiguate spatial perspectives. Previous work has shown how perspectives are coordinated actively during linguistic interaction. In this paper, we show that the connection between social and spatial processing may be deeper: *For whom* you are retrieving the salt may influence underlying spatial perspective-taking processing (e.g., an authority figure vs. someone of the same social status). In this paper we illustrate, using a dynamical systems model, that spatial cognition adapts in this way to social status. And in a human experiment, we show that this intuition bears out in real human performance.

Background

During linguistic interaction, a primary goal is to understand and to be understood. To do so, language users are likely to consider what another knows or is likely to know, as well as their needs or abilities, i.e., "common ground," to tailor messages that will maximize mutual understanding (Brennan, Galati, & Kuhlen, 2010; Clark, 1996). In doing so, speakers draw on a variety of pragmatic and contextual sources of information, as in whether their partner is young or old (Newman-Norlund et al., 2009; Horton & Spieler, 2007), or comes from the same geographic region (Isaacs & Clark, 1987). By doing so, partner cues allow for more fine-tuned expectations in how to best adapt to unique audiences (Fussell & Krauss, 1992)

One cue that is easily tracked is a person's social role. Social roles have been shown to be powerful memory cues in remembering who said what in a conversation (Senay & Keysar, 2009; Stangor, Lynch, Duan, & Glass, 1992), and are also responsible for source misattributions where statements from people of the same race or gender are often confused (Senay & Keysar, 2009). These studies suggest that expectations based on role are used to predict what another knows, or what perspective they are likely to take. Indeed, even in culturally imposed roles, as in between a boss and a employee, subordinates are quite accurate in gauging an authority's internal thoughts and feelings, and adjusting their own behavior to better accommodate the authority's perceived needs (Snodgrass, 1992).

Because social role is a categorical distinction that one either possesses or does not, it is a salient constraint on linguistic behavior, and is similar to cues of a statistical, phonological, or prosodic nature found in probabilistic models of sentence processing (MacDonald, Pearlmutter, & Seidenberg, 1994). Accordingly, being exposed to simple, often single "bits" of social information can easily guide perspective-taking modes of responding (see Brennan et al., 2010; Galati & Brennan, 2010).

To explore how partner information can influence both human and model spatial cognition, we developed a spatial perspective-taking task, motivated by Schober (1993), that allows systematic manipulation of the social role of a virtual partner. The task simulates a hypothetical interaction where a person, standing around a table, is asked by someone else for one of two objects laid out on the table. If the person making the request (i.e., speaker) says, "Give me the one on the right", and is standing directly across from the seated person (i.e., addressee), the proper response is ambiguous. Does the addressee assume their "right" (egocentric) or the speakers "right" (other-centric)? Moreover, if the participant shares a social role with a speaker, is the participant likely to assume that the speaker is more apt to take their perspective?

To create distinct and naturally occurring social roles, we turned to a novel paradigm based on Amazon's Mechanical Turk (www.mturk.com). This service allows users connected via Internet to participate in simple tasks, or "HITs", that require human expertise. In this set-up, those that post HITs are called "Requesters" and those that do the HITs are called "Workers." Thus, social roles that naturally occur in Mechanical Turk can be put to use. In the current study, participants are told that that they will be connected to either a fellow worker or to the requester. In actuality, both these social agents are simulated, but a carefully controlled ruse leads the majority of participants into believing that these agents are real. The social roles of worker and requester can be distinguished along a subordinate/authority dimension, where participants interact with a fellow worker ("same-status" condition), or with a requester ("authority-status" condition). We consider the requester to have greater authority because this is the person who recruits and pays participants for each HIT, and thus has the power to withhold payment if the HIT is not satisfactorily completed.

In the sections that follow, we provide several hypotheses on how participants might interpret ambiguous requests in the absence of explicit feedback. In general, if interpretations are made with mutual understanding in mind, we expect listeners to use partner specific information to guide spatial perspective-taking - both in the endpoint response (i.e., egocentric or other-centric) and the processing time involved in making the response. We then detail a dynamical systems model and summarize experimental results.

