

A visual model of collective behavior in crowds

by

Gregory C. Dachner

B.S., University of California, Merced, 2013

Sc.M., Brown University, 2015

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in the

Department of Cognitive, Linguistic, and Psychological Sciences at Brown University

Providence, Rhode Island

May 2021

© Copyright 2021 by Gregory C. Dachner

This dissertation by Gregory C. Dachner is accepted in its present form by the Department of Cognitive, Linguistic, and Psychological Sciences as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date \_\_\_\_\_

Dr. William H. Warren Jr., Advisor

Recommended to the Graduate Council

Date \_\_\_\_\_

Dr. Fulvio Domini, Reader

Date \_\_\_\_\_

Dr. Thomas Serre, Reader

Approved by the Graduate Council

Date \_\_\_\_\_

Dr. Andrew G. Campbell, Dean of the Graduate School

## Gregory Carter Dachner

### *Curriculum Vitae*

gregory\_dachner@brown.edu

### Education

- Brown University; Providence, RI  
Ph.D. in Cognitive Science  
Committee: William H. Warren Jr. (chair), Fulvio Domini, Thomas Serre
- 2015 Brown University; Providence, RI  
Sc.M. in Cognitive Science  
Thesis: Behavioral dynamics of heading alignment in pedestrian following
- 2013 University of California, Merced; Merced, CA  
B.S. in Cognitive Science  
*Highest Honors*
- 2010 Merced Community College; Merced, CA

### Publications

- Rio, K. W., **Dachner, G. C.**, Warren, W. H. (2018) Local interactions underlying collective motion in human crowds. *Proceedings of the Royal Society B: Biological Sciences*, 285(1878), 20180611.
- Dachner, G. C.**, & Kinateder, M. (2016). Effects of visual information on decision making during way-finding in emergency and non-emergency situations. *Transportation Research Procedia*, 185-189.
- Dachner, G. C.**, & Warren, W. H. (2014). Behavioral dynamics of heading alignment in pedestrian following. *Transportation Research Procedia*, 2, 69-76.

### Presentations

- Warren, W.H., **Dachner, G.C.**, & Wirth, T. (2020) A visual model of collective motion in human crowds. **Invited plenary lecture**, Workshop on Collective Information Processing, Humboldt University, Berlin.
- Warren, W. H., & **Dachner, G. C.** (2019). A Visual Model of Collective Motion in Human Crowds. **Talk**, *The European Conference on Visual Perception (ECVP)*; Leuven, Belgium.
- Dachner, G. C.**, & Warren, W. H. (2019). Dynamic occlusion reduces the influence of neighbors in human crowds. **Talk**, *International Conference of*

*Perception and Action (ICPA)*; Groningen, Netherlands.

- Warren, W. H., & **Dachner, G. C.** (2019). Multi-agent crowd dynamics: Visual coupling and network reconstruction. **Talk**, *International Conference of Perception and Action (ICPA)*; Groningen, Netherlands.
- Dachner, G. C.**, & Warren, W. H. (2019). Both optical expansion and depth information are used to control 2D pedestrian following. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Dachner, G. C.**, & Warren, W. H. (2018). A vision-based model of following in a human crowd. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Warren, W. H., **Dachner, G. C.** (2018). Comparing Disk and Doughnut Models of Collective Crowd Motion. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Wirth, T. D., **Dachner, G. C.**, Warren, W. H. (2018). Metric vs. Topological Models of Collective Motion in Human Crowds. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Warren, W. H., & **Dachner, G. C.** (2017). Simulating Collective Motion with a Pedestrian Model. Poster, *Sensori-motor control of animals and robots*; The Mathematical Biosciences Institute (MBI); Ohio State University; Columbus, OH.
- Dachner, G. C.**, & Warren, W. H. (2017). A vision-based model for the joint control of speed and heading in pedestrian following. Poster, *Pedestrian Dynamics: Modeling, Validation and Calibration*; The Institute for Computational and Experimental Research in Mathematics (ICERM); Brown University; Providence, RI.
- Dachner, G. C.**, & Warren, W. H. (2017). Visual control laws for pedestrian following. **Talk**, *International Conference of Perception and Action (ICPA)*; Seoul, South Korea.
- Dachner, G. C.**, & Warren, W. H. (2017). Vision-based model for the joint control of speed and heading in pedestrian following. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Warren, W. H., & **Dachner, G. C.** (2017). Simulating collective motion with a model of pedestrian following. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Dachner, G. C.**, & Kinateder M. (2016). Effects of visual information on decision making during way-finding in emergency and non-emergency situations. Poster, *Conference on Pedestrian and Evacuation Dynamics (PED)*; Hefei, China.
- Dachner, G. C.**, & Warren, W. H. (2016). Visual information for the joint control of speed and direction in pedestrian following. **Talk**, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.

Warren, W. H., & Rio, K. (2015). The visual coupling between neighbors in a virtual crowd. **Talk** presented in lieu of authors, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.

**Dachner, G. C., & Warren, W. H.** (2015). Behavioral dynamics of heading alignment in pedestrian following. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.

**Dachner, G. C., & Warren, W. H.** (2014). Behavioral dynamics of heading alignment in pedestrian following. **Talk**, *Conference on Pedestrian and Evacuation Dynamics (PED)*; Delft, Netherlands.

### **Teaching**

2018, Spring Instructor of Record; PSY 312, *Perception*; Wheaton College; Norton, MA.

2016, Fall Graduate TA; CLPS 0500, *Perception and Mind*; Brown University

2016, Spring Graduate TA; CLPS 0040, *Mind and Brain: Introduction to Cognitive Neuroscience*; Brown University

2015, Spring Graduate TA; CLPS 0510, *Perception, Illusion, and the Visual Arts*; Brown

2014, Fall Graduate TA; CLPS 0020, *Introduction to Cognitive Science*; Brown University

### **Selected Honors, Awards, & Funding**

2018 - 2019 Link Foundation – Modeling, Simulation, & Training Fellowship

2018 Hyundai Visionary Challenge: Pedestrians into Pilots

2018 Center for Vision Research (CVR) Pilot Research Grant: How visual information of goal angle and distance to the near end trade off during barrier avoidance

2017 - 2018 Brown / Wheaton Faculty Fellows Program

2016 - 2018 CVR Interdisciplinary Vision Training Program; National Eye Institute; Grant # 5T32EY018080-08, 2T32EY018080-11

2015 CVR Pilot Research Grant: Effects of visual information on decision making during way-finding in emergency and non-emergency situations

2013 Pamela Ann Stahl Prospective Teachers Scholarship

### **Symposia Organized & Hosted**

- 2016 Coffee Break Science: Communicating to Society; Brown University
- 2014 - 2015 Perception and Action Talk Series; Brown University

### **Academic Service**

- 2018 - 2020 Student and Employee Accessibility Services (SEAS), Accommodation Support; Brown University
- 2016 - 2018 Brown Graduate Resources Improving Professional Structures (B-GRIPS), Delegate; Brown University
- 2015 - 2018 Sheridan Center for Teaching and Learning, Department Representative; Brown
- 2015 - 2016 Swearer Center for Public Service, Social Innovation Fellowship, Project Analyst; Brown University
- 2014 - 2015 Department of Cognitive, Psychological, and Linguistic Sciences, Graduate Student Department Representative; Brown University

### **Employment**

- 2012 - 2013 Lab Manager, Dr. Michael Spivey's Lab University of California, Merced
- 2011 - 2013 Coordinator, Students Assessing Teaching and Learning (SATAL) University of California, Merced
- 2004 - 2007 Assistant Director, The Roleplay Tutoring Workshop Oakland, CA

### **Professional Development**

- 2017 Dealing with Conflict and Mediation Workshop, Ombuds Office; Brown
- 2016 Teaching Certificate IV, Teaching Consultant Program, Sheridan Center for Teaching and Learning; Brown University
- 2015 Teaching Certificate I, Reflective Teaching, Sheridan Center for Teaching and Learning; Brown University

## **Preface and Acknowledgements**

My grandpa Darrell once told me (well before I even had an inkling that graduate school even existed), “Greg, I got a Ph.D. To do it, I had to write a book about my research.” Here is my book. I hope you enjoy it.

I would like to thank everyone who got me here. Most importantly, thank you to my advisor, Bill Warren, for his continual guidance and support over these many, many years. From the first time I skyped with Bill (before even applying to Brown), I felt he was who I wanted to work with. And my intuition was right. I am grateful to have worked with such a dedicated and inspiring academic, who always had great advice and a kind ear. He has done the most of anyone to shape the academic I am today.

Thank you to my committee members, Fulvio Domini and Thomas Serre, who provided helpful guidance in the planning stages, useful feedback, and warm encouragement. I would also like to thank my undergraduate advisor, Michael Spivey; his influence drove me to majoring in Cognitive Science in the first place, which led me here. Also a big thanks to Theo Rhodes, whose lectures on the complexity of navigating static environments lead me to thinking about the complexity of navigating crowds. And thank you to Rolf Nelson, my faculty advisor from my wonderful time at Wheaton College; you reinforced and developed my love of teaching.

Importantly, huge thank you to all the members of Virtual Environment Navigation Laboratory that I have had the privilege of working with during my graduate career. I especially want to acknowledge Kevin Rio, who took time from writing his own dissertation to help me learn the VENLab ropes during my first year. Thanks, in addition, in the order of getting the pleasure to meet and work with them: Max Kinatered, Trent



Wirth, Arturo Cardenas, Brittany Baxter, Joey Bai, Meghan Willcoxon, and Brian Free. In addition, thank you to my cohort and the other students, researchers, and faculty in the Department of Cognitive, Linguistic, and Psychological Sciences at Brown University.

Next, a big thanks to the admin and staff of CLPS, the unsung heroes of all our graduate careers. We'll miss you, Michelle Ross! I would also like to thank my funding and support sources, who graciously buttressed the research presented here. These include the Brown Center for Vision Research, the Brown Center for Computation and Visualization, the Brown / Wheaton Faculty Fellows Program, the Link Foundation, the National Eye Institute, and the National Institutes of Health.

Thank you to all my family and friends, whose encouragement, patience, and support made this all possible. Graduate school is temporary. Family is forever.

Finally, I want to make a special note about this dissertation: the bulk of this text was written and edited while in self-isolation during the Covid-19 Pandemic of 2020. Some like to joke about how this gives more time to write, but it was ultimately a scary and lonely time. That said, none of this text would have been possible without Denise Werchan, my perennial best friend, roommate, home officemate, confidant, cheerleader, workout buddy, co-chef, editor, adviser, and soulmate. We leaned on each other through self-isolation together and none of this would have been possible without her.



Dedicated in memory of

Dr. Darrell B. Carter (1924-2020)

Emeritus, School of Optometry

University of California, Berkeley

Grandfather

## Table of Contents

<b>Tables .....</b>	<b>xiv</b>
<b>Figures.....</b>	<b>xv</b>
<b>CHAPTER 1 .....</b>	<b>1</b>
<b>CHAPTER 2 .....</b>	<b>13</b>
2.1 Introduction .....	14
2.2 The Visual Model .....	15
2.3 Experiment 1 - Leader / Follower Dyads .....	20
2.4 Experiment 1 - Results .....	23
2.5 Fitting the Model .....	24
2.6 Experiment 1 – Model Simulations .....	25
2.7 Experiment 2 - Following a Virtual Leader.....	29
2.8 Experiment 2 - Results .....	33
2.9 Experiment 2 – Model Simulations .....	39
2.10 Discussion.....	44
2.11 Conclusion.....	45
<b>CHAPTER 3 .....</b>	<b>46</b>
3.1 Introduction .....	47
3.2 Visual model for following a crowd.....	47
3.3 Experiment 3 - Effect of Distance in a Crowd .....	50
3.4 Experiment 3 - Results .....	52
3.5 Experiment 3 – Model Simulations.....	53
3.6 Experiment 4 - Effect of Distance to the Nearest Neighbor in a Crowd .....	57
3.7 Experiment 4 – Results.....	59

3.8 Experiment 4 – Model Simulations .....	61
3.9 Experiment 4 - Discussion.....	63
3.10 Conclusion.....	64
<b>CHAPTER 4.....</b>	<b>65</b>
4.1 Introduction .....	66
4.2 Experiment 5 - Effect of occlusion in the crowd.....	67
4.3 Experiment 5 - Results .....	70
4.4 Experiment 5 - Discussion.....	74
4.5 Modeling occlusion .....	75
4.6 Experiment 5 - Model Simulations.....	77
4.7 Re-simulating Experiments 3 and 4.....	79
4.8 Conclusion.....	84
<b>CHAPTER 5.....</b>	<b>85</b>
5.1 Introduction .....	86
5.2 Experiment 6 - Modeling human ‘swarms’ .....	86
5.3 Experiment 6 - Results .....	89
5.4 Multi-agent simulations.....	92
5.5 Multi-agent Simulations - Speed Results .....	94
5.6 Multi-agent Simulations - Heading Results.....	98
5.7 Multi-agent Simulations - Discussion .....	101
5.8 Conclusion.....	102
<b>CHAPTER 6.....</b>	<b>104</b>
6.1 Introduction .....	105
6.2 Experiment 7 - Depth information to following a single neighbor .....	105
6.3 Experiment 7 - Results .....	109

6.4 Experiment 7 - Discussion.....	112
6.5 Experiment 8 - Information Conflict in Crowds .....	113
6.6 Experiment 8 - Results .....	115
6.7 Experiment 8 - Discussion.....	118
6.8 Conclusion.....	119
<b>CHAPTER 7 .....</b>	<b>121</b>
7.1 Discussion.....	122
7.2 Theoretical Conclusions .....	124
7.3 Future Directions .....	127
<b>References .....</b>	<b>131</b>

## **Tables**

### **Chapter 2**

Table 2.1: Fits and errors for the visual model

## **Figures**

### **Chapter 2**

Figure 2.1: Information relationship between follower and neighbor

Figure 2.2: Changes in optical expansion and angular velocity by motion

Figure 2.3: Asymmetries in optical information

Figure 2.4: Representative following trials from Experiment 1

Figure 2.5: Representative side-by-side trials from Experiment 1

Figure 2.6: Computed speed and heading changes from Equations (3) & (4)

Figure 2.7: Modeled following trials from Experiment 1

Figure 2.8: Modeled side-by-side trials from Experiment 1

Figure 2.9: Diagram of Experiment 2

Figure 2.10: Representative trials from Experiment 2

Figure 2.11: Speed and heading responses to relative motion asymmetry

Figure 2.12: Speed and heading responses to distance asymmetry

Figure 2.13: Modeled trials from Experiment 2

Figure 2.14: Model means capturing the relative motion asymmetry

Figure 2.15: Model means capturing the distance asymmetry

Figure 2.16: Quantitative comparisons of the models

### **Chapter 3**

Figure 3.1: Mean final speed and heading from Experiment 3

Figure 3.2: Representative trials from Experiment 3

Figure 3.3: Mean model results for Experiment 3

Figure 3.4: Quantitative comparisons of the models

Figure 3.5: Diagram of Experiment 4

Figure 3.6: Mean final heading from Experiment 4

Figure 3.7: Representative trials from Experiment 4

Figure 3.8: Mean model results for Experiment 4

Figure 3.9: Quantitative comparisons for Experiment 4

## **Chapter 4**

Figure 4.1: Diagram of Experiment 5

Figure 4.2: Mean proportion of occlusion from Experiment 5

Figure 4.3: Representative trials from Experiment 5

Figure 4.4: Mean heading from Experiment 5

Figure 4.5: Modeling dynamic occlusion

Figure 4.6: Mean modeled heading from Experiment 5

Figure 4.7: Quantitative comparisons from Experiment 5

Figure 4.8: Re-simulated results from Experiment 3

Figure 4.9: Model comparisons for Experiment 3

Figure 4.10: Re-simulated mean results from Experiment 4

Figure 4.11: Model comparisons for Experiment 4

## **Chapter 5**

Figure 5.1: Representative segments from Experiment 6

Figure 5.2: Quantitative comparisons from Experiment 6

Figure 5.3: Representative trials from speed multi-agent simulations

Figure 5.4: Heatmaps of convergence of speed by varied initial conditions

Figure 5.5: Representative trials from heading multi-agent-simulations



Figure 5.6: Heatmaps of convergence of heading by varied initial conditions

## **Chapter 6**

Figure 6.1: Representative trials from Experiment 7

Figure 6.2: Mean participant speed comparing Experiments 2 and 7

Figure 6.3: Quantitative comparisons for Experiment 2 and 7

Figure 6.4: Representative trials from Experiment 8

Figure 6.5: Mean participant speed from Experiment 8

Figure 6.6: Quantitative comparisons for Experiment 8

## CHAPTER 1

### Introduction: Following a Crowd and Visual Information

Imagine the end of a football game. Everyone stands from their seats and attempts to exit the stadium. While it may seem each pedestrian in the crowd has individual locomotor behavior, their motion is influenced by the motion of their neighbors also trying to egress. If one attendee slows down relative to all the other people in the queue to exit, those attendees behind the first slow down as well. And so, all the pedestrians coordinate and walk together toward the exit. This global pattern of motion emerges from local interactions of that system's lower-level components (Couzin & Krause, 2003); this case, interactions between the individual pedestrians.

Humans navigate complex crowds like this on a daily basis. These can range from quick interactions between pedestrians on the sidewalk to coherent motion on a train station platform as a hundred people walk to board their train. While in a crowd, a pedestrian needs to generate paths and avoid collisions, all while moving with neighbors as a group to reach their destination. Human crowds exhibit complex dynamic behavior, but how collective motion emerges from the local interactions between individual pedestrians remains an open question. This question of interpersonal coordination is a perception and action problem that I am going to analyze and model.

