Intention-based Robot Control in Social Games

Christopher Crick (christopher.crick@yale.edu)
Yale Department of Computer Science, 51 Prospect St.
New Haven, CT 06511 USA

Brian Scassellati (scaz@cs.yale.edu)
Yale Department of Computer Science, 51 Prospect St.
New Haven, CT 06511 USA

Abstract

We present a novel, sophisticated intention-based control system for a mobile robot built from an extremely inexpensive webcam and radio-controlled toy vehicle. The system visually observes humans participating in various playground games and infers their goals and intentions through analyzing their spatiotemporal activity in relation to itself and each other, and then builds a coherent narrative out of the succession of these intentional states. Starting from zero information about the room, the rules of the games, or even which vehicle it controls, it learns rich relationships between players, their goals and intentions, probing uncertain situations with its own behavior. The robot is able to watch people playing various playground games, learn the roles and rules that apply to specific games, and participate in the play. The narratives it constructs capture essential information about the observed social roles and types of activity. After watching play for a short while, the system is able to participate appropriately in the games. We demonstrate how the system acts appropriately in scenarios such as chasing, follow-the-leader, and variants of tag.

Keywords: Artificial Intelligence, Interactive Behavior, Learning, Social Cognition, Robotics

Introduction

Humans have a powerful ability to make sense of the world using very rudimentary sensory cues. We can watch children from down the street, and know instantly whether they're playing amicably or if we need to prepare to deal with torn jeans and tears. We can sit in the nosebleed bleachers and enjoy a football game, even though the players are nothing more than small colored blobs. We can navigate the house by a four-watt nightlight and (usually) pilot automobiles through traffic in the dark and the fog. We usually can make do with even less. Two-thirds of a century ago, Heider and Simmel found that animated boxes on a flat white screen are enough to trigger this inference process (Heider & Simmel, 1944). We easily spin stories about sterile geometric shapes, assigning them intentions, personalities and goals. Given the chance, we happily take control of these nondescript avatars to play out our own intentions and desires, whether in the context of psychological research (Gigerenzer & Todd, 1999), or simply in relaxing video games.

Making sense of very low-context motion data is an important cognitive task that we perform every day, an irrepressible instinct that develops quickly in children, around the age of 9 months (Rochat, Striano, & Morgan, 2004). This low-level processing skill is quickly followed by the development of other social skills (Csibra, Gergely, Biro, Koos, & Brockbank, 1999), such as the attribution of agency and intentionality. It depends on very little information from the world

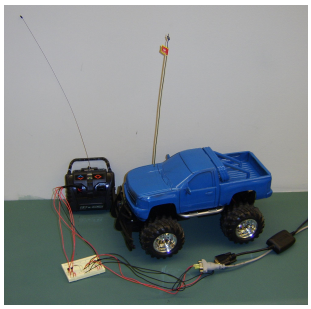

Figure 1: The robot-controlled toy truck

– so little, in fact, that we can have some hope at designing computational processes that can manipulate the manageable quantity of data to accomplish similar results. What's more, this can be accomplished quickly enough to serve as a control system for a robot, enabling us to explore the relationship between *watching* a game and *participating*. When taking an active part, the system can probe uncertainties in its learning, collapsing ambiguity by performing experiments, and explore how motor control relates to social interaction (Wolpert, Doya, & Kawato, 2003).

Our work also draws from and contributes to investigations of the fundamental cognitive processing modules underpinning perception and interpretation of motion. These modules appear responsible for our rapid and irresistable computation of physics-based causality (Choi & Scholl, 2006), as well as facile, subconscious individuation of objects in motion independently of any association with specific contextual features (Leslie, Xu, Tremoulet, & Scholl, 1998) (Scholl, 2004) (Mitroff & Scholl, 2004). Furthermore, different processing modules appear to attend to different levels of detail in a scene, including global, low-context motion such as used by our system (Loucks & Baldwin, 2008).