Hypotheses for Effect of Status

In the spatial perspective task described above, greater cognitive effort is involved in taking another's perspective. This is because, from trial to trial, the partner's location moves around the table while the participant's location remains fixed, creating a situation where the partner can be offset from the participant by 90 and 180 degrees. To take the other's perspective, the participant must "see" the table from the other's visual perspective, and this requires a participant to mentally rotate to the other's line of sight (also see Schober, 1995). Such an "other-centric" process of visuospatial transformation is cognitively taxing because the greater the rotation - as in going from 0 to 180 degrees versus 0 to 90 degrees - the greater the spatial distance to traverse (see Zacks & Michelon, 2005, for a review). As such, we expect other-centric responders to have generally slower processing times than egocentric responders. However, we also expect response choice and response time to be modulated by the social status of the conversational partner. But how might this unfold? Will participants interacting with an authority- or same-status partner be more or less other–centric, and will mental rotation become more or less easier? For possible answers, we consider several hypotheses motivated by past research.

Authority-Status vs. Same-Status Affiliation. In communicative interactions where partners are attempting to establish mutual understanding, they do so with the goal of minimizing the processing effort for each other (Clark, 1996). But if it appears that one of the partners is failing to understand, the other will expend greater cognitive effort to ensure mutual understanding (Clark & Wilkes-Gibbs, 1986, *principle of least collaborative effort*). Conversely, if one of the partners is deemed to be more competent or to be in greater control of the interaction, the other might assume that this partner will invest greater effort. In terms of the current study, this means that participants interacting with the authority-status partner (i.e., the "requester") are more likely be egocentric because the requester has greater perceived command over the interaction.

However, the principle of collaborative effort is not the only influence on communicative interactions. In other research, it has been shown that when a status difference exists between two social agents, the subordinate actor will allocate greater cognitive resources to assessing the other's intentions, at least to a greater extent than the authoritative actor will do for the subordinate (Rutherford, 2004). This account suggests there will be tension between the egocentric and other-centric response choice, with at least some participants willing to invest greater cognitive effort to take the requester's perspective (who is in the authority role). We predict that participants interacting with a requester are just as likely to be egocentric or other-centric. On the other hand, for participants interacting with a same-status worker, the social attributes of their partner are less likely to elicit a particular perspective choice. In the absence of any such constraints, and given that the egocentric response is easiest, we expect greater egocentric responding in the same-status condition.

We can also make predictions about the processing time required to take a partner's perspective. In general, egocentric responders will be much faster than other-centric responders. But for other-centric responders alone, those interacting with the authority-status requester will be faster in making mental rotations. One reason for this is that "authority-status" participants are simply more motivated to track the other's perspective.

In what follows, we briefly describe a dynamical systems model that can capture the predictions made here in an implemented framework that uses simple cognitive parameters. We then describe an experiment in which participants believe they are interacting with a same-status or authority-status partner.

Dynamical Systems Model

The model is an integration of a prominent dynamical system that captures two-choice decisions (Tuller, Case, Ding, & Kelso, 1994) with a diffusion-type model that accumulates activation over iterated time until a threshold is reached (Ratcliff, 1978). Duran, Dale, and Kreuz (2010) used the model to capture basic perspective-taking decisions (other- vs. egocentric choices), and found that it fit both endpoint-response distributions and the real-time dynamics of the decisions themselves. The model effectively treats a decision process as the dynamics of a single state, $x(t)$, as it descends over time into one or another fixed-point attractor (see Equation 2, explained below). As this descent occurs, the model accumulates activation to a threshold in order to determine the decision (akin to reaching a threshold in a diffusion model). Space restrictions limit our simulation of the current data, but the model permits an explicit computational instantiation of the social-role hypotheses.

The model is exemplified in Figure 1a-c as a potential landscape, characterized by Equation 1 and is solved by Equation 2.

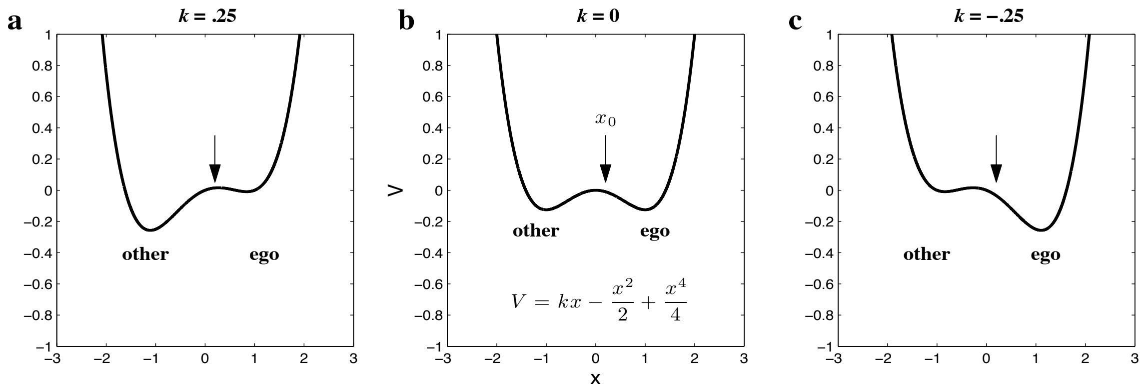

Figure 1: The potential landscape with two attractor basins.