Research on collective motion primarily began with a mathematical approach that sought to formalize rules for how an individual agent within the collective interacts with their neighbors. Very early work exploring these rules looked at how the collective behavior seen in fish schooling might emerge from simple rules between individual fish, in a position-based approach to measuring fish motion (Breder, 1954), as opposed to velocity-based approaches that rely on alignment of velocity vectors. This work demonstrated that macroscopic schooling behavior emerges from a bottom-up combination

of all the microscopic attraction and repulsion interactions between individual fish. Other work has shown how pursuit and escape interactions in insect swarms (analogous to attraction and repulsion between agents) also produced macroscopic collective motion (Romanczuk, et al. 2009). These findings showed that escape-inspired interactions led to a homogeneous distance between all agents within a simulation (as all agents would repel one another), while the pursuit-inspired interactions caused the formation of grouped clusters. Importantly, the combination of these interactions led to stable collective motion. However, both these works only examined how the cohesion was maintained given properties of attraction and repulsion, without explicit descriptions for alignment, a behavior clearly seen in the fish schools and bird flocks (Couzin et al., 2002).

Vicsek et al. (1995) took a different approach to simulations of collective behavior. Here the focus was on how alignment behaviors between individual agents (here termed “self-propelled particles”) in a collective led to self-ordered systems. Multiple simulations were performed where the initial particle positions were kept constant while the interaction parameters were varied. This was to find how the parameters lead to different patterns of motion. Each simulation varied particle direction of motion and distance (density) between particles. Speed was kept constant. Each particle matched the mean direction of all neighbors within a fixed radius on the next time step, and all particles were synchronously updated per time step. Simulations demonstrated that different forms of collective behavior emerged as density and direction of motion were varied. Low density (particles were spaced farther apart) and low variation in direction led to particles forming small groups of coherent motion, with each group having random mean direction; high density and low variation in direction led to ordered motion across all particles; and when both parameters

were high, it led to simulations with random motion and little coherence across all particles. These simulations demonstrated how varied parameters of initial conditions for bottom-up alignment interactions can lead to a wide range of macroscopic patterns of behaviors and are foundational for simulation work on collective motion. However, this work focused on only one potential alignment behavior between agents: alignment of direction, and neglected (for simplicity's sake) alignment of velocity.

Research since then has shown that agents in collective groups do align their speed with neighbors (Rio et al., 2014; Rio, Dachner, & Warren, 2018), and this alignment capability adds another degree of freedom to the macroscopic patterns of motion in a collective. In addition, this early form of self-propelled particles did not consider attraction and repulsion between agents, which is also a component of the interactions of individuals in collective behavior, and thus agents within this architecture are prone to collisions (although later accounted for, see Vicsek & Zafeiris, 2012).

Early computer animation work combined these four types of agent-based interactions (attraction, repulsion, speed alignment, heading alignment) into zonal models of collective behavior, as a means of avoiding the need to dictate explicit trajectories for a large number of animated agents (Reynolds, 1987). This work produced collective motion by modeling an agent's interactions with their neighbors as governed by three rules in concentric zones: repulsion, alignment, and attraction. In the first nearest zone, agents were repulsed by neighbors, preventing collisions. In the second zone, agents aligned their direction of motion with neighbors, producing collective motion. In the third and farthest zone, agents were attracted to their neighbors, ensuring the crowd would remain together.

These methods created qualitatively realistic looking bird flocks as a basis for computer animations of collective behavior in animals.

Later work used this zonal framework to model real biological systems that demonstrate collective behavior (Couzin et al., 2002). By varying the radii of the three zones, these agent-based models produced macroscopic behaviors that simulated state changes in fish schools, for example between torus-shaped schools and parallel motion in a specific direction. This work demonstrated how collective behavior likely emerges from individual interactions, not by direction from some form of hierarchical leadership. However, none of the discussed research attempted to apply these rules to pedestrian interactions.

The classic framework by which pedestrian agent-based interactions have been simulated is the Social Force Model (Helbing & Molnar, 1995). This framework was an attempt to simulate macroscopic collective motion in human crowds by simulating simple heuristic-based rules for individual pedestrian interactions. This modeling framework defined theoretical factors of social behavior as being analogous to physical forces acting on objects. For example, the social behavior of not standing near other people is analogous to the physical force of repulsion between two particles. The primary drive for these forces was the distance between simulated individual pedestrians, essentially following an attraction, repulsion structure. Simulations were able to demonstrate lane formation and congestion of a crowd into a narrow egress. Still, like the previously discussed distance-based models which only simulated attraction and repulsion between agents (Breder, 1954; Romanczuk & Couzin, 2009), the Social Force Model failed to account for the alignment of speed and heading behaviors between pedestrians. In addition, this model made no

considerations for how the information that controls these forces is perceived by the agents with the model, which likely led to unrealistic responses from agents. This leads to the most critical weakness of the early form of the Social Force Model: it was not based on real interactions between pedestrians in crowds. For example, while the model reduced the influence of neighbors behind a given agent, it was not zero. It is likely that neighbors outside the field of view of a given pedestrian do not influence that pedestrian's behavior - as any changes in behavior outside the field of view cannot be directly observed, only indirectly observed if crowd change information is transferred through mediating neighbors. Later versions of the model did incorporate a more theoretically accurate field of view (Moussaïd, et al., 2011). However, these additions do not make this model controlled by visual information. In addition, there is an even greater concern: the grounding of the Social Force Model. As the model was based on heuristic rules assumed by the authors and not grounded with empirical evidence, it cannot necessarily be generalized to real crowds.

The collective motion research presented thus far was conceptualized and investigated from an overhead third-person perspective and not in consideration for empirical evidence from crowds or the information used in collective behavior. Little research has explored the visual perspective of a pedestrian nested within a crowd. The body of research on crowd behavior to date has modeled how multiple agents interact with each other and whether collective behavior emerges from those interactions, but has largely failed to consider the sources of information that govern these interactions. While their primary focus was on exploring what rules generate global structure in crowds and flocks, they are an important groundwork to explore the complex behaviors observed in human

crowds. This modeling approach provides meaningful insights into the emergent behaviors of collective motion in crowds and how individuals within crowds may interact and coordinate with one another. However, pedestrians move through a crowd with a first-person perspective: they are embedded in and see the crowd, and use that information to govern their behavior. Thus, models with a consideration for the information used for collective behavior must be examined as well. This is especially true for visual information, as it is the primary sense a pedestrian would use in a crowd (Schellinck & White, 2011).

Recent work has proposed methods that simulate crowd behavior using input variables inspired by visual information (Ondřej et al., 2010; Moussaïd, et al., 2011; Dutra et al., 2017). The common thread for information input to these models was distance information between neighbors in crowds, i.e. the closest distance between the physical edges of two neighbors, or the time to closest contact. While these sources of distance information can be obtained visually, these are not true vision-based models for collective behavior. First, these models are not collective motion models in the classical sense. Under the constraints of the given models, collective motion is unlikely to actually occur in simulation, and usually only occurs when agents are already in near-alignment. This is even though collective motion in human crowds is ubiquitous (Warren, 2018). Second, they make assumptions about the information used to guide collective behavior. These models are more accurately described as distance-information models, not of visual models. Metric distance can be reconstructed from visual input, but it is possible that animals use other sources of visual information (that is informative about and correlated with distance) to guide behavior in collective motion scenarios, i.e. optical expansion, occlusion.



Consideration for the sources and uses of information in collective motion scenarios is vital for their modeling.

Other work has taken this ‘vision-inspired’ approach even farther. Bastien & Romanczuk (2020) derived a minimal generalized first-person perspective model for collective motion. This model used the binary presence or absence of neighbor edge and shape by location in the visual field as input, and no other visual properties (i.e. motion or occlusion). By shifting the response weights to neighbor presence by spatial location, multiple forms of collective motion were simulated, from non-convergent motion, to rotating disks, to swarms. The advantage of this model is it is generalizable, allowing it to be applied to collective motion seen in simple forms of animal life. However, while it is a theoretical approach motivated by visual systems, more grounding with empirical study would be necessary to demonstrate its applicability. In addition, it may be too simple regarding human pedestrians, as it does not use motion nor occlusion as forms of input.

Strandberg-Peshkin et al., (2013) is perhaps the closest current approach to a visual model for collective motion. Here, occlusion (calculated as the minimum threshold for detecting the subtended visual angle of a neighbor on the retina) and proportion of neighbors that previously responded was input to determine the likelihood of response by a simulated individual agent. Results were compared to experiments investigating information transfer in fish schools. Their findings show better performance with a visual model, compared to other collective motion models that relied on metric or topological neighborhood information. This implies that visual information and the proportions of neighbor behavior are combined in fish schooling and that these same concepts may be good candidates for modeling human locomotor behavior in crowds.

Unfortunately, the modeling approaches presented thus far have not provided a proper account of collective behavior human crowds. These models have typically contained two related issues: (1) the agents responded to their neighbors based on heuristics determined by the designers, because (2) the models lacked experimental evidence to ground the formulations of their microscopic interactions. The agents in these models often interact with neighbors in ways that would not be visible to a real pedestrian. While descriptively useful, this ignores the fact that in a crowd, a pedestrian can likely only respond to available visual information. Failing to incorporate visibility can create neighbor interactions that would otherwise be implausible. For example, an agent in a model may respond to a neighbor that is occluded by closer neighbors, or to a neighbor located outside that agent's field of view. Moreover, visual information is generated in accordance with the laws of optics, such as linear perspective, and may be provide indications for how distance information is derived and used. A computational dynamic model of crowd motion grounded in visual information could better describe the locomotor behaviors of an individual pedestrian, and thus be more accurate at predicting the interactions between all agents in a human crowd. Without an examination of the sensory information that guides local interactions between individuals, it is hard to gauge how well these models capture actual human crowd behavior. This is especially true when the models discount visual perception, which is likely to be the dominant source of information governing locomotor behaviors in human crowds (Bonneaud & Warren, 2014).

Previous research in visual perception has explored many visual mechanisms that may play a role in individual locomotion in a crowd, including motion processing, object tracking, distance perception, and higher-order optical variables (Burr & Thompson, 2011;

Chen, 2012; Geisler, 2008; Gibson, 1979). But these lines of research have primarily focused on simple visual stimuli, typically in situations where the participant is stationary. However, a human crowd is visually complex, involving many of the previously mentioned optical variables, in addition to the visual motion generated by the individual, who is also moving and embedded within the crowd itself. Both the crowd's motion and an individual's movement create a complex locomotor-control scenario, in which the individual uses continuously updated information about the immediately visible neighbors to move with a crowd.

The failure to consider visual information in most crowd models, and consequently the control laws that govern behavioral responses to that information, leads to multiple unknowns about the mechanics of human crowds. The purpose of this dissertation is to investigate and model the visual information that governs pedestrian following and to simulate collective motion using this information.

Previous work has found that people can generate collective motion by 'following' their neighbors; specifically, by aligning their locomotion vector with a weighted average of neighbor vectors in a neighborhood (Dachner & Warren, 2014; Rio, Dachner, & Warren, 2018; Warren, 2018). Other collective motion models have found that following a single neighbor through locomotion alignment was not required for coordination (Romanczuk & Couzin, 2009; Moussaid et al., 2011). However, these models need more empirical grounding, models of fish schools to perform better when alignment is considered (Strandberg-Peshkin et al., 2013), and pedestrians successfully walk together in groups, demonstrating cooperation in crowds (Rio, Dachner, & Warren, 2018).

The purpose of the VENLab's current research is to study if following behavior is sufficient to explain collective crowd motion, through the study of the dynamics of perception and action. This is accomplished by describing the interactions between an individual and their environment and then characterizing the evolution of this interaction over time (Fajen & Warren, 2003). This framework has successfully simulated the behavioral evolution of following a crowd. This has been accomplished through a bottom-up modeling approach of simulating a given agent in the crowd by sampling the means of their neighbors' speeds and heading (direction of motion) as input to compute how that simulated agent should change locomotion on the next time step. Comparisons to empirical and observational data have produced accurate results, suggesting that pedestrians can use neighbor speed and heading to follow a crowd (Rio, Dachner, & Warren, 2018; Warren & Dachner, 2017, 2018).

However, such crowd models are based on physical variables, not the visual information available to a pedestrian in a crowd. As human pedestrians will largely rely on visual information to navigate, it necessitates exploring how this information is perceived and then used for action. There are multiple visual variables humans could be using to locomote in crowds (e.g. edges and shapes; Bastian & Romanczuk, 2020). However, here we focus on three motion-based variables: the (1) optical expansion and (2) angular velocity (change in bearing direction) of neighbors as they change speed or turn relative the follower, and the loss of information about a neighbor due to (3) occlusion. These have been chosen as humans have been previously shown to be sensitive to these variables (Regan & Beverley, 1973; Regan & Beverley, 1979; Regan & Hamstra, 1993; Hoffmann, 1994; Wann, et al., 2011; Trent & Warren, 2017). Incorporating these optical variables into

a model of collective motion would significantly advance our understanding of the visual control of pedestrian behavior, strengthen crowd simulation for a variety of applications, and make models of collective motion more ecologically valid.

The remaining chapters of this dissertation are organized as follows: Chapter 2 develops, fits, and tests a model for following a single neighbor that relies on visual information. This work begins in this way so as to take a bottom-up approach, where once the interactions of a pedestrian and a single neighbor are understood, these rules can be generalized to multiple neighbors by the concept of superposition. Chapter 3 expands the model to include multiple neighbors (by the superposition of interactions) to account for collective motion and following a crowd. Chapter 4 experimentally tests whether neighbor occlusion influences following a crowd and then adds and tests this as a component to the model. Chapter 5 examines the model's performance using multi-agent simulations of large crowds with varied initial conditions to test the robustness of the model for collective motion and to find novel predictions. Chapter 6 tests other sources of visual information that a pedestrian may use in a crowd. Finally, Chapter 7 concludes the dissertation by discussing the overarching themes found with visual information and following a crowd and outlining future work.

## CHAPTER 2

### The Visual Model: Following a Leader

## 2.1 Introduction

In previous research on following in pairs of pedestrians (dyads), we have found that a pedestrian aligns their velocity (speed and heading direction) with a leading neighbor during following (Dachner & Warren, 2014; Rio, et al., 2014). When this coupling is replicated across multiple neighbors, the follower's trajectory in a crowd can be closely simulated (Rio, Dachner, & Warren, 2018). This is consistent with the principle of superposition, in which the response to a group is the linear combination of the response to each neighbor. In addition, these previous works found that an individual trajectory of a follower (with one neighbor or a crowd) can be simulated using models inspired by mass-spring equations. These are the Behavioral Dynamics Models for following a single or multiple neighbors.

A similar method can be employed to understand what visual information a pedestrian in a crowd uses to follow their neighbors. Chapter 2 will describe and model the visual control of following a single neighbor, so it can be replicated across all the neighbor interactions in a larger crowd, based on principles of superposition. I will begin by describing a behavioral dynamics model for following a single leader that uses optical variables as input, fit that model using real pedestrian dyads, and then test whether the model generalizes across a larger range of distances and eccentricities (the relative angle between a modeled agent's heading and a neighbor, where  $0^\circ$  is the follower's current heading direction) in a controlled virtual reality study.

## 2.2 The Visual Model

Two sources of visual information that human pedestrians are sensitive to and may use for following a single neighbor are optical expansion (change in visual angle over time, or  $\dot{\theta}$ ) and angular velocity (change in eccentricity [angle direction to a neighbor, or  $\beta$ ] over time, or  $\dot{\psi}$ ) of the followed neighbor (Regan & Hamstra, 1993). Through the application of Euclid's Laws of Visual Angles, equations can be derived that describe visual control laws for speed and heading to follow a moving neighbor or crowd of neighbors.

In order to understand how a follower uses these optical variables, it needs to be understood how these descriptive variables change as a leader moves relative to a follower. This change is dependent on how the leader's relative speed or heading offsets from zero (when both follower and leader have aligned locomotion), based on their eccentricity from the follower. To demonstrate this tradeoff in Figure 2.1, a follower is following a leader at various eccentricities. If the leader is at  $0^\circ$  eccentricity (directly in front of the follower's heading direction) and slows down, the leader optically expands, but their angular velocity is zero (Figure 2.1, panel a). However, if that same leader turns left, their optical expansion is now zero, but their angular velocity starts to increase (Figure 2.1, panel b).

This relationship is reversed if that leader is at  $+90^\circ$  eccentricity (directly to the right of the follower's heading). A leader turning left optically expands, while their angular velocity remains zero (Figure 2.1, panel c), whereas a leader that slows down increases angular velocity, while optical expansion is zero (Figure 2.1, panel d). At  $+45^\circ$  eccentricity, turning left or slowing down both create optical expansion and angular velocity (Figure 2.1, panels e & f). These optical variables tradeoff with eccentricity to specify whether a leader has changed relative speed, direction, or both, from zero. By nulling (i.e. canceling)



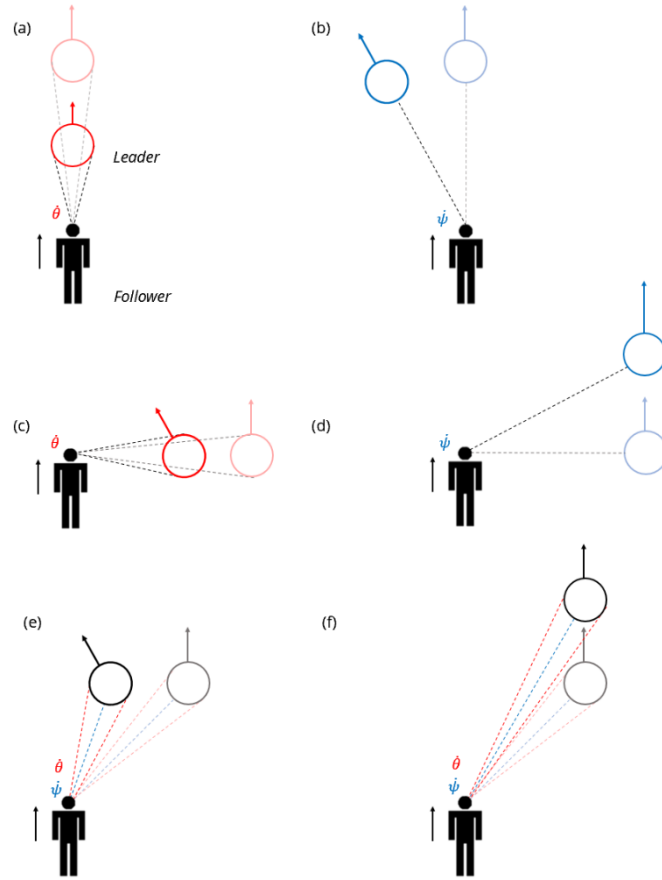


Figure 2.1. Information relationship between a follower and a leader. As relative speed or heading of the leader change, optical expansion ( $\dot{\theta}$ ) and angular velocity ( $\dot{\psi}$ ) change based on eccentricity ( $\beta$ ) from the follower.

changes in optical expansion and angular velocity, a pedestrian follower can approximately match speed and heading of a leader. This leads to the follower aligning their locomotion with a pedestrian neighbor using only optical variables.