The specific analysis undertaken by our system, hypothesizing vectors of attraction and repulsion between agents and objects in the world in order to explain the causal relationships we note in an interaction, relates to the dynamics-based model of causal representation proposed by Wolff (Wolff, 2007) and on Talmy's theory of force dynamics (Talmy, 1988). As Talmy notes, the application of force has a great impact (no pun intended) on our understanding of the semantics of interaction, and on our ideas about causality, intention and influence. Humans can explain many events and interac-

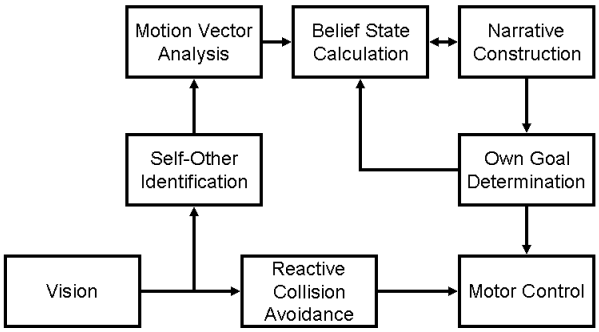

Figure 2: System components and information flow

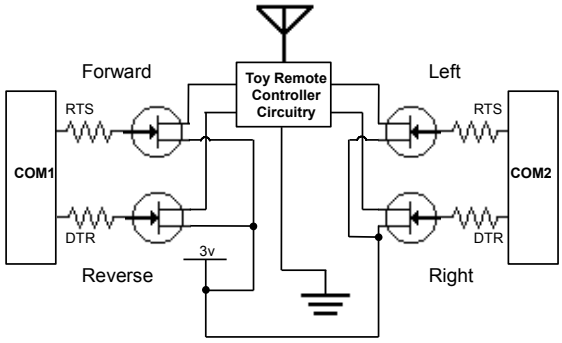

Figure 3: Circuit diagram for computer-controlled toy RC car

tions by invoking a folk-physics notion of force vectors acting upon objects and agents. This holds not only for obviously physical systems (we talk easily of how wind direction affects the motion of a sailboat), but for social interactions as well (the presence of a policeman can be interpreted – and in fact is described by the same vocabulary – as a force opposing our desire to jaywalk). Our system explicitly generates these systems of forces in order to make sense of the events it witnesses.

This work represents the latest step in our efforts to model a computationally tractable piece of human social cognition and decisionmaking. Within the constraints of its conceptual framework, our robot comprises a complete functional entity, from perception to learning to social interaction to mobility. Earlier versions of this system – lacking the ability to participate bodily in the observed games – are fully described in (Crick, Doniec, & Scassellati, 2007) and (Crick & Scassellati, 2008).

System Description

The system involves a number of interconnected pieces, depicted in Figure 2. Each component is described below in turn.

Vision

The system employs a simple but robust method to tracking the players as they move through the play space. Using an inexpensive USB webcam mounted on a stand in such a way as to provide a complete image of the floor of the room, the system uses naive background subtraction and color matching to track the brightly-colored vehicles. Before play begins, the camera captures a 640x480 pixel array of the unoccupied room for reference. During a game, 15 times a second, the system examines the raster of RGB values from the webcam and looks for the maximum red, green and blue values that differ substantially from the background and from the other two color channels of the same pixel. These maximum color values are taken to be the positions within the visual field of the three vehicles – one painted red, one blue, and one green (by design). Obviously, this is not a general-purpose or sophisticated visual tracking algorithm, but it is sufficient to generate the low-context percepts that are all our cognitive model requires.

Note that the camera is *not* overhead. The information coming to the robot is a trapezoid with perspective foreshortening. It would be possible to perform a matrix transformation to convert pixel positions to Cartesian geospatial ones, but our system does not go to the computational expense of doing so. The image may be distorted, but only in a linear way, and the vector calculations described below work the same, whether in a perspective frame or not.

Motor Control

In order not only to observe but to participate in activities, we provided our system with a robotic avatar in the form of a $20 toy remote-controlled car. By opening up the plastic radio controller and wiring in transistors to replace the physical rocker switches that control the car's driving and steering, and connecting these wires to controllable voltage pins on a computer's serial ports, we turned the system into a high-speed (7 m/s) robot. See Figure 3 for wiring details.

The controller is quick and reactive. The system maintains the position history over the previous $\frac{1}{5}$ second – three position reports, including the current one. With this information, it computes an average velocity vector and compares it with the intended vector given by the own-goal system described further below. Depending on the current direction of drive and the angle of difference between the actual and intended vectors, a set of commands is sent to the robot as shown in Figure 4.