$$V = kx - \frac{x^2}{2} + \frac{x^4}{4} \quad (1)$$

$$x_{t+1} = x_t + \lambda(-k_t + x_t - x_t^3) + \xi \quad (2)$$

The $x_t = x_0$ parameter value represents the model's initial state bias, reflecting whether models are slightly more egocentrically or other-centrically biased at the time-scale of individual decisions. Under a subtle noise signal, ξ, the model descends from the saddle point between potential wells until it falls into one or another attractor. Under noise, $x(t)$ may briefly fluctuate around the saddle point before this occurs, capturing initial indecision when the model first faces the task. Once a decision is made, the potential parameter k, which produces a bias to respond on one side or another (see Equation 3), is shifted from trial to trial in response to one choice or another (e.g., descending on the left causes k to bias the system towards future leftward responses).

$$k_{t+1} = k_t + \delta \quad (3)$$

Trial-by-trial change in k can reflect the gradual commitment of strategy that participants tend to exhibit in spatial perspective-taking tasks (Carlson, 1999).

These models capture a range of two-choice behaviors (Raczaszek-Leonardi, Shapiro, Tuller, & Kelso, 2008; Rooij, Bongers, & Haselager, 2002), and we used it to model the choice between other-centric and egocentric choices in the current task. The purpose is to examine the influence of different sets of parameters on (1) the response distribution of participants in two different social-role conditions, and (2) the response time of the individual decisions in these two conditions. The parameter sets we chose explicitly represent the hypotheses as framed in the previous section: same-status affiliation vs. authority-status affiliation. These hypotheses lead to the questions: Does working with a perceived authority figure result in competing biases for other-centric and egocentric responding that are not present with someone of a similar social status? Does one get more efficient at responding in the spatial perspective of an authority figure, or someone of the same status? Of course, if these social identities do not influence spatial perspective-taking biases or processes, then no effects could be modeled (nor present in the study). A set of illustrative parameters (shown in Figure 2) shows how the model behaves if social status causes a shift in these cognitive parameters of the model. In summary, this model (Duran et al., 2010) naturally captures the intuition that two interactive systems converge dynamically to produce endpoint-response and response-time changes in ways consistent with what we (verbally) predicted in the previous section.

Human Experiment

Method

Participants were led to believe that they were being connected to another Mechanical Turk worker (at random), or to the requester (who, incidentally, was paying the participant for their service). These partners could communicate to participants in a "one-way chat" environment. The partners explained the task as an attempt to debug software within an arbitrary instruction-following task environment. Both partners were recordings of the same male speaker, but as the worker, the speaker claimed he was reading directions from a script - another factor that gives the worker less authority when compared to the requester.

A trial began by automatically placing a participant's mouse cursor underneath an empty table (Figure 3, Position 1). Identical verbal instructions were then given by one of the two simulated partners that directed the participant to select one of the folders placed on the tabletop. At this time, the participant pressed a "GO" button (also located at the bottom of the screen), and the folders and the partner's location appeared. Folders were arranged diagonally (as shown in Figure 3, Position 2 and 3), vertically, or horizontally. The participant was to then drag the selected folder to the simulated partner who was located somewhere around the table, either at the other sides of the table (as shown in Figure 3, Position 4). Based on the layout in this ambiguous example, if the participant heard, "Give me the folder on the right," and selected the folder at Figure 3, Position 3, they would have interpreted