Figure 2.2 demonstrates simple examples of how a leader's change in relative motion produce changes in these optical variables as observed by a follower. Here, a pedestrian is following a 0.4 m wide leader that is 1 m distant, both moving an initial speed of 1.0 m/s, and at an initial heading of  $0^\circ$  (the same direction). Panel (a) of Figure 2.2 shows potential positions of the leader relative to the follower, along a  $180^\circ$  eccentricity arc ( $-90^\circ$  to the left to  $+90^\circ$  to the right). Here the neighbor slows down relative to the follower by -

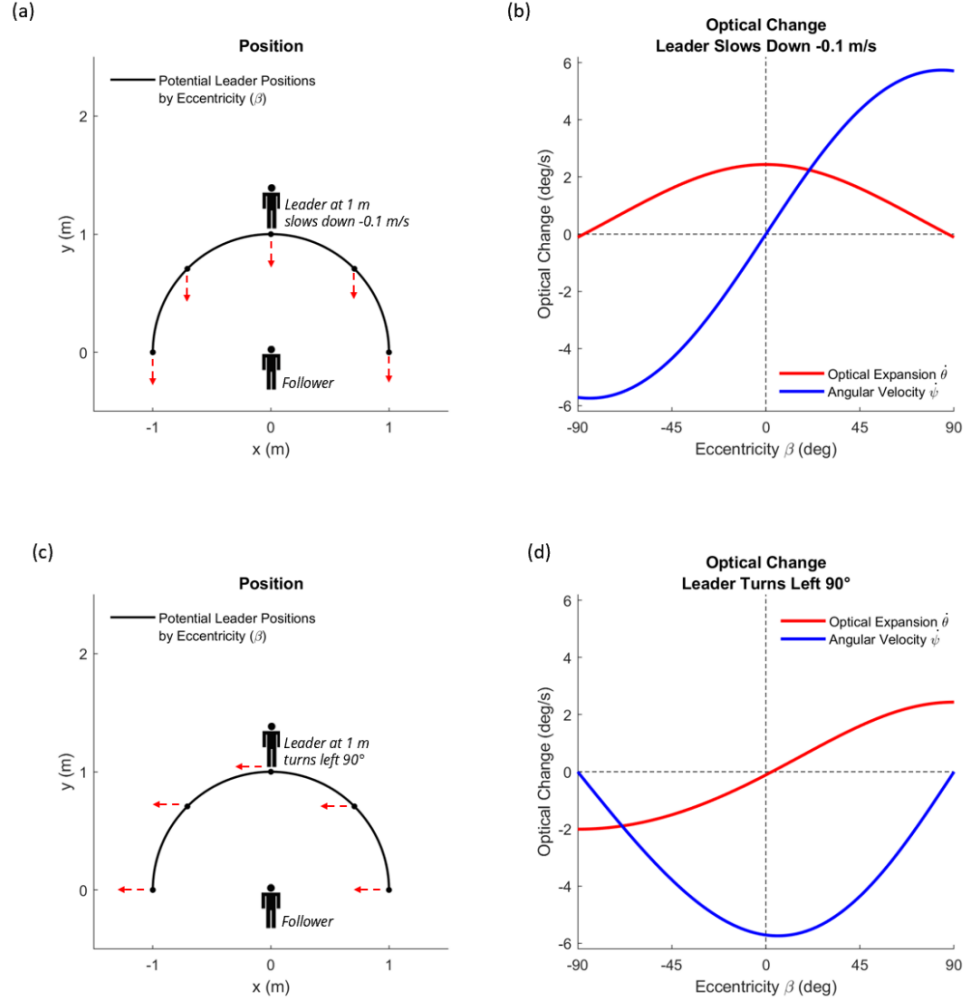


Figure 2.2. Changes in optical expansion and angular velocity of a 0.4 m wide leader that is 1 m distant across eccentricities ranging from  $-90^\circ$  to left to  $90^\circ$  to right.

0.1 m/s. Panel (b) of Figure 2.2 shows how optical expansion and angular velocity change by eccentricity as the follower observes that neighbor slowing down. Optical expansion is numerically computed by the change over time of visual angle:

$$\theta = 2 \cdot \tan^{-1}(c/2d) \quad (1)$$

and angular velocity is computed by the change over time of bearing:

$$\beta = \tan^{-1}(x_L - x_F, y_L - y_F) - \phi_F \quad (2)$$

where  $\theta$  is the visual angle the leader as observed by the follower,  $\beta$  is the eccentricity angle to the leader from the follower,  $c$  is the diameter of the leader (0.4 m),  $d$  is the

distance to the leader (1 m at timestep 1),  $(x_L, y_L)$  is the leader's position,  $(x_F, y_F)$  is the follower's position, and  $\phi_F$  is the follower's heading direction. Note there is zero optical expansion at  $\pm 90^\circ$  eccentricity, and zero angular velocity at  $0^\circ$  eccentricity. Panels (c) & (d) of Figure 2.2 show the optical computations of a neighbor turning left  $90^\circ$ , so that their speed 0.0 m/s in the 'x' dimension and 1.0 m/s in the 'y' dimension. The opposite result is found compared to panel (b) of Figure 2.2; now there is zero optical expansion at  $0^\circ$  eccentricity and zero angular velocity at  $\pm 90^\circ$  eccentricity. These two sources of optical information, across a range of eccentricities, can serve as control laws for following.

Based upon Euclid's Laws of Visual Angles, two asymmetries in the visual information available to a follower are evident. The asymmetry is that a follower will observe different magnitudes of optical change dependent on whether a leader's relative motion moves them closer or farther from the follower. In a simple example in Figure 2.3 panel (a), a leader (0.4 m wide, 1 m distant) speeds up by +0.1 m/s or slows down by -0.1 m/s. While the magnitude of relative speed change is identical, the magnitude of optical

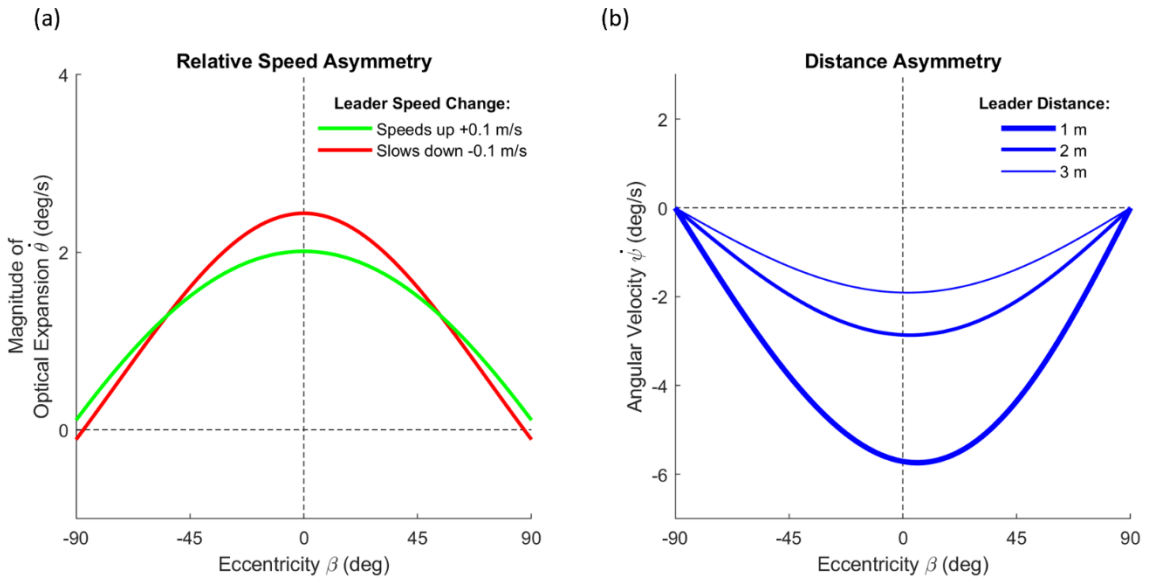


Figure 2.3. Asymmetries in optical information. Panel (a) demonstrates that identical magnitudes in leader speed change produce asymmetrical magnitudes of optical expansion. Panel (b) demonstrates an inverse relationship between angular velocity and distance.

expansion is not: a neighbor slowing down produces greater overall expansion than a neighbor speeding up (this is true for angular velocity as well). The second asymmetry is that a follower will observe different magnitudes of optical change dependent on a leader's distance. Figure 2.3 panel (b) shows a simple example where three leaders (each 0.4 m wide) turn left 90° at three distances: 1 m, 2 m, and 3 m. While the direction of turn is identical, there is an inverse relationship between distance and optical change: the shorter the distance, the greater the angular velocity (as is also true for optical expansion). These asymmetries in optical information make a novel predication: asymmetries in optical information may lead to asymmetries in human following behavior.

To build a model for following a single leader based on optical variables, these optical variables are transformed and scaled, so that two properties are derived: (1) depending on eccentricity, optical expansion and angular velocity trade off influence to determine the follower's change in speed, and conversely (2) angular velocity and optical expansion trade off as complementary functions of eccentricity to determine the follower's change in heading. These transformations lead to the following control laws for speed alignment:

$$\ddot{r} = -c_1 \cdot \sin(\beta) \cdot \dot{\psi} - c_2 \cdot \cos(\beta) \cdot \dot{\theta} \quad (3)$$

and for heading alignment:

$$\ddot{\phi} = -c_3 \cdot \cos(\beta) \cdot \dot{\psi} + c_4 \cdot \sin(\beta) \cdot \dot{\theta} \quad (4)$$

where  $\ddot{r}$  is the radial acceleration of the follower,  $\ddot{\phi}$  is the angular acceleration of the follower's heading direction,  $\beta$  is the eccentricity of the leader,  $\dot{\psi}$  is the angular velocity of the leader, and  $\dot{\theta}$  is the expansion rate of the leader.  $c_{1-4}$  are free parameters, fit to data from a leader/follower dyad experiment. To explain why each component of the equation

is multiplied by the sine or cosine of eccentricity, consider following scenario: there is a leader at  $-45^\circ$  eccentricity (to the left) or a leader a  $+45^\circ$  eccentricity (to the right). If each leader expands equivalently toward a follower, they specify identical amounts of expansion. However, the heading response of the follower will vary dependent on which side (eccentricity) of the follower the leader is on. When the leader is expanding on the left side, the follower should turn right in response. Conversely, when the leader is expanding on the right side, the follower should turn left in response. The scaled term of eccentricity produces the appropriate magnitude of response for a given leader eccentricity.

To fit the free parameters for Equations (3) and (3), Experiment 1 was conducted.

### **2.3 Experiment 1 - Leader / Follower Dyads**

Experiment 1 investigated the locomotion dynamics of following and was conducted as a part of a larger study. In this experiment, we recorded leader-follower dyads in which the leader was a confederate given pre-trial instructions to either change speed or direction.

#### **Participants**

12 participants were recruited at Brown University, 5 female and 7 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

## **Apparatus**

The experiment was conducted in the Sayles Hall space at Brown University. Sixteen infrared motion capture cameras (Qualisys, Deerfield, IL) were placed around a 14 x 20 m tracking area and used to record head position and orientation at 60 Hz. Each participant wore a lightweight bicycle helmet with five passive reflective markers on protruding stalks in a unique configuration, so each participant could be independently identified, and position and orientation could be derived.

## **Procedure**

In two separate blocks, one participant (designated the follower) was either directly behind or side-by-side another participant (designated the leader; a confederate). The follower was instructed to walk with the leader and stay with them if they changed speed or direction, while maintaining a constant distance. On each trial, the participants walked to starting marks on the floor that specified the initial interpersonal distance (1, 2, or 4 m). Before each trial, the leader received covert written instructions, directing them to make two changes during that given trial. These instructions contained directions to either change speed or direction (but not both). In the speed change trials, the leader received one of the four following instructions: (1) speed up and then speed up again, (2) speed up - slow down, (3) slow down - slow down, (4) slow down - speed up. In the direction change trials, the leader received one of the four following instructions: (1) turn right and then turn right again, (2) turn right - turn left, (3) turn left - turn left, (4) turn left - turn right. Blocks also contained two control trials in which the leader received covert written instructions to not change speed nor direction. The leader was trained and performed two practice trials on these instructions before the follower entered the hall and data was recorded.

To initiate the trial, the experimenter gave a verbal “begin” command to both participants, who started walking forward. The leader then made their instructed changes while walking about 20 m across the hall, varying the timing and magnitude of each change at will. The ten possible sequences and three initial distances were presented once in a random order, yielding 30 trials per block.

### **Design**

The experiment had a 2x3x10 factorial design: 2 walking scenarios (Following, Side-by-side), 3 initial distances (1, 2, 4 m), and ten instructed sequences. This was repeated across three sessions with different participants, for a total of 180 trials.

### **Data processing**

Due to tracking errors in the large hall, head positions and orientations were successfully recovered on 85% of the trials. There was a total of 153 usable trials, 72 trials in the Following scenario and 81 trials in the Side-by-Side scenario. The time series of head position in the horizontal plane were filtered using a 4th-order low-pass Butterworth filter with a cutoff frequency of 0.6 Hz. This was to remove tracker noise and reduce oscillations due to the step cycle. The first two seconds of every trial was removed, so that the participant was at a steady-state of locomotion for analysis and modeling. The last second of every trial was removed, to avoid any errors due to filtering at the edges of recorded data (see Howarth & Callaghan, 2009). Speed and heading of each participant were computed from the filtered position data as the displacement between successive time steps.

## 2.4 Experiment 1 - Results

Representative trials for the Following condition, in which the leader changed speed or heading, appear in Figure 2.4 and for the Side-by-side condition in Figure 2.5. These plots show the position, speed, and heading for both leader and follower. In all cases, the leader initiates the change in locomotion first and then the follower aligns their locomotion with the leader. The coupling strength was high, with a mean correlation between leader and follower for speed of Pearson's  $r = 0.87$  in speed changing trials and a mean correlation for heading of  $r = 0.85$  in heading changing trials. These experimental results provide clear evidence that leader and follower are strongly coupled, and thus are satisfactory data for fitting the free parameters in the visual model.

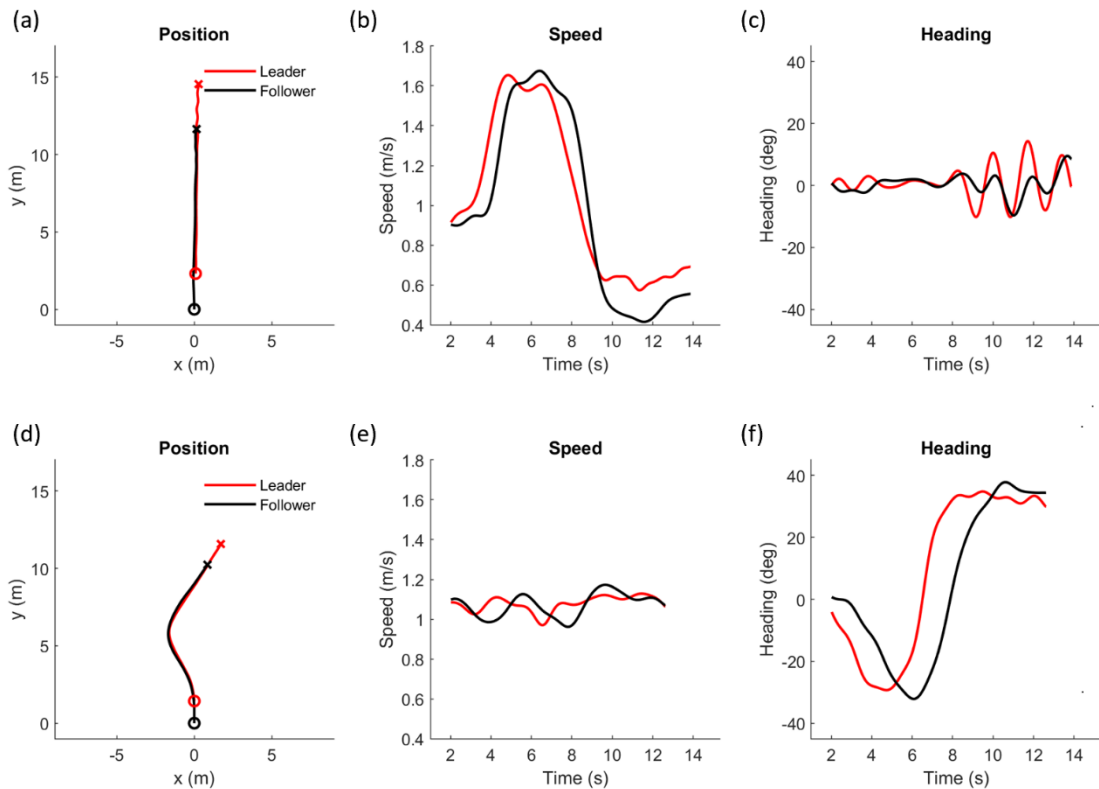


Figure 2.4. Representative trials from the Following condition in Experiment 1. Panels (a-c) show position (circle is the starting point, x is the end point of the trial), speed, and heading for a speed up - slow down 2 m initial distance condition. Panels (d-f) show position, speed, and heading for a right - left 1 m initial distance condition.