Reactive collision avoidance

The room's walls obviously have an effect on the motions of the players, since their actions are constrained by the physical dimensions of the space. We chose to deal with wall avoidance in simple fashion. If the robot approaches too near the edge of the play area, a reactive behavior emerges that is independent of the goal state: if the robot is located within a certain number of pixels of the edge of the play area, an emergency goal vector pointing straight out from the wall or corner supercedes whatever the robot had been trying to do beforehand. This danger area ranged from 30 pixels wide at

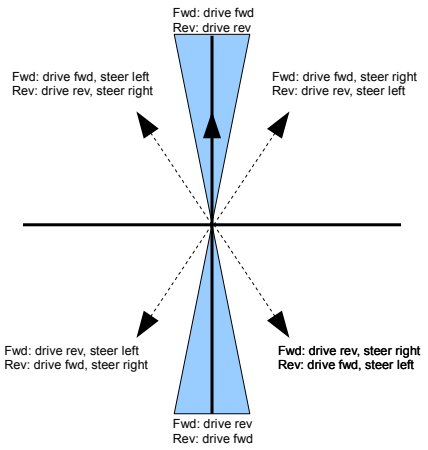

Figure 4: Robot directional control. The goal vector is compared to the computed vector of motion.

the bottom of the image (closest to the camera) to 18 near the top. Interestingly, several study participants noted the robot's ability to avoid running into walls, claiming that the robot was a much better driver than they were!

Self-other Identification

The system does not immediately know what salient object in its visual field "belongs" to itself. The playing area contains three different-colored toy cars, but it controls only one. Using a technique described in (Gold & Scassellati, 2005) for robotic self-recognition, the system sends out a few random motor commands and detects which of the perceived objects responds in a correlated fashion. The system sends a brief pulse (200 ms) of the command for "forward", followed by a similar command for "back", repeating as necessary. At the same time, the system inspects the visual field for the positions of the three salient colorful objects, looking for one moving predictably forward and back in time with the commands (finding and computing the necessary motion vectors are a byproduct of the analysis described in the next section). In this way, the system identifies itself for the duration of the exercise. Although the process would theoretically continue for as long as necessary, we found that throughout our experiments it never took more than one forward and reverse command for reliable identification.

Notably, this is precisely the same procedure invariably used by the human participants, who were each handed a remote controller without being told which of the three cars they were to drive. Invariably, the participant worked the controls forward and backward, watching the playing area to note which car acted as directed. The system has access to no privileged information about what it sees, no more than an undergraduate test subject walking into the lab space for the first time.

Motion Vector Analysis

Having determined which vehicle it is driving, the system begins to observe the behavior of the others to begin working out the rules of the game. For each of the other two participants in the game, the system calculates the "influence" of the remaining players (including itself) on the first person's perceived two-dimensional motion, expressed as constants in a pair of differential equations:

$$V_{x_i^n} = \frac{c_{x_j}(x_j^n - x_i^n)}{d_{ij}^n} + \frac{c_{x_k}(x_k^n - x_i^n)}{d_{ik}^n} + \cdots \quad (1)$$

(and similarly for the y dimension). It obtains the (noisy) velocities in the x and y direction, as well as the positions of the other vehicles, directly from the visual data:

$$V_{x_i^n} = \frac{x_i^{n+1} - x_i^n}{t_{n+1} - t_n} \quad (2)$$

(again, also for the y dimension). Here, $V_{x_i^n}$ represents the x component of agent i's velocity at time n. x_i^n, x_j^n and x_k^n are the x coordinates of agents i, j and k respectively, at time n. Likewise, d_{ij}^n and d_{ik}^n are the Euclidean distances between i and j or i and k at time n.

This results in an underconstrained set of equations; thus to solve for the constants we collect all of the data points falling within a short window of time and find a least-squares best fit. The visual system runs at 15 Hz; we found that a window of 220 milliseconds (about three position reports) worked best – coincidentally near the accepted average human reaction time (Laming, 1968).

Belief State Calculation

Each constant determined by the process described above represents in some fashion the influence of one particular player on the motion of another at a particular point in time. Some of these may be spurious relationships, while others capture something essential about the motivations and intentions of the agents involved.

To determine the long-term relationships that do represent essential motivational information, we next assemble these basic building blocks – the time-stamped pairwise constants that describe instantaneous attraction and repulsion between each agent and object in the room – into a probabilistic finite state automaton, each state representing a set of intentions that extend over time. At any particular point in time, any particular agent may be attracted or repelled or remain neutral with respect to each other object and agent in the room; this is characterized by the pairwise constants found in the previous step. The system assumes that the actors in the room remain in a particular intentional *state* as long as the pattern of hypothesized attractions, repulsions and neutralities remains constant, discounting noise. A particular state, then, might be that Red is attracted by Blue and neutral toward Green, Blue is repelled by Red and neutral toward Green, and Green is repelled by red and neutral toward Blue. This state might

occur, for instance, in the game of tag when Red is "it" and has decided to chase Blue.