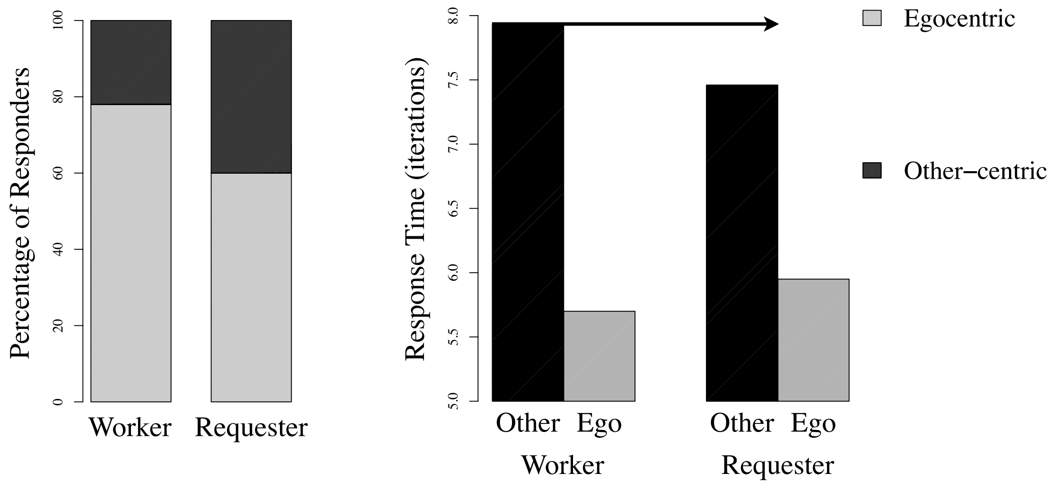

Figure 2: Sample model results, Worker: $x_0 = .2$, $k_0 = .1$; Requester: $x_0 = .2$, $k_0 = .2$. The endpoint response distribution do not vary for authority-status ("requester") condition, and fewer cycles are required to process other-centric responses.

the instructions egocentrically. However, if the participant selected the folder at Figure 3, Position 2, he or she would have interpreted the instructions other-centrically.[1] As participants were moving toward and selecting a folder, we were rapidly sampling the x,y coordinates of their mouse cursor at every 25 ms (a sampling rate of 40Hz). Response time analysis are based on the time between pressing "GO" and selecting an initial folder.

The partners verbal instructions ("right," "left," "front," and "back") were strategically paired with particular folder/partner location combinations to create 20 shared-perspective and 20 critical ambiguous trials. It is during the critical trials that a participant's perspective interpretation is revealed. The analysis below are based on these trials.

Participants

Eighty Amazon's Mechanical Turk "workers" participated in this study and were paid 75 cents for their effort. We only included participants that believed that they were interacting with a real partner. In a series of follow-up questions after the experiment, participants who indicated in any way that the speaker was not real had their data discarded. We achieved a nearly 75% deception rate, retaining 59 participants. In addition, 60 trials, or 2.58% overall, were removed because they exceeded three standard deviations above the response time mean for selecting an initial folder (i.e., 3892.06 ms).

Results

Participant's Response Choice

Participants were identified as being consistently other-centric, egocentric, or a mixture of these two response types. A proportion score for egocentric responding was computed for each participant, with a score over 70% resulting in an egocentric identification, and a score below 30% resulting in an other-centric identification. A score between 30% and 70% was mixed. In the same-status/worker condition, there were 15 egocentric, 7 other-centric, and 1 mixed responder. In the authority-status/requester condition, there were 23 egocentric, 10 other-centric, and 3 mixed responders. Figure 4 shows the proportion of these responder types across social status.

Chi-square tests conducted between egocentric and other-centric responders in the same-status/worker condition show a significant preference for responding egocentrically,

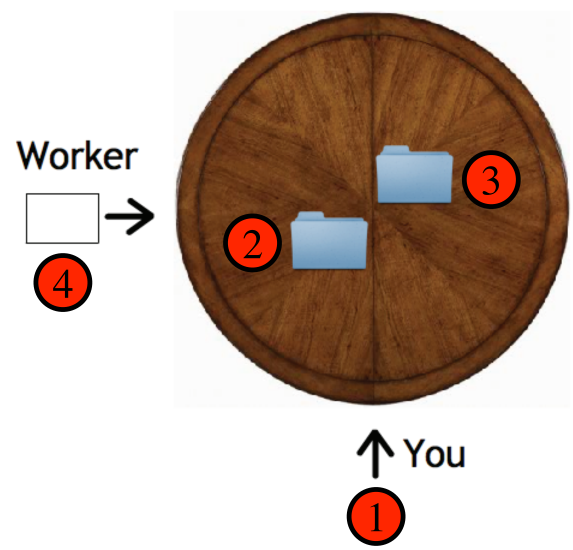

Figure 3: The experimental interface for capturing egocentric or other-centric behavior in a single computerized task. *Note: Numbered circles were not seen by participants but are shown for explanatory purposes only.*

[1] A version of the game can be accessed at: cognactive.org/perspectiveTask

$\chi^2(1) = 5.12$, p = .02, and in the authority-status/requester condition there was no strong preference for a particular perspective choice, $\chi^2(1) = 2.91$, p = .09, although the preference for egocentrism is approaching significance. As predicted, the two conditions do vary slightly in their relative distributions. It should also be noted that a fairly large proportion of participants were willing to take their partner' s perspective, comparable to rates reported in related spatial perspective-taking studies (see Tversky & Hard, 2009).