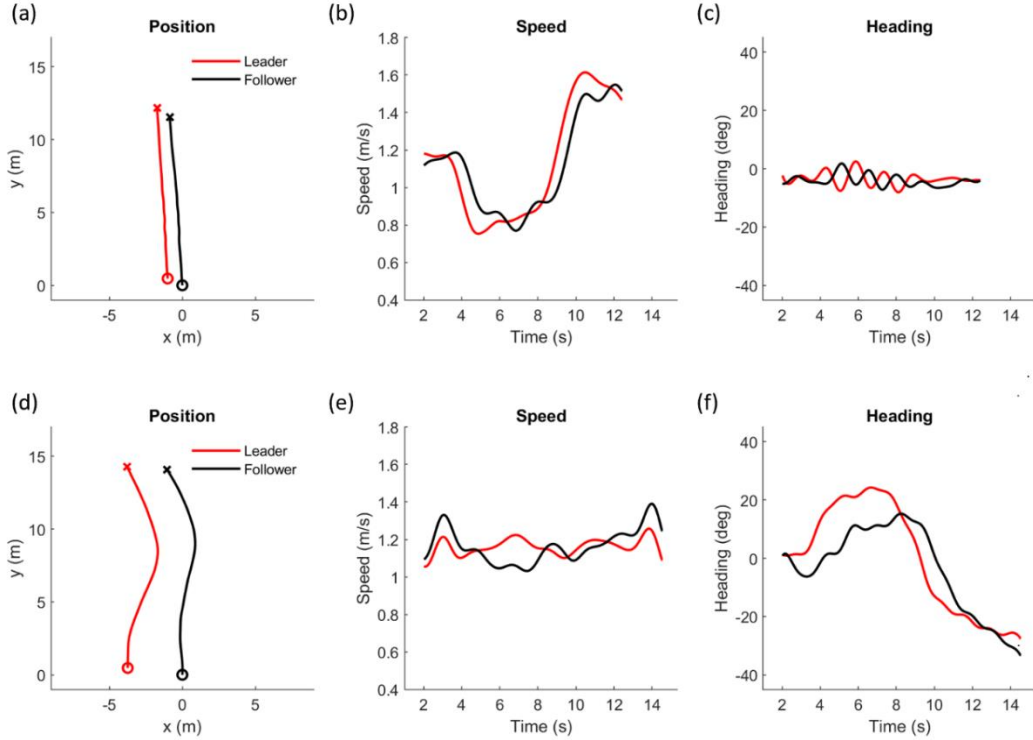


Figure 2.5. Representative trials from the Side-by-side condition in Experiment 1. Panels (a-c) show position (circle is the starting point, x is the end point of the trial), speed, and heading for a slow down – speed up 1 m initial distance condition. Panels (d-f) show position, speed, and heading for a left-right 4 m initial distance condition.

### 2.5 Fitting the Model

To find the best fitting parameters for the speed and heading equations of the visual model, the experimental trials were simulated, and the resulting time series compared to the human data. Each trial was simulated by taking the leader’s time series of position and the follower’s initial position, speed, and heading as inputs to Equations (3) and (4).

Model performance was evaluated on each trial by computing the root-mean-squared-error (RMSE) between the model and human time series for each equation independently. Performance for the two parameters for Equation (3) (radial acceleration) were evaluated by taking RMSE between the follower’s and model’s speed, in speed change trials only. Performance for the two parameters for Equation (4) (angular acceleration) were evaluated by taking RMSE between the follower’s and model’s heading,

in heading change trials only. Thus, the performance of each equation's two parameters was evaluated independently. This was due to heading not varying in the speed condition and vice versa. The parameters were fit using a cross-validation criteria. On each iteration, trials were randomly distributed into a training set (made up of 75% of trials) and a test set (made up of the remaining 25%). The Broyden-Fletcher-Goldfab-Shanno Method (Shanno, 1985) for numerical optimization was used to find the set of parameter values for a given training set that minimized the RMSE for each fit parameter. These training parameters were then used to simulate the trials in the test set and used to calculate the mean RMSE across the trials in the test set. A new iteration was then begun, randomly distributing trials between training and test sets, until 100 iterations were conducted. The final fixed parameters were calculated by taking the value that corresponded to the mean RMSE from the test sets across all iterations. The final parameter values and corresponding RMSE and mean correlations across all trials can be found in Table 2.1. The resulting parameter values were fixed for all subsequent simulations of the visual model in this dissertation, to avoid overfitting and to produce a predictive model that will generalize to new experimental conditions.

## **2.6 Experiment 1 – Model Simulations**

The mean correlation (Pearson's  $r$ ) between follower and model for speed and heading were strong (speed:  $r = 0.75$ , heading:  $r = 0.77$ ). Thus, the visual model closely reproduces the time series, and accounts for the coordination of speed and heading of following a single leader using only optical expansion, angular velocity, and eccentricity as inputs.

Table 2.1. Mean and standard deviation of RMSE, mean and standard deviation of correlation (Pearson's  $r$ ), and fit parameters from the optimization process using Experiment 1 as input, for Equations (3) and (4).

Equation	Mean RMSE	SD RMSE	Mean $r$	SD $r$	Angular Velocity Parameter	Expansion Parameter
Speed	0.167 m/s	0.152	0.75	0.55	$c_1 = 0.180$	$c_2 = 0.72$
Heading	5.972 deg	0.503	0.77	0.44	$c_3 = 14.38$	$c_4 = 59.71$

To demonstrate how these fit parameters tradeoff in their respective control laws, Figure 2.6 takes the numerically computed optical expansions and angular velocities from the simple examples in Figure 2.2 and computes the respective simulated speed and heading changes using Equations (3) & (4) and the fit parameters. The first row of panels in Figure 2.6 show the computed speed and heading changes from a leader slowing down

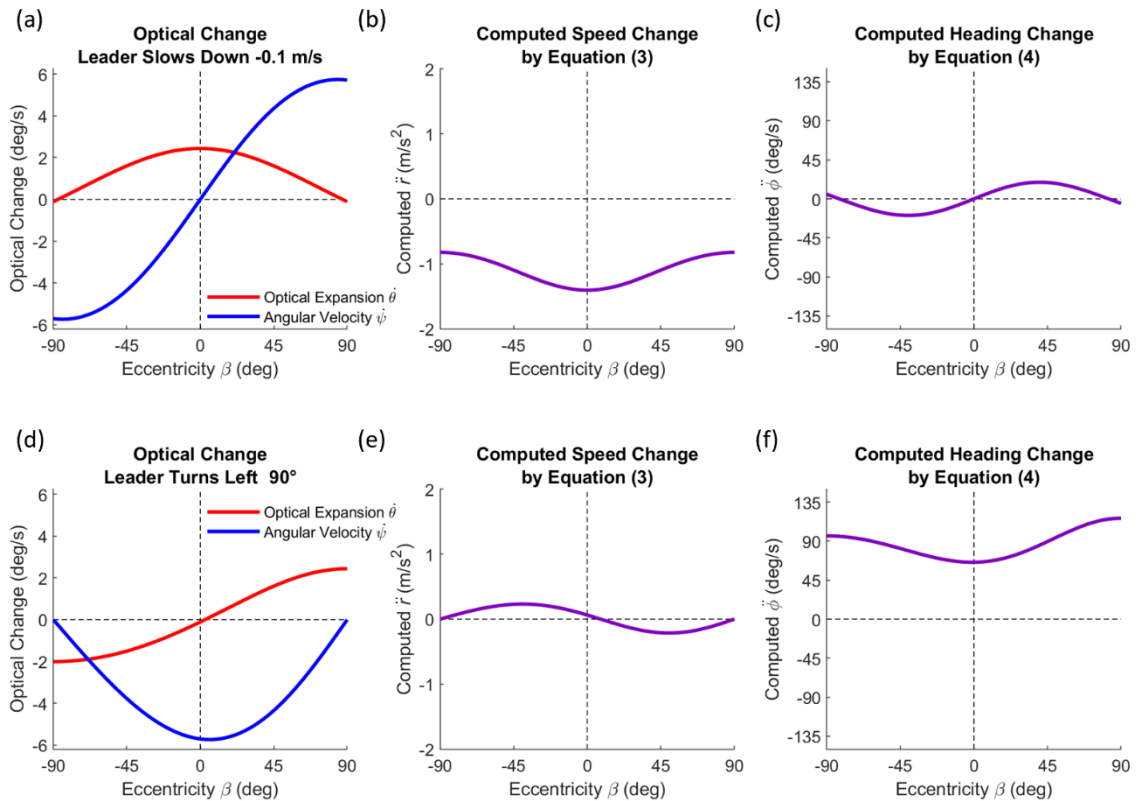


Figure 2.6. Computed speed and heading changes using Equation's (3) & (4), using the toy examples from Figure 2.2. Panel's (a-c) show a neighbor slowing down -0.1 m/s,

-0.1 m/s. Regardless of eccentricity, a leader slowing down causes the model to slow down (panel b), with little change in heading (panel c). The second row of Figure 2.6 shows the respective computations of speed and heading changes of a neighbor turning left by  $90^\circ$ . Here, regardless of eccentricity, a leader's left turn causes little change in speed (panel e) and a larger change in heading (panel f).

Figure 2.7 and Figure 2.8 show the representative trials from Experiment 1, together with simulations of the follower using Equations (3) and (4) and the fit parameters. These show the time series for position, speed, and heading for four trials. In each, the blue dashed curve is the trajectory of the model, which accurately simulates the follower's locomotor behavior.

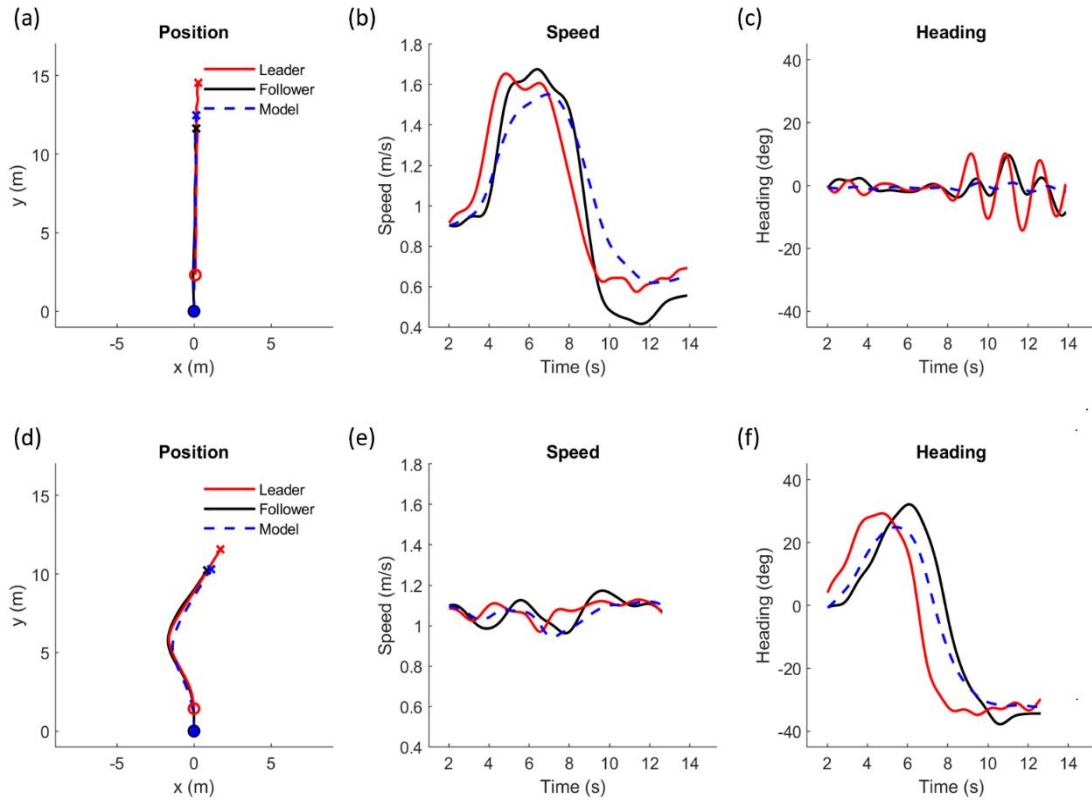


Figure 2.7. Representative simulations from the following condition of Experiment 1. The model (blue dashed curve) simulates follower trajectory.

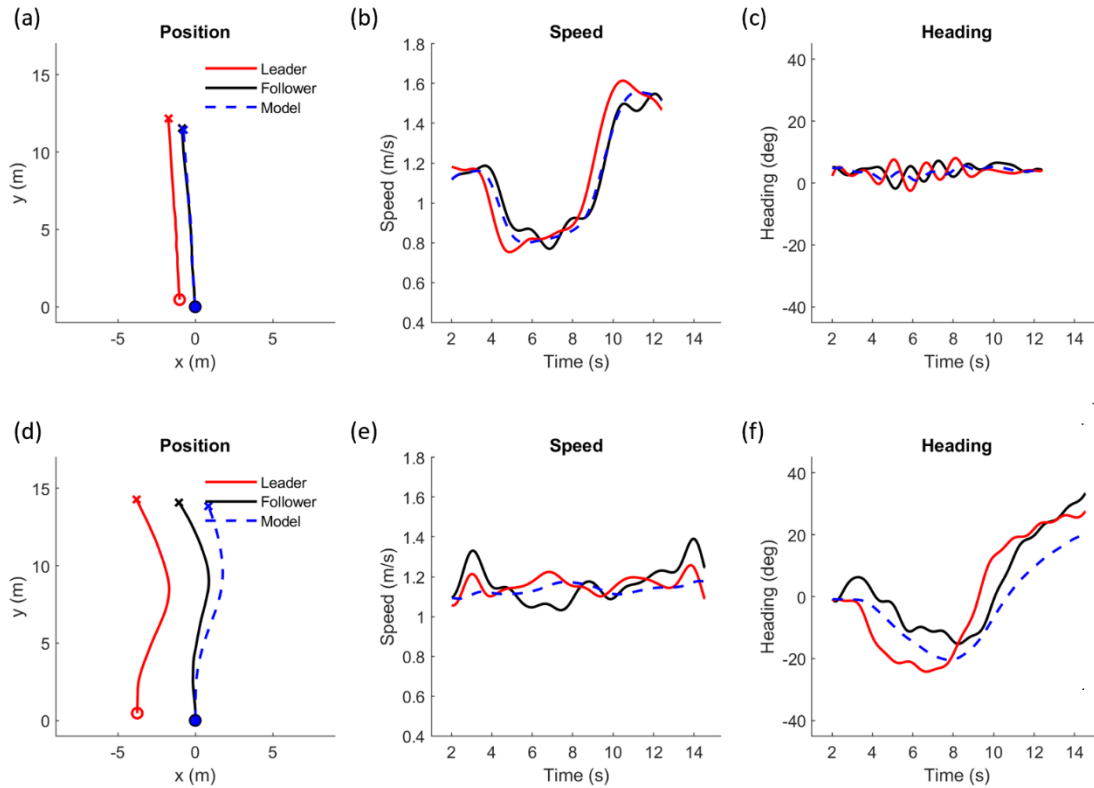


Figure 2.8. Representative simulations from the side-by-side condition of Experiment 1. The model (blue dashed curve) simulates follower trajectory.

However, Experiment 1 alone does not validate that model. While the model uses optical variables as input, the follower participants in Experiment 1 may have been using other sources of information to control locomotion and follow the leader. In addition, due to the constraints of the experiment, the behavior of the leader cannot be perfectly controlled and may vary in unknown ways. However, this is a benefit for parameter fitting, as it gives more variation to fit with. To demonstrate the generalizability of the visual model, Experiment 2 was conducted. This experiment brought following one neighbor into a virtual reality environment, ensuring that pedestrians can follow a controlled single target using only the 2D optical information of expansion and angular velocity.

## **2.7 Experiment 2 - Following a Virtual Leader**

Experiment 2 investigated the locomotion dynamics of following using virtual reality techniques. The purpose of the experiment was to test the visual model by manipulating the optical information presented to the participant during walking, to see whether the model generalizes to a wider range of conditions, and to see if the computed asymmetries in visual information produce asymmetries in human following. Specifically, to test the efficacy of the proposed optical variables, the participant followed a virtual target pole in empty space, while the optical expansion and/or the angular velocity of the target was perturbed, and the participant's walking trajectory was recorded. Based on Figure 2.1, we predict that if the target's optical expansion is perturbed when the target is at  $0^\circ$  eccentricity from the participant, the participant should change speed. While  $\pm 90^\circ$  eccentricity was not possible in this experiment (given the limited field of view of the head-mounted display), if the target's optical expansion is perturbed when the target is at  $\pm 60^\circ$  eccentricity from the participant, the participant should change speed and heading.