The system maintains an evolving set of beliefs about the intentions of the people it observes, modeled as a probability distribution over all of these possible states. As new data comes in, the current belief distribution is adjusted, and the system assumes that the most likely alternative reflects the current state of the game.

$$Bel_n(S) = \frac{Bel_{n-1}(S)(1+\lambda\sum_{c\in S} s(c_n))}{Z} \quad (3)$$

Here, the belief in any particular state S at time n is the belief in that state at time $n-1$, modified by the current observation. c_n is the value at time n of one of the pairwise relationship constants derived from the data in the previous step; the function s is a sign function that returns 1 if the constant's sign and the intention represented by the current state agree, -1 if they disagree, and 0 if the state is neutral toward the pairwise relationship represented by the constant. λ is a "learning rate" constant which affects the tradeoff between the system's sensitivity to error and its decision-making speed. The magnitude of this factor ranges between 0.04 and 0.12, depending on whether the system is simply observing or is actively participating and trying out hypotheses (see the following section). Finally, Z is a normalizing constant obtained by summing the updated belief values across all states.

Own Goal Determination

As the system begins to observe its human partners, it develops a belief distribution over their possible intentional states. Because it controls a robot of its own, the system is then able to *probe* the likeliest candidate states. It chooses the belief state it has rated most likely, and acts in such a way to confirm or reject the hypothesis. It adjusts its beliefs accordingly, and more decisively than if it was not participating.

For example, say that the system had the highest degree of belief in the following state: Green was chasing Red and ignoring Blue, while Red was fleeing from both Green and Blue. To probe this state of affairs, the system would drive Blue toward Red. If Red continued to move away from Blue and Green did not react, the system's degree of belief in this state would further increase; if the other players reacted in some other way, the belief would subside, eventually to be replaced by another belief state judged more likely.

The ability to participate in and change the course of the game is a powerful tool for efficient learning. Machine learning theory is full of algorithms which perform much better when they are allowed to pose queries, rather than simply passively receiving examples (Angluin, 1988). Our system possess an analogous ability, able to query its environment and settling ambiguities in its beliefs by manipulating its own intentions and behaviors. At the same time, it watches for the effects on others' behaviors of the social forces brought into play by its actions. We show the effectiveness of such participation below.

Narrative Construction

The process described in the preceding sections converts instaneous velocity vectors derived from somewhat noisy video into sustained beliefs about the intentional situation that pertains during a particular phase of an interaction. As the action progresses, so too do the system's beliefs evolve, and as those beliefs change, the sequence of states becomes a narrative describing the scenario in progress. This narrative can be analyzed statistically to identify the action in progress, differentiate it from other possible activities, and also provide the system with clues to use in unsupervised feature detection. It can collect statistics about which states commonly follow which others (a prerequisite for developing the ability to recognize distinct activities). And it identifies points in time where important events take place, which will allow the system to notice information about the events themselves.

For this particular set of scenarios involving playground-like games, we set the system to look for game rules by observing the relative positions of the participants during the crucial moments of a belief state change, and to search for correlations between the observed distances and the particular state change. Distance is only one feature that could be considered, of course, but it is a common-enough criterion in the world of playground games to be a reasonable choice for the system to focus on. If the correlations it observes between a particular state transition and a set of relative distances are strong enough, it will preemptively adjust its own behavior according to the transition it has learned, thus playing the game and not only learning it.

Experiments

We tested the system in a 20x20-foot lab space with an open floor. We ran trials on three separate occasions, with two human subjects driving the red and green remote-controlled cars and the system controlling the blue one. We also ran one additional control trial with three human drivers and no robot-controlled car. The subjects themselves were in the room with the vehicles, but seated against the wall behind the camera's field of view. Each set of trials involved different people as drivers. Data from the first experiment were collected during each trial; the final experiment involving modified tag was conducted only during the last trial.

Chasing and Following

The first game we tested was simple. Each player had only one unchanging goal. The driver of the red car was asked to stay as far away from the others as possible, while the green car gave chase. In each trial, the behavior of the system was consistent. Within less than a second, the system determined the intentional states of Green and Red with respect to each other. It then proceeded to generate and test hypotheses regarding their intentions toward itself, by approaching each of the two cars. Within a few seconds more, it was able to determine that Red was fleeing from both, and Green was indifferent to Blue. Since the intentional state never changed, no positional information was ever recorded or analyzed.