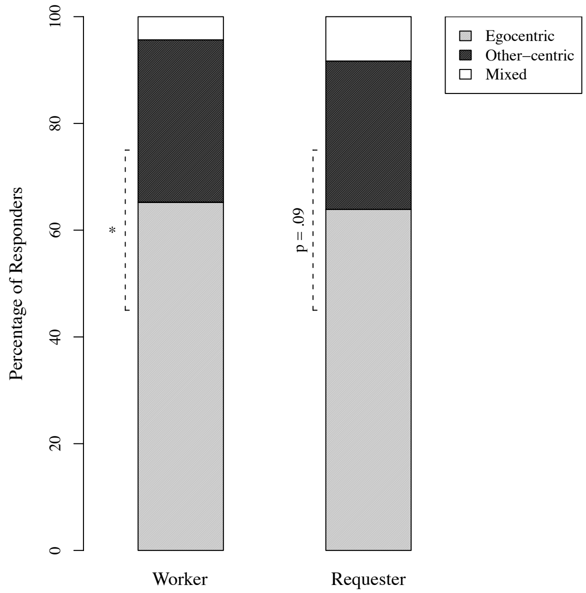

Figure 4: The response choices of human participants as a function of interacting with another Mechanical Turk worker or requester.

Participant's Response Times

The time taken for consistent egocentric and other-centric responders to select a folder in the critical trials was recorded and subjected to a mixed-effects statistical analysis, comparing differences between responder type and social status (as fixed factors).[2] We also included random factors for subject and item,[3] ensuring that the addition of each variable resulted in a best-fitting model through a stepwise model comparison, using log-likelihood ratio tests (see Baayen, Davidson, & Bates, 2008). The analysis were conducted using the lmer package in the R statistical software. In this package, p-values are computed with 10,000 Monte Carlo Markov Chain simulations, using lmer's pvals.fnc function. We report these p values, as well as the standardized effects from each model.

The overall model revealed a significant main effect for responder type (egocentric vs. other-centric), *beta* = 0.59, $p < .01$, and a significant interaction between responder type and social status (same-status/worker vs. authority-status/requester), *beta* = 0.82, p = .01. Given the significant interaction between the two factors, follow-up simple main effects were conducted for social status at each responder type. For the same-status/worker condition, other-centric responders were 1002.13 ms slower than egocentric responders, *beta* = 1.28, $p < .001$, and in the authority-status/requester condition, other-centric responders were 394.32 ms slower than egocentric responders, *beta* = 0.77, $p < .01$. Lastly, we compared differences between responder type across social status. We found that other-centric responders are 549.65 ms slower when interacting with what they believe to be a Mechanical Turk worker, *beta* = 0.74, p = 0.02. There are no statistically significant differences for egocentric responders.

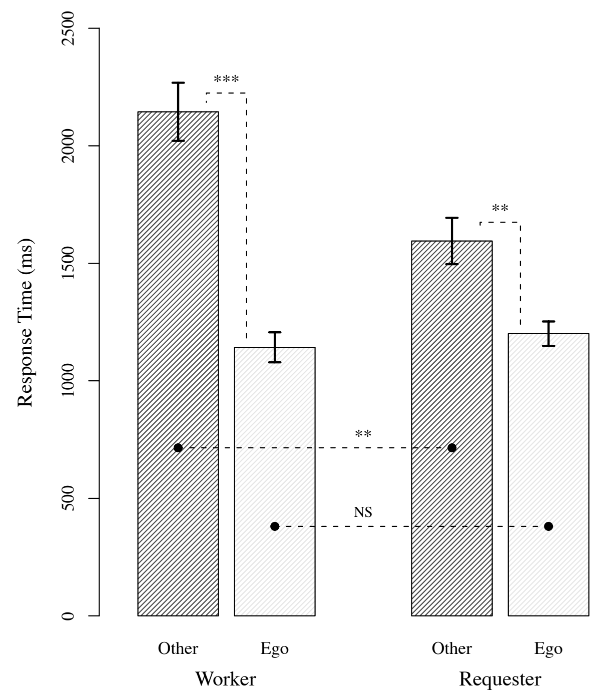

Figure 5: Participant response times and standard errors in selecting an initial folder during critical trials. Other-centric responders take greater time, particularly when interacting with a Mechanical Turk worker, the same-status partner.