### **Participants**

12 participants were recruited at Brown University, 8 female and 4 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

### **Apparatus**

The experiments were conducted in the Virtual Environment Navigation Laboratory (VENLab) at Brown University. Participants walked in a 12 x 14 m tracking

area while wearing a stereoscopic head-mounted display (HMD, Oculus Rift DK1, 640 x 800 pixels per eye, 90° view, 60 Hz frame rate, 100% binocular overlap). Head position and orientation were recorded with an ultrasonic/inertial tracking system (Intersense IS-900, 60 Hz sampling rate) and used to update the display (50–67 ms latency).

### **Procedure**

Participants began each trial by facing toward an orientation pole floating in space. After 3 seconds, the orientation pole disappeared and the target to be followed appeared within their field of view, moving away from the participant. Participants were instructed to over headphones to “walk with the target as if walking down the street with it, while maintaining a consistent distance and orientation to the target”.

### **Displays**

The target was a virtual pole that appeared to be 0.4 m in width, 2 m in height (with approximately 1.6 m below eye level, dependent on participant eye height), floating in a black featureless space. The target appeared at five initial eccentricities relative to the participant’s initial facing direction: -60° and -30° (to left of the participant), 0° (directly in front of the participant), and +30° and +60° (to the right of the participant) and it could appear at 2 distances (1 m, 4 m). To remove distance information, there was no ground plane and the target was presented synoptically. Specifically, the calculated target position was at an infinite distance, and its subtended visual angle and visual direction were dynamically computed to simulate an object with a consistent size and eccentricity from the participant's viewpoint. Thus, the only information available to the participant was 2D

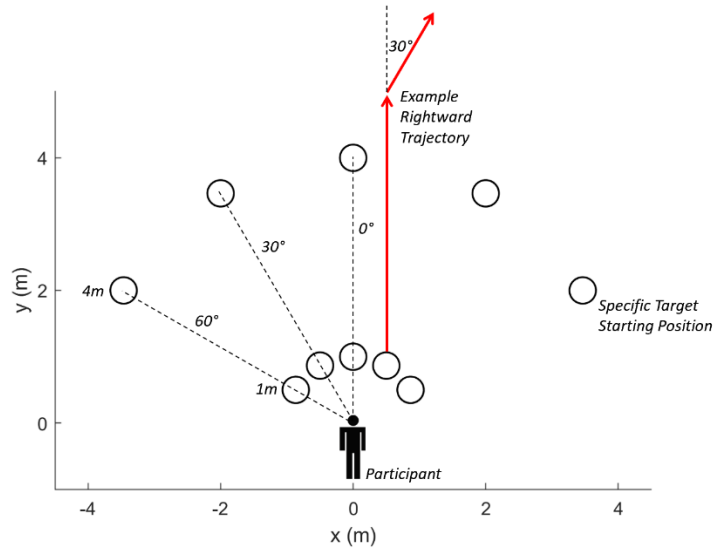


Figure 2.9. Diagram of the 10 target starting positions of Experiment 2 by distance (1 m, 4 m) and eccentricity ( $0^\circ$ ,  $\pm 30^\circ$ ,  $\pm 60^\circ$ ) relative to the participant, with one rightward trajectory plotted.

optical variables of expansion and angular velocity. At the beginning of the trial, the target pole would appear to move forward on a straight path (parallel to the participant's facing direction) at 0.8 m/s. 4.5 seconds into the trial, the pole's visual angle and/or angular velocity was perturbed to simulate a change in the pole's speed (-0.2, 0, or +0.2 m/s) or direction ( $-30^\circ$  leftward,  $0^\circ$ , or  $+30^\circ$  rightward) or both. For example, with a target ahead of the participant ( $0^\circ$  eccentricity), the visual angle of the pole was perturbed to create an optical expansion rate specifying a -0.2 m/s decrease in the pole's speed, or the angular velocity of the pole was increased to specify a  $30^\circ$  turn to the right, or both. With a target at an eccentricity of  $60^\circ$ , a particular combination of optical expansion and rightward angular velocity would specify a -0.2 m/s decrease in the pole's speed, while a different combination would specify a  $30^\circ$  turn to the right. Figure 2.9 diagrams all target starting positions relative to the participant and a rightward trajectory for one target. After the participant walked the length of the room (approximately 10 m), the participant heard a voice over headphones that said 'end', that trial would be over, and the target would



disappear. A new orientation pole would then appear, signifying the beginning of the next trial.

## **Design**

The experiment had a 5x2x3x3 factorial design: 5 target eccentricities (-60°, -30°, 0°, +30°, +60°), 2 target distances (1 m, 4 m), 3 simulated speed perturbations (-0.2 m/s, no change, +0.2 m/s), 3 simulated heading perturbations (+30°, no change, -30°). This was replicated twice over two blocks, for a total of 180 trials.

## **Data processing**

2160 trials were recorded across participants. Of those, 214 had to be thrown out due to tracking errors, leaving 1949 usable trials (90% of trials). The time series of head position (x, y) in the horizontal plane were filtered using a 4th-order low-pass Butterworth filter with a cutoff frequency of 0.6 Hz. This was to remove tracker noise and reduce oscillations due to the step cycle. The first two seconds of every trial was removed, so that the participant was at a steady-state of locomotion for analysis and modeling. The last second of every trial was also removed, to avoid any errors due to filtering at the edges of recorded data (see Howarth & Callaghan, 2009). Time series of speed and heading for each trial were computed from the filtered position data as the displacement between successive time steps.

Because the stimuli target only changed in optical size and angular velocity (instead of physically moving through space away from the participant), the target's specified position, speed, and heading were used for analysis and modeling. This specified position was calculated by taking the distance specified by the visual angle of a 0.4 m object at an angle specified by the target's eccentricity from the participant's known position at every

time step (this causes the oscillations seen in target leader time series (red) in Figure 2.10, as the target's position is computed from participant position, which oscillates).

## 2.8 Experiment 2 - Results

Representative trials from Experiment 2 appear in Figure 2.10. These plot the position, speed time series, and heading time series for the participant and the target pole

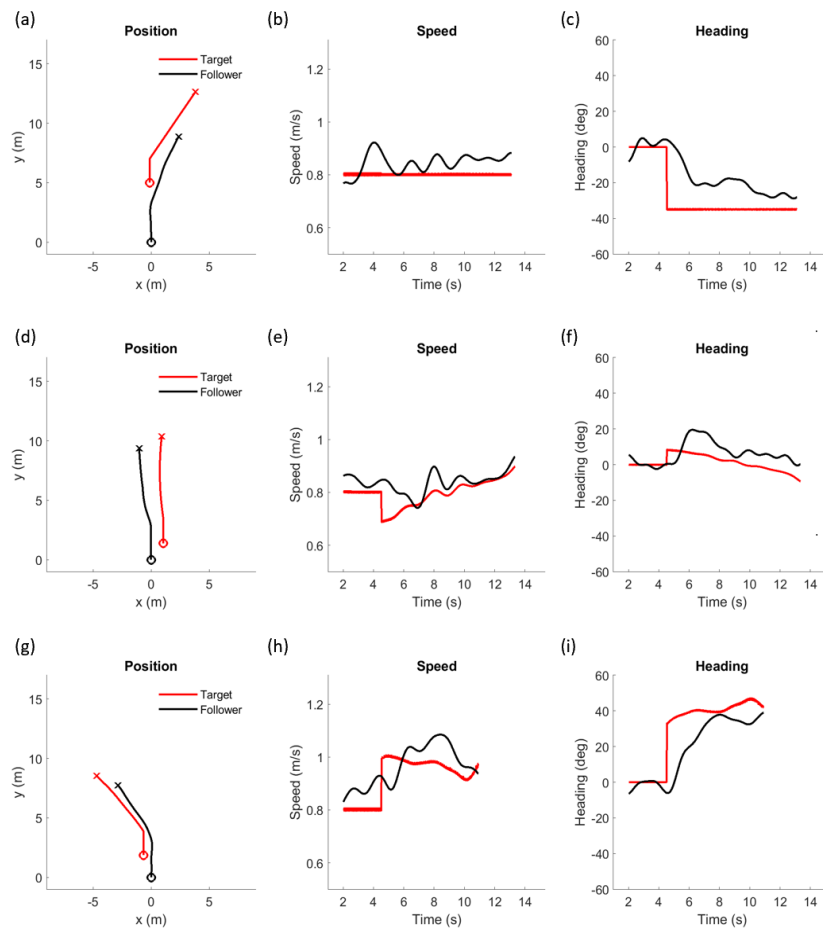


Figure 2.10. Representative trials from Experiment 2. Panels (a-c) shows a 4 m initial distance,  $0^\circ$  eccentricity trial, where the target's angular velocity specified a right turn, with constant optical size. Panels (d-f) shows a 1 m initial distance,  $60^\circ$  eccentricity trial, with constant angular velocity and a change in optical size that specified the target slowed down. Note that the target expanded while at an eccentricity  $\neq 0^\circ$ . To the participant this specified a target slowing down and turning slightly inward, which caused an accompanying change in participant trajectory. Panels (g-i) shows a 1 m initial distance,  $-30^\circ$  eccentricity trial, where angular velocity specified a left turn and the optical size changed to specify the target sped up.

as specified by the 2D optical variables of optical expansion and angular velocity. The coupling strength was high, with a mean correlation between participant and target of  $r = 0.51$  for speed and  $r = 0.53$  for heading across all trials.

As demonstrated by Euclid's Laws and Figure 2.3, identical magnitudes of a leader's relative change in speed or heading can be observed by a follower as asymmetrical magnitudes of optical expansion and angular velocity. These asymmetries are produced by two methods of motion: (1) whether the leader's relative motion brings them toward or away from the follower, and (2) whether the leader's motion is at a close or far distance. In order to measure if the asymmetries computed in Figure 2.3 cause asymmetrical locomotor responses when following a single leader in Experiment 2, the mean time series was computed in each condition for each participant and compared across conditions that match up with the computed asymmetries.

Figure 2.11 addresses the asymmetry of optical information produced by motion toward or away from a follower. Panel (a) shows the mean speed change across all participants in response to the target's change in optical expansion at an eccentricity of  $0^\circ$ . Comparing when the target contracted (specifying that it sped up; green curve) to when the target expanded (specifying that it slowed down, red curve), participants responded more quickly and to a greater degree when the target expanded, which is relative motion toward the participant. A repeated measures ANOVA comparing the mean final participant speeds across the last two seconds (shaded yellow region) found a significant difference across the three conditions ( $F(2, 33) = 288.55, p < 0.001$ ). Post hoc comparisons (Figure 2.11, panel c) found a significant differences between the contraction condition and the control ( $t(22) = 15.32, p < 0.001$ ) and between the expansion condition and control ( $t(22) = 21.30, p <$

0.001), demonstrating that optical expansion controls participant speed at 0° eccentricity. In addition, a post hoc comparison between the expansion and contraction conditions found participants changed speed more in response to a target expanding compared to contracting ( $t(22) = -4.87, p < 0.001$ ).

Figure 2.11, panel (b) shows the mean heading change across all participants in response to the target's change in angular velocity at an eccentricity of 60°. Here the participant's heading has been collapsed across 0° when the target turned toward or away

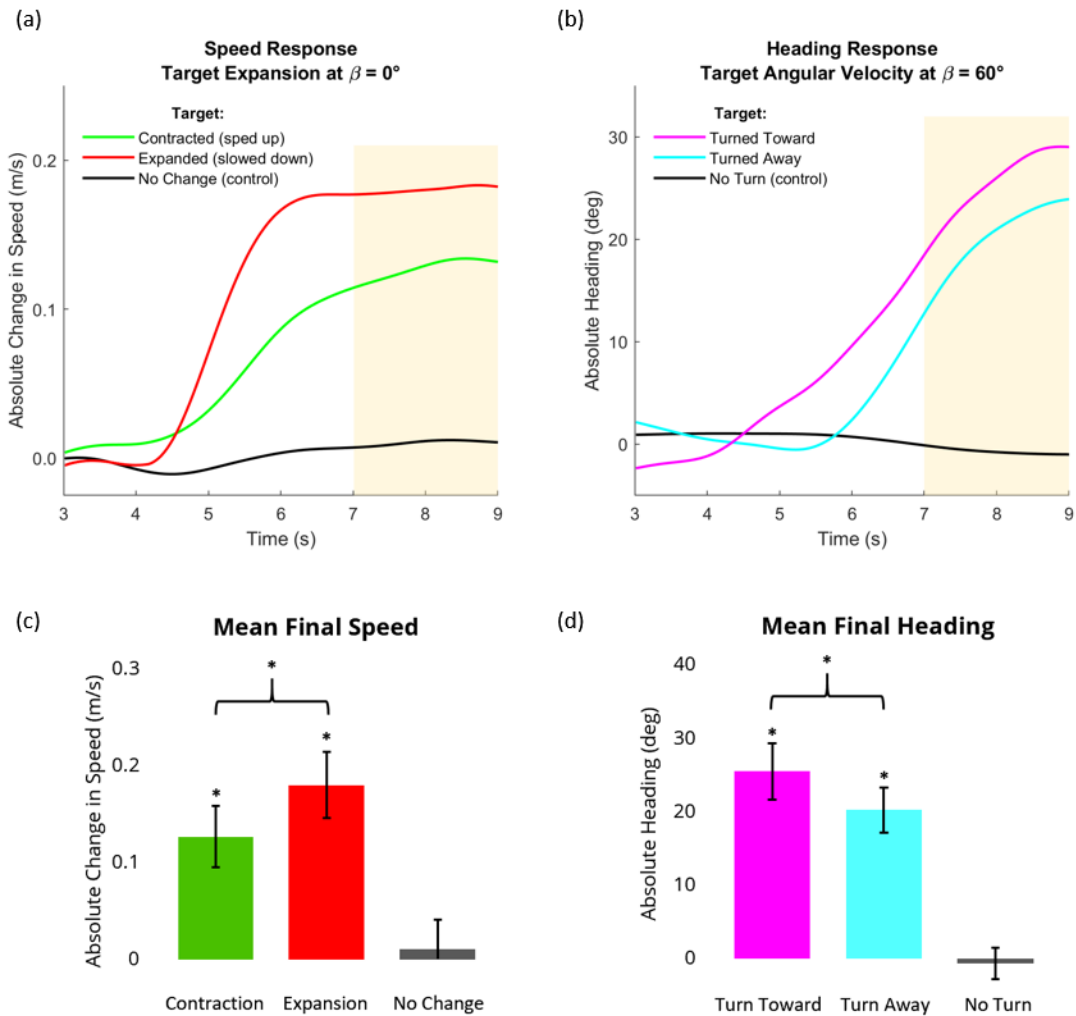


Figure 2.11. Speed and heading response to target optical expansion and angular velocity changes. Participant data demonstrates that the asymmetry of optical information caused by the relative direction of motion of a leader causes an asymmetrical speed and heading change in a follower.

from the participant. Comparing when the target turned toward the participant (pink curve) to when the target turned away (light blue curve), participants responded again responded more to relative motion toward them. A repeated measures ANOVA comparing the mean final participant heading across the last two seconds (shaded yellow region) found a significant difference across the three conditions ( $F(2, 33) = 145.81, p < 0.001$ ). Post hoc comparisons (Figure 2.11, panel d) found a significant differences between the target turning toward condition and the control ( $t(22) = 21.55, p < 0.001$ ) and between the target turning away condition and control ( $t(22) = 17.98, p < 0.001$ ), demonstrating that angular velocity controls participant heading at  $60^\circ$  eccentricity. In addition, a post hoc comparison between the turned toward and turned away conditions found heading changed more in response to a target's relative motion toward the participant ( $t(22) = 3.8259, p = 0.002$ ).

The results of panels (c) and (d) together demonstrate that optical information can act as control laws in speed for following and that the asymmetry of optical information caused by the relative direction of motion of a leader causes asymmetrical speed and heading changes in a follower.

Figure 2.12 addresses the asymmetry of optical information produced by motion at a close or far distance. Panel (a) shows the mean speed change across all participants in response to the target's change in optical expansion when the target was at 1 m distance compared to 4 m. Comparing when the target changed rate of expansion at 1 m (red curve) to when the target was at 4 m (orange curve), participants responded more quickly and to a greater degree when the target was closer. A repeated measures ANOVA comparing the mean final participant speeds across the last two seconds (shaded yellow region) found a significant difference across the four conditions ( $F(3, 44) = 76.76, p < 0.001$ ). Post hoc

comparisons (Figure 2.12, panel c) found a significant differences between the 1 m change condition and the control ( $t(22) = 14.82, p < 0.001$ ) and between the 4 m change condition and control ( $t(22) = 16.68, p < 0.001$ ), demonstrating that changes in optical expansion cause a reduced influence on participant speed at greater distances. In addition, a post hoc comparison between the 1 m and 4 m optical expansion conditions found participants changed speed more in response to a closer target compared to a farther one ( $t(22) = 6.73, p < 0.001$ ).

Figure 2.12, panel (b) shows the mean heading change across all participants in response to the target's change in angular velocity when the target was at 1 m distance compared to 4 m. Here the participant's heading has been collapsed across  $0^\circ$ , across left and right participant turns. Comparing when the target changed angular velocity at 1 m (light blue curve) to when the target was at 4 m (dark blue curve), participants responded more quickly and to a greater degree when the target was closer. A repeated measures ANOVA comparing the mean final participant headings across the last two seconds (shaded yellow region) found a significant difference across the four conditions ( $F(3, 44) = 210.84, p < 0.001$ ). Post hoc comparisons (Figure 2.12, panel d) found a significant differences between the 1 m change condition and the control ( $t(22) = 18.37, p < 0.001$ ) and between the 4 m change condition and control ( $t(22) = 16.34, p < 0.001$ ), demonstrating that changes in angular velocity cause a reduced influence on participant heading at greater distances. In addition, a post hoc comparison between the 1 m and 4 m angular velocity conditions found participants changed heading more in response to a closer target compared to a farther one ( $t(22) = 4.56, p < 0.001$ ).