Table 1: Results from Chase and Follow

	No. of Games	Average Time	σ
Chase	6	7.5 sec	1.41
Follow	4	33.5 sec	4.66
Chase (observe)	3	29.3 sec	10.69

The fact that the robot can *participate* in the game provides it with significant added power to probe the players' intentional states. For comparison, we also ran versions of the game that involved three human drivers, relegating the robot system to the role of passive observer. Still, the system applied the same algorithms to hypothesize the intentions of the players, and eventually converged on a stable, correct belief state. But it took much longer.

The second game, Follow the Leader, increased the complexity somewhat. The driver of the red car was instructed to drive wherever he or shey wished, and the green car was to follow, but remain a foot or two away – stopping when the red car stopped, reversing if it got too close. Success in this game came when the system understood this: it should approach the red car from across the room, but avoid it close in. In this game, the system was only successful in four runs of the game, out of six. In both of the other two trials, it formed the belief that the game was Chase, just as in the previous experiment, and never noticed the change from following to ignoring or avoiding.

Tag

Having confirmed that the system was able to understand and participate in simple games, we asked our subjects to play the somewhat more sophisticated game of tag. In previous research that involved the system merely watching people play, rather than attempting to participate, we enjoyed a great deal of success (Crick et al., 2007; Crick & Scassellati, 2008). However, several factors conspired against us. The RC cars are not nearly as agile as actual humans, and our subject drivers had significant difficulties controlling the vehicles well enough to conduct the game. In addition, one of the three participants in the game – the robot – had no idea how it should be played, and the two human players were unable to demonstrate gameplay adequately by themselves. We asked a pair of students unconnected to our tests to watch the videos of the tag attempts, and neither of them were able to identify the game being played, either.

Since freeform tag was too difficult for all involved, we developed rules for a tag-like game in order to test the system's ability to understand turn-taking and role shifts within the context of a game. In the modified game, only one person was supposed to move at a time. The player designated as "it" picked a victim, moved toward it, tagged and retreated. Then the new "it" repeated the process. Figure 5 depicts a set of stills from one of these modified tag games. A frame-by-frame description of the game is depicted in Table 2.

At each time point, the table includes a human-constructed

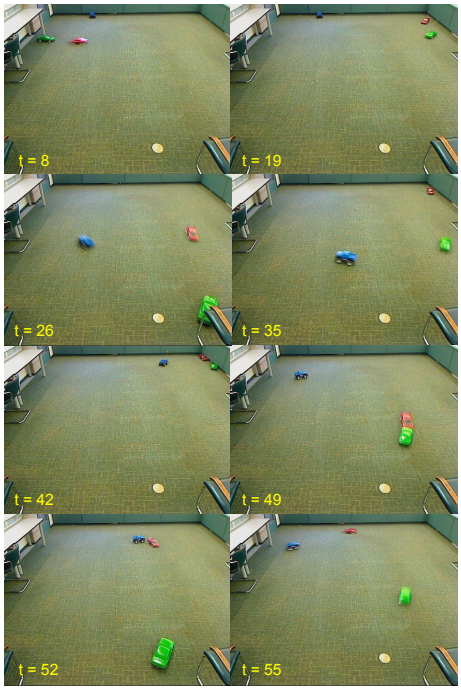

Figure 5: Succession of images from modified tag. See text and Table 2 for details.

verbal description of the action of the game, as well as the textual description produced by the system itself. This comes from the robot's own actions (which it knows absolutely and need form no beliefs about), and its belief in the intentional states of the players during a particular narrative episode. We can evaluate the system's success in ascribing intentions by comparing these human descriptions with the intentional states posited by the robot. Furthermore, we can identify points at which the system establishes rules that coincide with human understanding of the game. At the start, the robot watches the other two players each tag one another, without participating. Then, not knowing what its own role in the game is, it begins to move toward and away from the other players, observing their reactions. Because both of the human players are currently ignoring the robot, these actions are inconclusive. However, by second 42, the system has accumulated enough data to know that intentional shifts are signalled by close proximity. In the fifth frame, it recognizes the tag and reverses its own direction at the same time. By the seventh frame, it is testing to see whether approaching the red car will cause it to reverse course. By the end of this sequence, the system still has not determined that there is only one player ("it") with the chasing role, but it is well along the way – it understands tagging and the ebb and flow of pursuit and evasion.