General Discussion

In this paper, we take seriously the notion of Brennan and colleagues that single sources of partner-oriented information are rapidly assimilated during language use, influencing both perspective choice and processing demands (Brennan et al., 2010; Brennan & Hanna, 2009). We did so by uniquely focusing on an addressee's ability to interpret another's perspective in a visuospatial mental rotation task - with particular emphasis on how interpretations are shaped by a partner who shares

[2] For time taken, we collapsed over partner positions at 0, 90, and 180 degrees. In a separate analysis not reported here, there was evidence for mental rotation by other-centric responders, with a monotonic increase of response time from 0 to 90 to 180 degrees.

[3] Item corresponds to the four orientations of the folder configuration.

the same social role, or is in a position of authority.

The first set of hypotheses we introduced was that participants would exhibit greater egocentric responding when interacting with what they believed to be an authority, a partner who presumably could take responsibility in working toward mutual understanding. However, we also acknowledged that in the authority-status condition there might be an opposing force balancing out the processing task. Participants are also likely to assume that they should take on an authority figure's perspective, which as been discussed by previous research (as reviewed in Schober, 1993). For those in the same-status condition, there is no similar opposition at play. And given that there was no "correct" perspective and that the egocentric was easiest, perspective choice should be weighted toward egocentric responding. These hypotheses were somewhat supported by the data, but were weakened by the marginally significant tendency for those in the authority-status condition to prefer egocentric responses. Future work will explore the social modulation in greater detail, using partners who are more "authoritative" (thus increasing egocentric responding), or partners who is more subordinate (thus increasing other-centric behavior).

For our second set of hypotheses, the results were more conclusive. Here, our theoretical claim was that some participants will orient toward an other-centric response mode, and when they do, response times will be facilitated when interacting with a authority-status partner compared to a same-status partner. Indeed, this was found, and was expected as subordinates have been shown to be highly motivated to respond to a superiors' perceived intentions. Indeed, this could not have been merely *overall* motivation to satisfy the authority's needs, because the same was not true of those who responded egocentrically: The conditions did not differ. Thus, the speed with which spatial cognition was operating seems to have more rapidly adapted to the partner with an authoritative social status.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *The Psychology of Learning and Motivation: Advances in Research and Theory*, *53*, 301-344.

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science*, *1*, 274-291.

Carlson, L. A. (1999). Selecting a reference frame. *Spatial Cognition and Computation*, *1*, 365-379.

Clark, H. H. (1996). Using language. *Computational Linguistics*, *23*.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1-39.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, *62*, 378-391.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*, 35-51.

Horton, W. S., & Spieler, D. H. (2007). Age-related differences in communication and audience design. *Psychology and Aging*, *22*, 281.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676-703.

Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A., Ruiter, J. P., Hagoort, P., et al. (2009). Recipient design in tacit communication. *Cognition*, *111*, 46-54.

Raczaszek-Leonardi, J. R., Shapiro, L. P., Tuller, B., & Kelso, J. A. S. (2008). Activating basic category exemplars in sentence contexts: A dynamical account. *Journal of Psycholinguistic Research*, *37*, 87-113.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Rooij, I. van, Bongers, R. M., & Haselager, W. (2002). A nonrepresentational approach to imagined action. *Cognitive Science*, *26*, 345-375.

Rutherford, M. D. (2004). The effect of social role on theory of mind reasoning. *British Journal of Psychology*, *95*, 91-103.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*, 1-24.

Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations? *Discourse Processes*, *20*, 219-247.

Senay, I., & Keysar, B. (2009). Keeping track of speaker's perspective: The role of social identity. *Discourse Processes*, *46*, 401-425.

Snodgrass, S. E. (1992). Further effects of role versus gender on interpersonal sensitivity. *Journal of Personality and Social Psychology*, *62*, 154-158.

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, *62*, 207-218.

Tuller, B., Case, P., Ding, M., & Kelso, J. A. (1994). The nonlinear dynamics of speech categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 3-16.

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking. *Cognition*, *110*, 124-129.

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, *4*, 96-118.