The results of panels (c) and (d) together demonstrate that the asymmetry of optical information caused by the distance of a leader cause asymmetrical speed and heading changes in a follower.

Taken together, the results from Experiment 2 provide clear evidence optical expansion and angular velocity are effective information for the control of pedestrian following. The asymmetries present in Experiment 2 further demonstrate that pedestrians use optical information (as opposed to direct changes in speed and heading) to align locomotion with a leader when following. This is a critical property of pedestrian

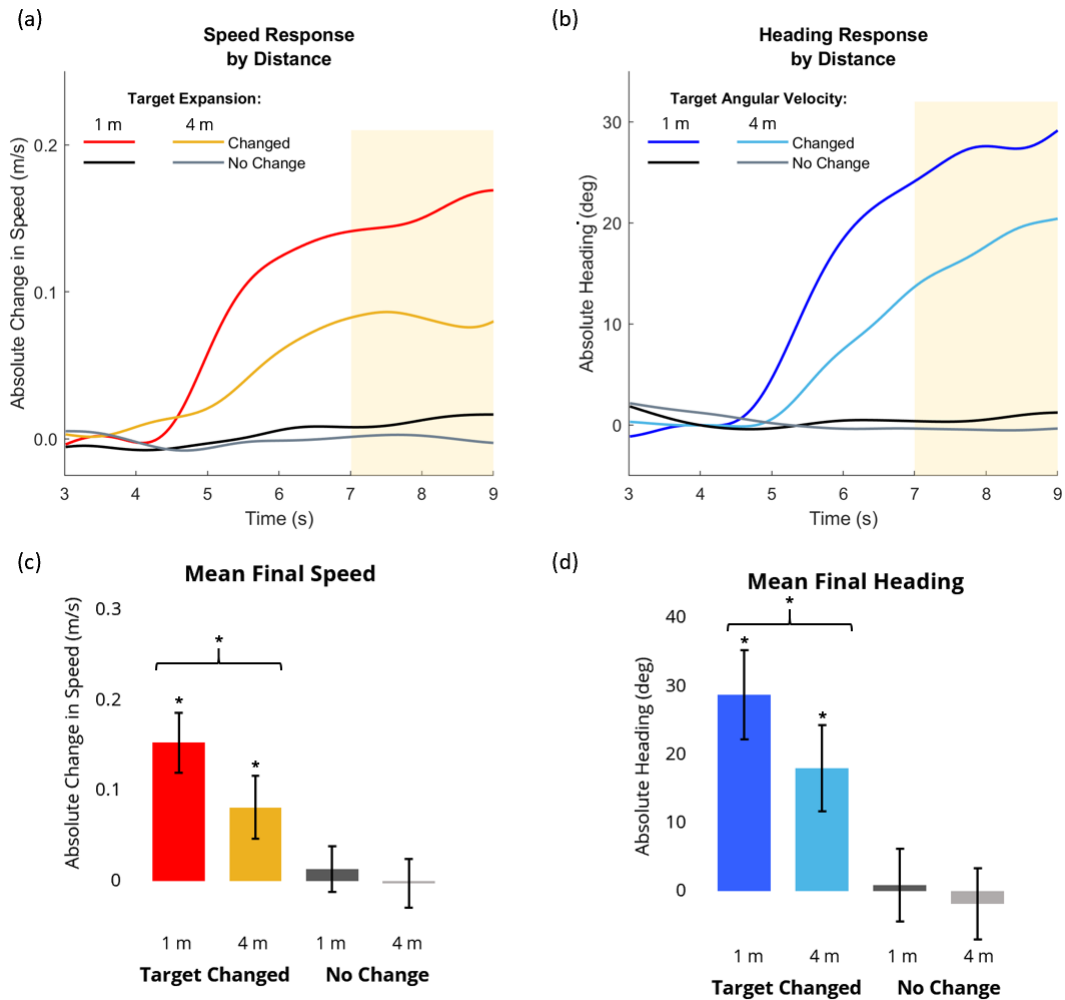


Figure 2.12. Speed and heading response to target changes at 1 m compared to 4 m. Participant data demonstrates that the asymmetry of optical information caused by the distance to a leader causes an asymmetrical speed and heading change in a follower.

following, one that the visual model of following is equipped to reproduce. This allows Experiment 2 to serve as a test for the visual model for following, based on optical expansion and angular velocity.

## 2.9 Experiment 2 – Model Simulations

To compare the behavioral and visual models, all trials in Experiment 2 were simulated using both models with fixed parameters, taking the participant's initial conditions as input. The behavioral model (Dachner & Warren, 2014; Rio et al., 2014) relies on the physical variables of speed, heading, and distance of the leader and follower as input. A theoretical limitation of the behavioral model is that it cannot account for the

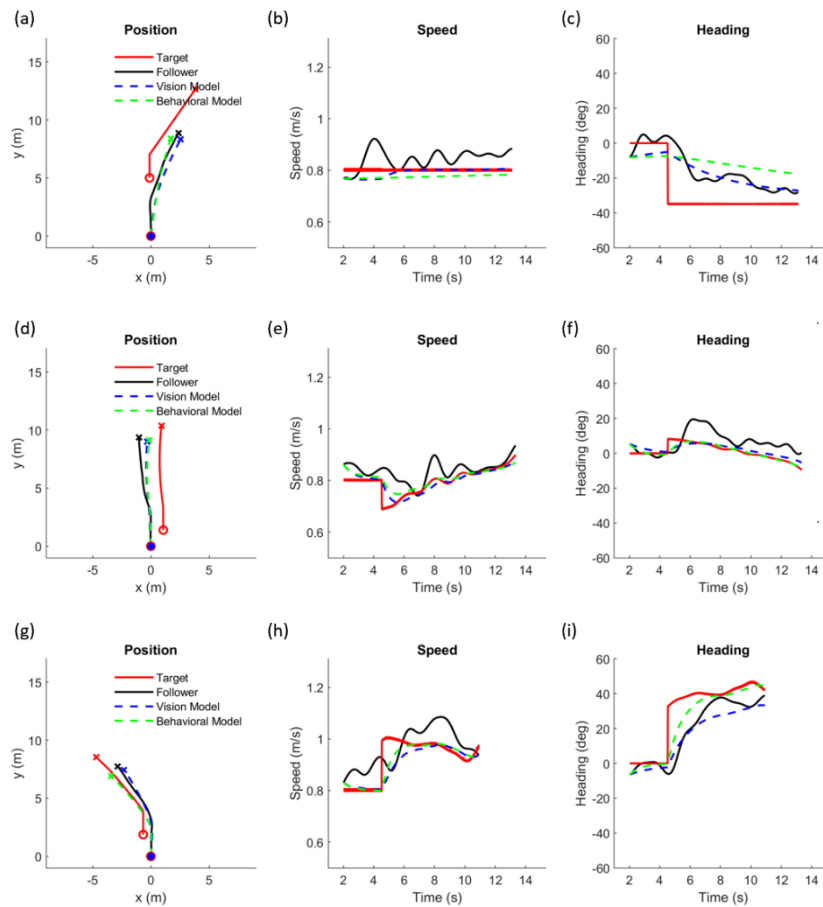


Figure 2.13. Representative trials from Experiment 2. Each trial was simulated using the visual model and the behavioral model.



asymmetries found in follower response in Experiment 2 and instead relies on the physical variables that are not directly perceived. Thus, the computed physical variables served as input to the behavioral model, whereas the target's computed visual angle and angular velocity were taken as input to the visual model. Figure 2.13 shows representative trials from Experiment 2 modeled using both the visual model and the behavioral model.

The mean time series for both models was taken across conditions for comparison. Figure 2.14 plots the model means from the conditions previously plotted in Figure 2.11, which looked at the optical asymmetry due to whether a leader's relative motion is toward or away from a follower. Panel (a) shows the mean speed change across both models, simulating a follower's response to the target's change in optical expansion at an eccentricity of  $0^\circ$ . Critically, the visual model (dashed curves) captured the asymmetrical response to a target contracting (green; specifying the target sped up) compared to expanding (red; specifying the target slowed down), while the behavioral model (dotted

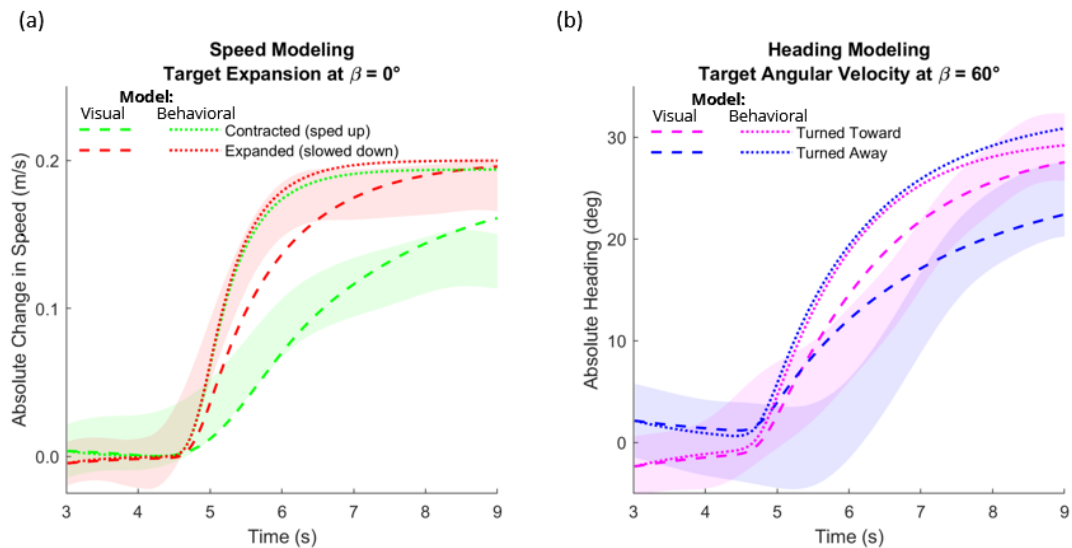


Figure 2.14. Mean speed and heading time series of both models for changes in optical expansion and angular velocity for motion toward and away from the follower. Shaded regions are the 95% confidence interval of participant data. Results demonstrate that the visual model better simulates asymmetrical response caused by the relative direction of motion of a leader.

curves) did not capture this asymmetry. The shaded regions are the 95% confidence interval of the participant data, further demonstrating the visual model's speed simulations perform more closely to participant data.

Figure 2.14, panel (b) shows the mean heading change across both models, simulating a follower's response to the target's change in angular velocity at an eccentricity of  $60^\circ$  (collapsed across  $0^\circ$ ). As with the speed time series, the visual model (dashed curves) captured the asymmetrical response to a target turning toward (pink) compared to turning away (blue), while the behavioral model (dotted curves) did not. Here again the asymmetry is simulated by the visual model, where a target turning away caused a reduced maximal turn compared to a target turning toward the model. The shaded regions are the 95% confidence interval of the participant data, again demonstrating the visual model's heading simulations also perform more closely to participant data.

Together, these simulations demonstrate a follower's asymmetrical response to a leader's relative motion toward or away from the follower is likely driven by the asymmetry in optical variables, as predicted by Euclid's laws of optics.

Figure 2.15 plots the model means from the conditions previously plotted in Figure 2.12, which looked at the optical asymmetry due to the distance between leader and follower. Panel (a) shows the mean speed change across both models, simulating a follower's response to the target's change in optical expansion when the target was at 1 m distance compared to 4 m (collapsed across the control speed of 0.8 m/s). As with the previous asymmetry, the visual model (dashed curves) captured the asymmetrical response between a close target (red, 1 m) and a far target (orange; 4 m), while the behavioral model (dotted curves) did not capture this asymmetry. The shaded regions are the 95% confidence

interval of the participant data, further demonstrating the visual model’s speed simulations perform more closely to participant data.

Figure 2.15, panel (b) shows the mean heading change across both models, simulating a follower’s response to the target’s change in angular velocity when the target was at 1 m distance compared to 4 m. Here again, the visual model (dashed curves) captured the asymmetrical heading response between a close target (dark blue, 1 m) compared to a far target (light blue, 4 m), while the behavioral model (dotted curves) did not. The asymmetry due to distance is simulated by the visual model. The shaded regions are the 95% confidence interval of the participant data, again demonstrating the visual model’s heading simulations also perform more closely to participant data.

Due to Euclid's law of visual angles, closer targets produce larger visual change, than farther targets, eliciting a faster and larger response. This asymmetry is what is captured by the visual model, allowing it to simulate asymmetries in both speed and heading that the behavioral model cannot capture.

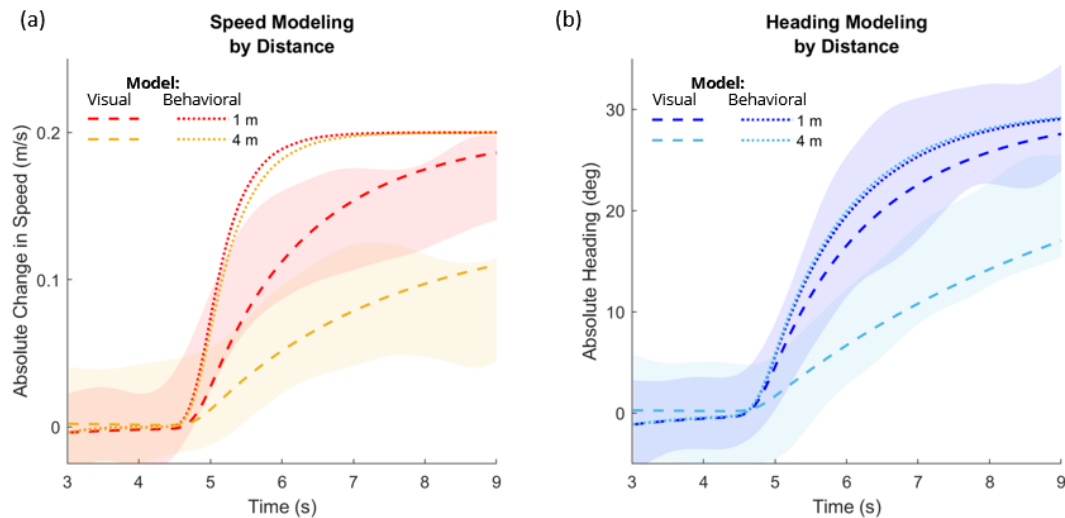


Figure 2.15. Mean speed and heading time series of both models for changes in optical expansion and angular velocity, comparing effect of distance. Shaded regions are the 95% confidence interval of participant data. Results demonstrate that the visual model better simulates asymmetrical response caused by distance to the leader.

Quantitative measures of error between the participants and models for speed time series and heading time series were taken in order to test which model best captured participant behavior (Figure 2.16). Tests comparing the mean metric distance between participant position and model position over time revealed that the visual model has significantly less RMSE for speed ( $t(22) = 2.21, p < 0.001$ ), and less RMSE for heading ( $t(22) = 10.05, p < 0.001$ ) than the behavioral model. Bayesian model comparison (Nuijten, et. al, 2015) yielded a scaled JZS Bayes Factor of  $BF_{10} = 3.13$  for speed, indicating substantial evidence for support of the speed component of the visual model, and a scaled JZS Bayes Factor of  $BF_{10} = 5649731$  for heading, indicating decisive evidence for support of the heading component of the visual model. There was also less error in measured position (x, y) distance between the participant data and the visual model compared with the behavioral model, but quantitative analysis was not conducted on these means, as position is not independent from speed and heading. These results demonstrate that the visual model performs better than the behavioral model on each metric. This finding implies that the visual model relies on the same optical variables that human pedestrians

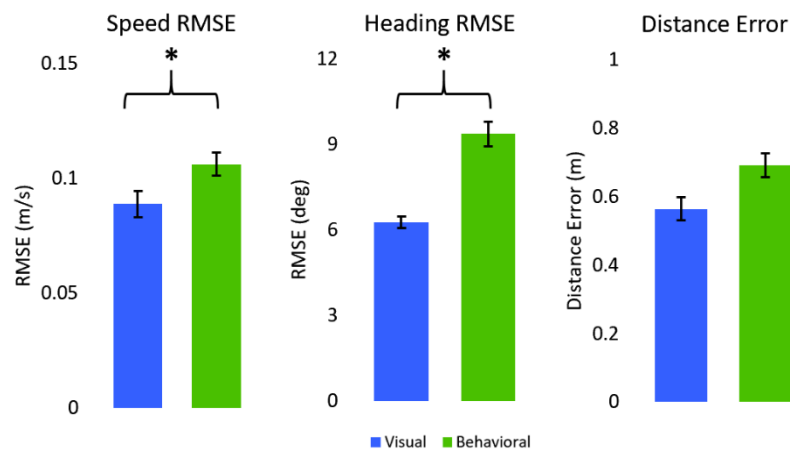


Figure 2.16. Comparison of error measurements between the participant and each model: mean metric distance between participant position and model positions, RMSE for speed, and RMSE for heading. In all cases, the visual model produces less error.

use to control following (optical expansion and angular velocity) as opposed to the physical variables of speed and heading, which themselves must be derived from optical variables by the visual system. This property allows the visual model to capture asymmetries and effects of distance the behavioral model cannot.

## **2.10 Discussion**

The presented results provide clear evidence for the strength of a model for following a leader based on visual input. Experiment 1 demonstrated that pedestrian followers successfully align their speed and heading with that of a leader, and that dyad pedestrians are tightly coupled in their locomotor behaviors. Experiment 2 demonstrated that optical variables alone are sufficient to explain following behavior. Specifically, optical expansion and angular velocity trade off in controlling the follower's speed and heading as a function of the leader's eccentricity. The results from both studies were then used to test a visual model for following a single target. This model was fit to the trajectories from Experiment 1, and generalized to the trajectories from Experiment 2 with the fixed parameters found in Experiment 1. This model can capture aspects of pedestrian following (such as effects of distance and asymmetries in alignment response) that a behavioral model cannot. This power of the visual model is likely due to the fact it is driven by visual information, which is inherently asymmetrical, as derived by Euclid's law of visual angles and demonstrated by Figure 2.3. This information asymmetry drives asymmetrical locomotor following behaviors in pedestrians.