Conclusion

Biological beings excel at making snap decisions and acting in a complex world using noisy sensors providing information both incomplete and incorrect. In order to survive, hu-

Table 2: Action and narrative during modified tag. Each time point corresponds to a still from Figure 5.

t (sec)	Action Description	System narrative
8	R approaches and tags G	R chases G.
19	R withdraws G approaches and tags R	G chases R, and R runs from G.
26	G withdraws R approaches and tags G B approaches G and R	R chases G, and G runs from R and me, and I chase G.
35	R withdraws G approaches R B approaches G and R	G chases R, and R runs from G, and I chase R.
42	G tags R B approaches G and R	G chases R, and I chase R.
49	G and B run from R R tags G	R chases G, and G runs from R, and I run from R.
52	R withdraws B approaches and tags R	R chases G, and G runs from R, and I run from R.
55	G approaches B and R B withdraws	R runs from me, and G runs from R, and I chase R.

mans must engage and profit from not only their physical environment, but a yet-more-complex social mileu erected on top. One of our most powerful and flexible cognitive tools for managing this is our irrepressible drive to tell stories to ourselves and to each other. This is true even or perhaps especially when we have only sparse information to go on. And beyond the telling, we take great delight in participating. We play games, we act, and the stories we love most are the ones in which we are the central characters.

We have developed a system that takes advantage of the very fact that it receives only rudimentary sensory impressions, and uses them to weave a story in which it can take part. The relative positions of moving objects are more than enough data for a human observer to begin making sense of the interaction by imagining their intentions and goals. By applying force dynamics to hypothesize about such human intentions, and by acting on those hypotheses to explore and verify its beliefs about the world, our system attempts to do the same.

The system has to figure out for itself how its motor controls correspond to action in the world. It theorizes about and tries to learn the intricate rules to games it knows nothing about. The verisimilitude of the data thus collected enables us to draw stronger conclusions with respect to real human interaction and interpretation, in contrast to data derived from simulation or computer-mediated play. And it does it in the real world, in real time, at human speed, using few shortcuts.

Acknowledgements

Support for this work was provided by National Science Foundation awards #0534610 (Quantitative Measures of Social Response in Autism), #0835767 (Understanding Regulation of Visual Attention in Autism through Computational and Robotic Modeling) and CAREER award #0238334 (Social Robots and Human Social Development). Some parts of the architecture used in this work was constructed under the DARPA Computer Science Futures II program. This research was supported in part by a software grant from QNX Software Systems Ltd, hardware grants by Ugobe Inc., and generous support from Microsoft and the Sloan Foundation.

References

Angluin, D. (1988). Queries and concept learning. *Machine Learning*, *2*, 319–342.

Choi, H., & Scholl, B. J. (2006). Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. *Perception*, *35*, 385–399.

Crick, C., Doniec, M., & Scassellati, B. (2007). Who is it? inferring role and intent from agent motion. In *Proceedings of the 11th ieee conference on development and learning.* London UK: IEEE Computational Intelligence Society.

Crick, C., & Scassellati, B. (2008). Inferring narrative and intention from playground games. In *Proceedings of the 12th ieee conference on development and learning.* Monterey CA: IEEE Computational Intelligence Society.

Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition*, *72*, 237–267.

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

Gold, K., & Scassellati, B. (2005). Learning about the self and others through contingency. In *Aaai spring symposium on developmental robotics.* Palo Alto, CA: AAAI.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*, 243-259.

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.

Leslie, A. M., Xu, F., Tremoulet, P. D., & Scholl, B. J. (1998). Indexing and the object concept: developing 'what' and 'where' systems. *Trends in Cognitive Sciences*, *2*(1), 10–18.

Loucks, J., & Baldwin, D. (2008). Sources of information in human action. In *Proceedings of the 30th annual conference of the cognitive science society* (pp. 121–126). Austin TX: Cognitive Science Society.

Mitroff, S. R., & Scholl, B. J. (2004). Forming and updating object representations without awareness: evidence from motion-induced blindness. *Vision Research*, *45*, 961–967.

Rochat, P., Striano, T., & Morgan, R. (2004). Who is doing what to whom? young infants' developing sense of social causality in animated displays. *Perception*, *33*, 355–369.

Scholl, B. J. (2004). Can infants' object concepts be trained? *Trends in Cognitive Sciences*, *8*(2), 49–51.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, *12*, 49-100.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology*, *136*, 82–111.

Wolpert, D. M., Doya, K., & Kawato, M. (2003, March). A unifying computational framework for motor control and social interaction. *Philisophical Transactions of the Royal Society B*, *358*(1431), 593–602.