## 2.11 Conclusion

We believe that the ability of a pedestrian to follow their neighbor is a critical component of collective crowd behavior. Thus far we have derived a visual control model that accounts for following a single neighbor. We now propose to scale up this model to explain following a crowd. The subsequent chapters will determine how the visual model accomplishes this task. If the model scales up to larger crowds, it can be applied to large-scale simulations of collective motion and tackle open questions about the information pedestrians use to follow not just one neighbor, but a whole crowds. The aim of this project is a simple empirically-grounded model of the visual coupling that can serve as a basis for understanding self-organized collective motion.

## CHAPTER 3

### The Visual Model: Following a Crowd

### **3.1 Introduction**

It is widely believed that collective motion in human crowds emerges from the local interactions between individual pedestrians. In many models of collective motion, a critical interaction is following a neighbor, in which the velocity vector (speed and heading direction) is aligned with the neighbor's velocity vector. Based on experiments on pedestrian following, Chapter 2 formalized this interaction as a visual control law in which the follower's speed and heading are controlled by the optical expansion and angular velocity of a leader. In addition, Chapter 2 showed how the visual model, which uses optical variables as input, performs better than a behavioral model that uses the physical variables of speed and heading as input. Evidence shows that the optical information of a leader is inherently asymmetrical when observed by a follower. These asymmetries lead to asymmetrical following behaviors, which are captured by the visual model.

The visual control law for following a single neighbor may be extended to following a crowd in a crowd by the principle of superposition (Rio, Dachner, & Warren, 2018). Thus, the interaction rule for a single neighbor found in Chapter 2 can be applied to the interactions in a whole crowd. Chapter 3 will derive an averaging version of the visual model and experimentally test it using three datasets of crowd behavior, to test whether the visually-guided location principles found in Chapter 2 also apply across a crowd.

### **3.2 Visual model for following a crowd**

Equations (3) and (4) in Chapter 2 describe a visual model for following a single neighbor that takes their optical expansion, angular velocity, and eccentricity as input. Similar equations will be used for following a crowd, due to the assumption that the



superposition principle applies to following in a crowd (Rio, Dachner, & Warren, 2018). This assumption states that the response to multiple neighbors is the sum (or average) of the responses to each individual neighbor. That is, the responses are independent, and can be linearly combined. Chapter 3 will take the equations that describe the control law for an individual neighbor (as derived in Chapter 2), and apply them to each neighbor in a crowd, where the visual control law for speed is given by:

$$\dot{r} = \frac{1}{n} \sum_{i=1}^n (-c_1 \cdot \sin(\beta_i) \cdot \dot{\psi}_i - c_2 \cdot \cos(\beta_i) \cdot \dot{\theta}_i) \quad (5)$$

and the visual control law for heading is given by:

$$\ddot{\phi} = \frac{1}{n} \sum_{i=1}^n (-c_3 \cdot \cos(\beta_i) \cdot \dot{\psi}_i + c_4 \cdot \sin(\beta_i) \cdot \dot{\theta}_i) \quad (6)$$

where  $\dot{r}$  is the follower's radial acceleration,  $\ddot{\phi}$  is the follower's angular acceleration,  $\beta_i$  is the eccentricity neighbor  $i$ ,  $\dot{\psi}_i$  is the angular velocity of neighbor  $i$ , and  $\dot{\theta}_i$  is the optical expansion rate of neighbor  $i$ .  $c_{1-4}$  are the same parameters as found in Chapter 2, which are held fixed to avoid overfitting and to demonstrate the visual model generalizes from single to multiple neighbors. The optical variables are computed from the follower's observation point for each neighbor in the field of view, summed, and then divided by the number of neighbors within the field of view, yielding the mean of optical influence across the crowd in view.

Note, the model will ignore any neighbors that are beyond  $\pm 90^\circ$  eccentricity reflecting the fact that a pedestrian in the crowd cannot perceive neighbors outside the field of view (FOV). This is important, because it restricts information transfer in the crowd to

passing from the front of the crowd toward the back, and does not allow for the back-most neighbors in a crowd to directly influence those in the front.

Also note that, in the visual model, a pedestrian's neighborhood does not have a limited range of interaction. With the behavioral model, neighbors beyond 5 m no longer influence the pedestrian, reflecting an exponential decay with distance observed in empirical data (Rio, Dachner, & Warren, 2018). In the visual model, the observed decay with distance is a natural consequence of Euclid's law of visual angles. When computing the average influence, the visual model divides by  $n$ , where  $n$  is the total number of neighbors in the field of view. However, it may become necessary to define a cutoff for  $n$ , based on detection thresholds for optical expansion (Regan & Beverley, 1979a; Wann, et al., 2011; Hoffmann, 1994) and angular velocity (Regan & Beverley, 1973b; Trent & Warren, 2017), below which a neighbor's relative motion is not perceived. This threshold would capture three interacting variables for how object motion translates to optical expansion and angular velocity: target size, target distance, and magnitude of motion change.

In this chapter, I test whether the visual model (Equations 5 & 6) generalizes from single to multiple neighbors by simulating two datasets from previous crowd experiments. Experiment 3 models the effect of neighbor distance within a virtual crowd on the participant's control of heading and speed, and Experiment 4 models the effect of distance to the virtual crowd as a whole.

### **3.3 Experiment 3 - Effect of Distance in a Crowd**

Experiment 3 is crowd data taken from Experiment 2 in Rio, Dachner, & Warren (2018). The original purpose of the experiment was to find if the influence of perturbed neighbors in a crowd decreased with their distance from the participant. This effect was modeled by weighting the influence of each neighbor in the behavioral model, with a weight decayed exponentially with distance (Rio, Dachner, & Warren, 2018). The data will be simulated with both the behavioral and the visual model in order to compare the two models. In theory, the visual model can naturally capture an effect of distance. As metric distance increases to a neighbor moving at a fixed speed, their relative angular velocity and optical expansion rate decreases in accordance with Euclid's law of visual angles. Specifically, as distance increases, the visual angle subtended by an object (as well as its rate of change for a moving object) decreases with the arctangent of the inverse of distance, which is approximated by an exponential function. This should lead to naturally reduced influence of distant neighbors in the visual model, without requiring the explicit distance weight in the behavioral model.

#### **Participants**

10 participants were recruited at Brown University, 6 female and 4 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

#### **Apparatus**

The experimental apparatus was identical to Experiment 2.

## **Procedure**

Participants were instructed to walk as naturally as possible, to treat the virtual pedestrians as if they were real people, and to stay together with the crowd. On each trial, the participant walked to the start pole and faced the orienting pole. After 2 seconds, the poles disappeared and the virtual crowd appeared; 1 second later, a verbal command (“Begin”) was played over headphones and the virtual crowd began walking. The display continued until the participant had walked about 12 m; a verbal command (“End”) signaled the end of the trial. In each experiment, there were eight heading trials and eight speed trials per condition, presented in a randomized order, with 80 trials in each 1-hour session.

## **Displays**

The virtual environment consisted of a granite-textured ground plane with a green start pole and a red orienting pole (3 m high, 0.2 m radius, 12.7 m apart) and blue sky. The virtual crowd was generated using 3D human models (WorldViz Complete Characters), animated with a walking gait at a randomly varied phase. Thirty virtual humans were positioned on two circles (radius 1.5 m, 3.5 m) with the participant at the center to enhance the sense of immersion. Twelve neighbors ( $N = 12$ ) were experimentally manipulated, and appeared on two 90° arcs centered on the initial walking direction, within the HMD field of view. Five of these neighbors were placed at equal intervals on the 1.5 m radius arc (near zone), and seven on the 3.5 m radius arc (far zone); their positions were then subjected to Gaussian jitter in polar coordinates (distance  $r$  standard deviation = 0.15 m; eccentricity  $\beta$  standard deviation = 8°). The remaining 18 neighbors were also placed at equal intervals and similarly jittered. A different configuration was generated for each trial; all participants

received the same set of configurations, but virtual humans were randomly assigned to the positions.

During a trial, all virtual humans accelerated from a standstill (0 m/s) to a speed of 1.0 m/s over a period of 3 seconds in the participant's walking direction, following an ogive function ( $\mu = 0$ ,  $\sigma = 0.5$  seconds) fit to previous data. On perturbation trials, after 2 seconds a subset (S) of the 12 neighbors then changed their heading direction ( $\pm 10^\circ$  left or right) or speed ( $\pm 0.3$  m/s) over a period of 0.5 seconds, following another ogive function ( $\mu = 0$ ,  $\sigma = 0.083$  seconds).

### **Design**

Experiment 3 had a  $2 \times 2 \times 2 \times 5$  factorial design: 2 perturbations (heading or speed) x 2 direction (left/right or fast/slow) x 2 distance (near or far row) x 5 number perturbed (0, 3, 6, 9, 12 in subset), for a total of 40 conditions. There were 4 repetitions per condition, for a total of 160 trials per participant. These trials were run across two sessions on consecutive days, in a randomized order.

### **Data processing**

Data was processed using the same procedures from Experiment 2. Trials were simulated using the visual and behavioral models as before.

## **3.4 Experiment 3 - Results**

To quantify the results of Experiment 3, participant mean final speed and heading was taken across the last two seconds of every trial in response to crowd perturbations. Both mean final speed and heading were collapsed across the control (1.0 m/s and  $0^\circ$ ,

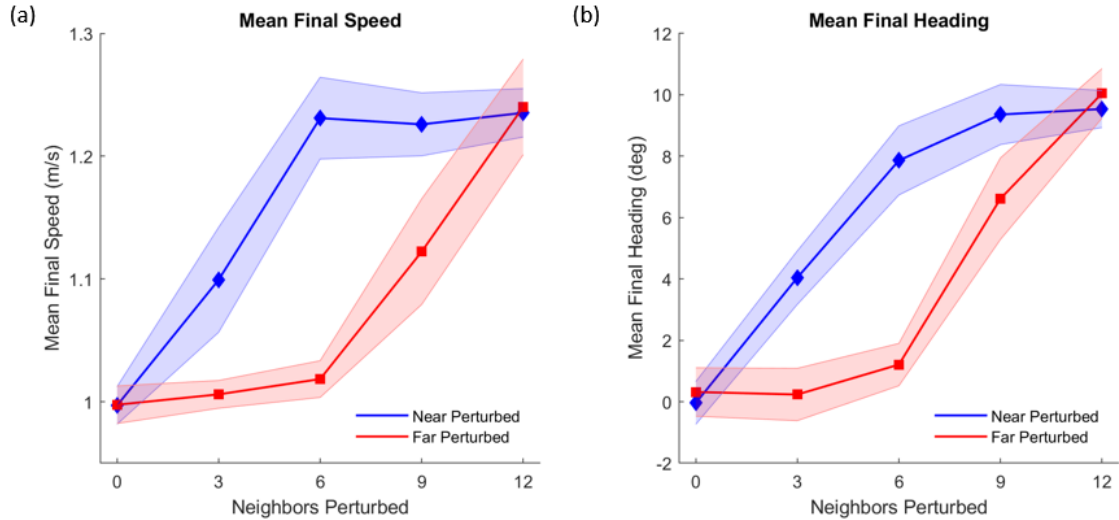


Figure 3.1. Mean final speed (a) and heading (b) of the last two seconds of participant data from Experiment 3, broken down by distance conditions and number of neighbors perturbed. Results are consistent with superposition and distance hypotheses.

respectively), broken down by distance conditions and the number of perturbed neighbors (Figure 3.1). The response to the near row is greater than the far row for both speed and heading perturbations, and the curves converge when all 12 neighbors are perturbed. Results showed a significant effect on participant speed from neighbor distance ( $F(1, 9) = 22.93, p < 0.001$ ) and number of perturbed neighbors ( $F(4, 36) = 34.28, p < 0.001$ ). Similarly, they showed a significant effect on participant heading from neighbor distance ( $F(1, 9) = 71.57, p < 0.001$ ) and number of perturbed neighbors ( $F(4, 36) = 244.66, p < 0.001$ ). Taken together, these results are consistent with the superposition and neighbor distance hypotheses for following in a crowd and provide a good dataset to test the visual model for following in a crowd.

### 3.5 Experiment 3 – Model Simulations

To compare the behavioral and visual models, all trials in Experiment 3 were simulated using both models, taking the participant's initial conditions as input. The visual

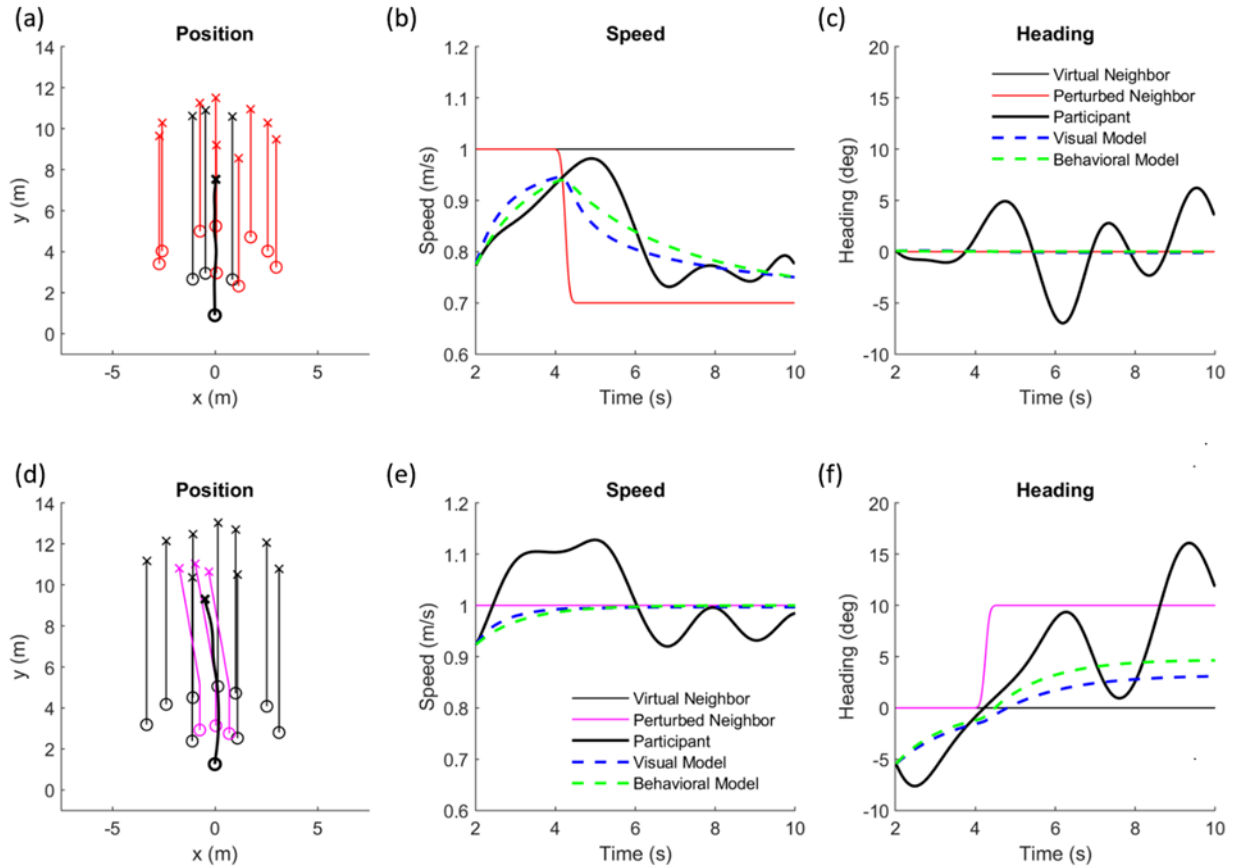


Figure 3.2. Representative trials from Experiment 3, plotting the time series of position, speed, and heading over time for the participants, the virtual neighbors, and the visual and behavioral models. Panels (a-c) demonstrates 9 perturbed neighbors slowing down in the far row. Panels (d-f) demonstrates 3 perturbed neighbors turning left in the near row.

model used equations (5) and (6) and took the fixed parameters found in Chapter 2 for the visual model of following a single leader. The behavioral model used Equations (4.1) and (4.2) and the fixed parameters from Rio, Dachner, and Warren (2018). Representative trials and simulations from Experiment 3 appear in Figure 3.2. These plots show the position, speed, and heading for the participant, 12 virtual neighbors, and the visual and behavioral models, with one trial for the perturbation of speed and another for perturbation of heading.

The mean final 2 seconds of speed and heading for each trial was computed to compare model performance to participant data. Results of participants and simulations by

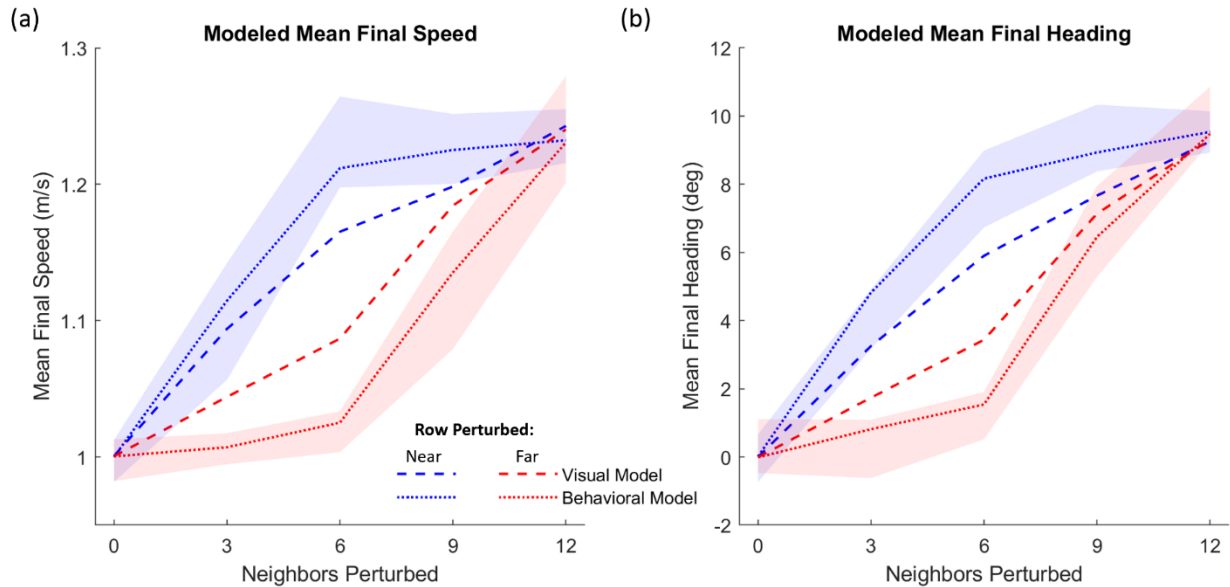


Figure 3.3. Mean final speed (a) and heading (b) of the last two seconds of participant data, and simulations using the visual and behavioral models, taken from Experiment 3. The shaded regions are the 95% confidence interval of participant means. Note that in both cases, the visual model does not capture participant behavior.

number of neighbors perturbed and distance for speed and heading can be found in Figure 3.3. The shaded regions are the 95% confidence intervals for participant means. While the behavioral model for near and far distances and both speed and heading (dotted curves) fall within the confidence interval of human data, the visual models (dashed curves) do not. Tests comparing the RMS error between participant final means and both models across speed change trials finds a significant difference of speed between models ( $t(18) = 2.244$ ,  $p = 0.038$ ), and across heading change trials finds a significant difference of heading between models ( $t(18) = 3.202$   $p = 0.005$ ). These are found in Figure 3.4. Bayesian model comparison yielded a scaled JZS Bayes Factor of  $BF_{10} = 2.05$  for speed, indicating anecdotal evidence for the speed component of the behavioral model, and a scaled JZS Bayes Factor of  $BF_{10} = 8.73$  for heading, indicating substantial evidence for the heading component of the behavioral model.



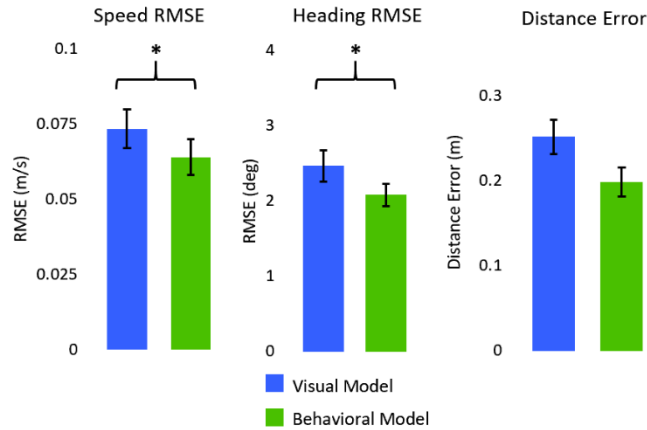


Figure 3.4. Comparison of error measurements between the participant and each model: RMSE for speed, RMSE for heading and mean metric distance between participant position and model positions, for Experiment 3.

Overall, these tests demonstrate the visual model does not explain following a crowd as well as the behavioral model. These results are opposite to what was found in Chapter 2 for following a single neighbor. While both models capture effects of distance in following of a single neighbor, the visual model appears to over-weight the influence of far neighbors in a crowd. This can be seen in Figure 3.3, in which the visual model overshoots the human response when far neighbors are perturbed (red), indicating they are weighted too strongly. Conversely, the visual model undershoots the human response when near neighbors are perturbed (blue), again indicating that (in this case, unperturbed) far neighbors are weighted too strongly.

There are two possible explanations for this phenomenon. One explanation for the performance of the behavioral model is that far neighbors are drifting outside the neighborhood (the maximal distance of influence of 5 m), and so only closer neighbors are influencing the behavioral model. In the case of visual model, the neighborhood has no maximal distance of the neighborhood, and so farther neighbors are still influencing that model, thus increasing their weight. This implies there might be a limited distance range

of interaction that the behavioral model captures, and the visual model does not. Another explanation is that the influence of far neighbors depends not only on distance, but also on occlusion. In Experiment 3, neighbors in the far row are not just at a farther distance, but as a consequence are also more likely to be partially or fully occluded by neighbors in the near row. This could be an instance of a more general rule about crowds: the farther the neighbor within a crowd, the greater the likely degree of occlusion by nearer neighbors. Partially occluded neighbors may have reduced influence on a pedestrian in a crowd. The absence of occlusion in the visual model could be causing the over-weighting of far neighbors in simulations of Experiment 3. Simulating the data of Experiment 4 will seek to differentiate these explanations.

### **3.6 Experiment 4 - Effect of Distance to the Nearest Neighbor in a Crowd**

Experiment 4 simulates data from the Honors Thesis of Emily Richmond (2016), reported in Wirth & Warren (2016). The experiment's goal was to determine whether the radius of a pedestrian's neighborhood depends on distance to the nearest neighbor in the crowd, distance within the crowd, or both. Simulations of behavioral and visual models will help understand the role of neighbor distance in the visual model.

#### **Participants**

12 participants were recruited at Brown University, 8 female and 4 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

## **Apparatus**

The experimental apparatus was identical to Experiment 3, except for the HMD was an Oculus Rift DK2 (960 x 1080 pixels per eye, 100° field of view, 75 Hz frame rate, 100% binocular overlap), which sports a larger field of view and higher frame rate.

## **Procedure**

Participants were instructed to "walk with" 12 virtual neighbors, and to "treat them as if they were real people". To begin each trial, participants stood at a green start pole and faced a grey orienting pole. After 2 seconds, the virtual neighbors appeared in front of the participant. After 1 second, a verbal command "begin" was heard over a loudspeaker, and the virtual pedestrians began walking. When the participant had walked for 12 seconds, a second verbal command "end" signaled the end of the trial. The participant then began the next trial by walking to a new green start pole, which appeared near the stopping point of the previous trial.

## **Display**

The virtual neighbors appeared in three arcs (near, mid, far arcs), with 4 neighbors in each arc. The initial distance between the participant and the 'near arc' varied by condition (2 m, 4 m, 6 m) with the successive mid arc spaced 2 m beyond the near arc and the far arc spaced another 2 m beyond that (Figure 3.5). A manipulated arc was randomly chosen on each trial and the heading direction of all 4 neighbors in that arc was perturbed ( $\pm 10^\circ$  left / right). Participant response trajectory was recorded.

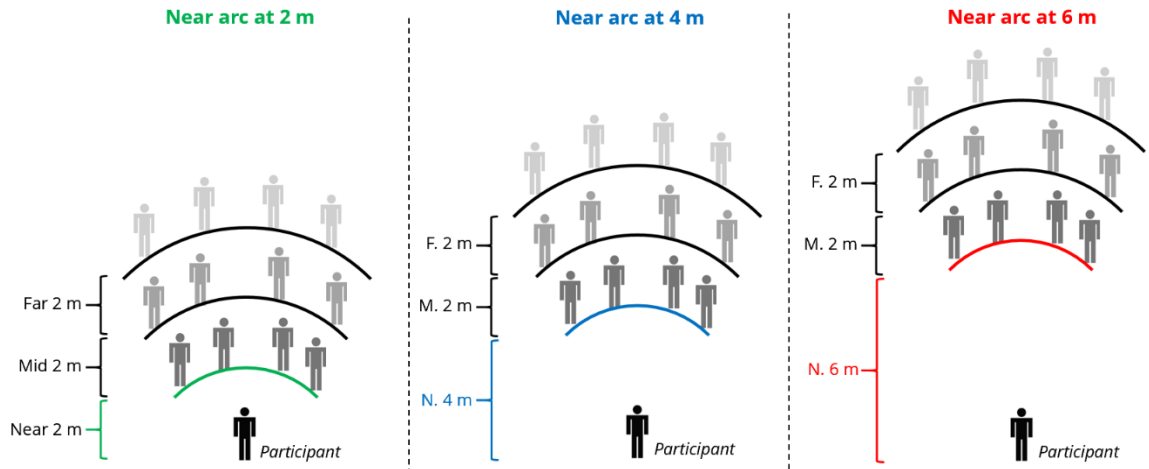


Figure 3.5. Diagram of near, mid, and far arcs by distance to near arc for Experiment 4.

## Design

Experiment 4 had a 3x3x3 factorial design: 3 directions of change (left, no change, right) x 3 distances to the near arc (2 m, 4 m, 6 m) x 3 perturbed arcs (near, mid, far), for a total of 27 conditions. There were 3 repetitions per condition, for a total of 81 trials per participant.

## Data processing

Data was processed using the same procedures from Experiment 3. Trials were simulated using the visual and behavioral models as before.

## 3.7 Experiment 4 – Results

To quantify the results of Experiment 4, participant mean final heading was taken across the last two seconds of every trial in response to arc perturbations. The mean final heading was then collapsed across left / right turns ( $0^\circ$ ), and broken down by the distance to the near arc and the distance to the perturbed arc (Figure 3.6). This plot shows that as the distance to the near arc increases (comparing the first point of the near, mid, and far

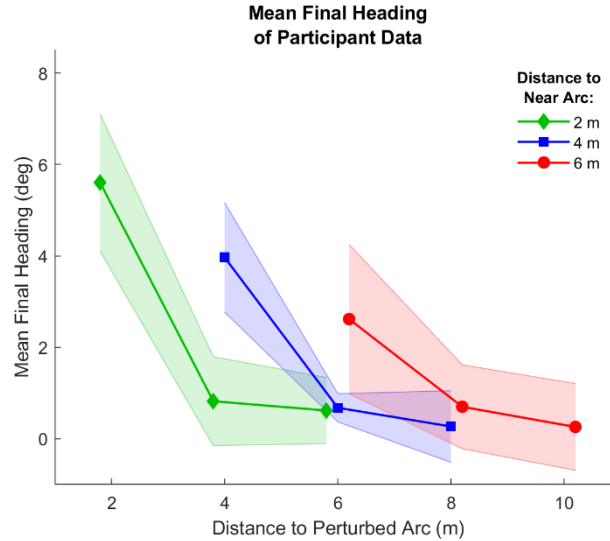


Figure 3.6. Mean final heading of the last two seconds of participant data from Experiment 4, broken down by the distance to the near arc and the distance to the perturbed arc. Results are consistent with the double decay effect of distance hypothesis.

arcs), response to perturbation decreases - this is the known effect of distance. However, as distance to the perturbed arcs increases within a crowd (comparing the second and third points in an arc to the first point), a sharp drop off is also found. Taken together, the participant results for Experiment 4 shows that influence of neighbors drops off gradually to the nearest neighbor, and then that influence drops off rapidly to neighbors within the crowd. Comparing the condition mean where the near arc is at 2 m (green curve) and the perturbed arc is at 6 m (the third point along the green curve) to the condition mean where the near arc is at 6 m (red curve) and is also the perturbed arc (the first point along the red curve), demonstrates this property. Both perturbed arcs are at the same distance (6 m), but when that is also the distance to the near arc, it produces the largest response in the participant. Results found a significant effect distance to the near arc ( $F(2, 81) = 5.91, p < 0.01$ ), and a significant effect of perturbed arc ( $F(2, 81) = 59.75, p < 0.001$ ), with a significant interaction between the two ( $F(4, 81) = 2.622, p < 0.05$ ). Taken together, these results are consistent with the hypothesis that influence of neighbors drops off gradually to

the nearest neighbor, and then that influence drops off rapidly to neighbors within the crowd, as a double decay with distance influence.

### 3.8 Experiment 4 – Model Simulations

To compare the behavioral and visual models, all trials in Experiment 4 were simulated using both models, taking the participant's initial conditions as input. The simulations were conducted in the same manner as the simulations for Experiment 3. Representative trials and simulations from Experiment 4 appear in Figure 3.7. These plots show the position, speed, and heading for the participant, 12 virtual neighbors, and the visual and behavioral models. Panels (a-c) show a trial in which the near arc is at 2 m, and

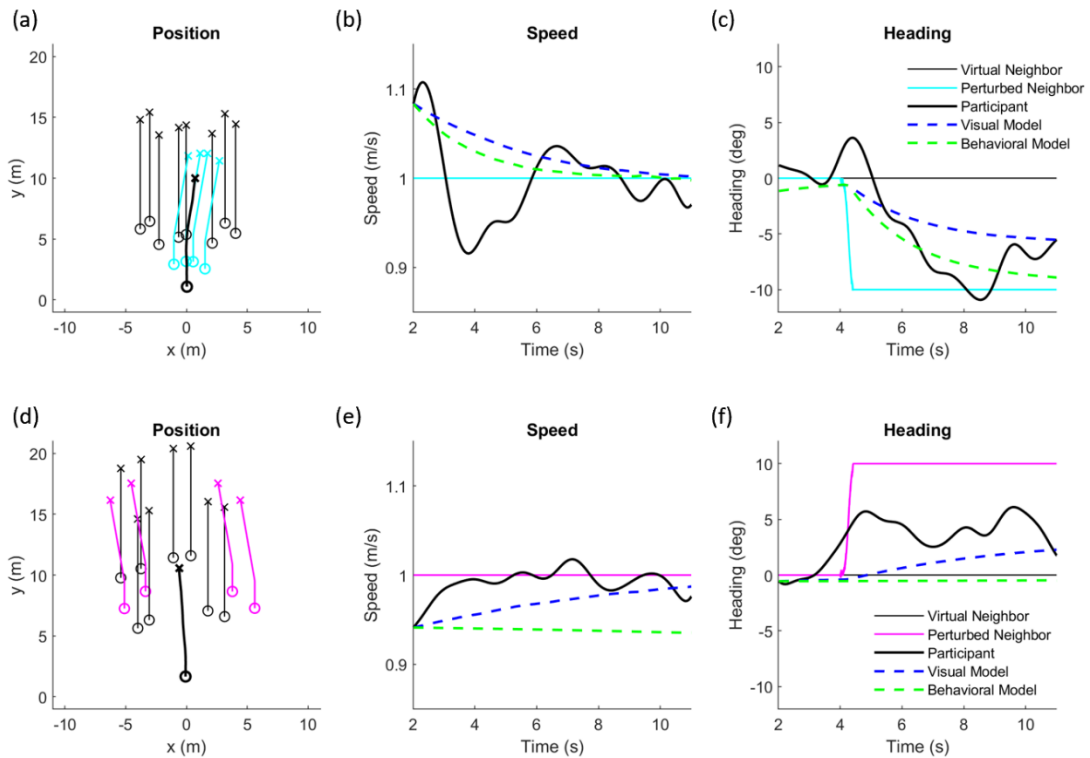


Figure 3.7. Representative trials from Experiment 4, plotting position, speed, and heading for participants, virtual neighbors, and simulations of both models.

is also the perturbed arc, which turns right. Panels (d-f) show a trial in which the near arc is at 6 m, and the mid row is perturbed to turn left.

The mean final 2 seconds of speed and heading for each trial was computed to compare model performance to participant data. Results of participants and simulations broken down by distance to the near arc and the distance to the perturbed arc can be found in Figure 3.8. The shaded regions are the 95% confidence intervals for participant means. In simulations of Experiment 4, both models perform poorly, however in different condition sets. The visual model captures the participant data well when the near arc is also the perturbed arc, regardless of distance to the near arc (Figure 3.8, the first points on the green dashed (2 m), blue dashed (4 m) and red dashed (6 m) curves). However, in conditions where the mid or far arc was perturbed (second and third points on the dashed curves), the visual model overperforms and does not match the participant data. In the case of the behavioral model, it performs well when the near arc is at 2 m (green dotted curve),

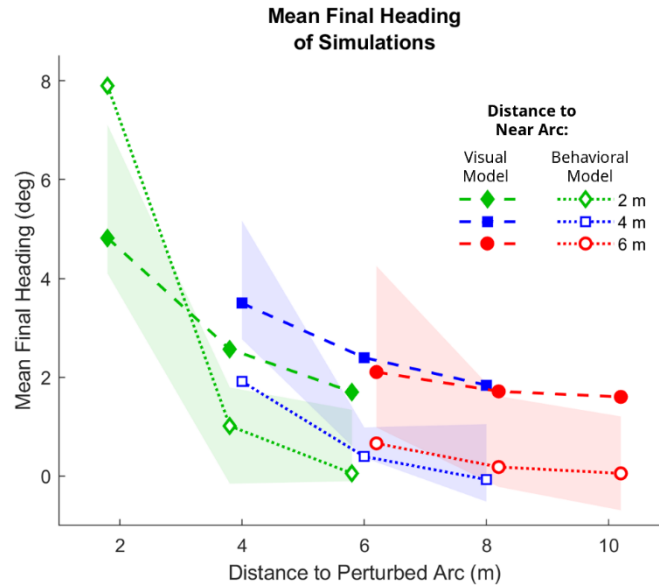


Figure 3.8. Mean final heading of the last two seconds of participant data, and simulations using the visual and behavioral models, taken from Experiment 4. The shaded regions are the 95% confidence interval of participant means. Note that neither model captures participant behavior.

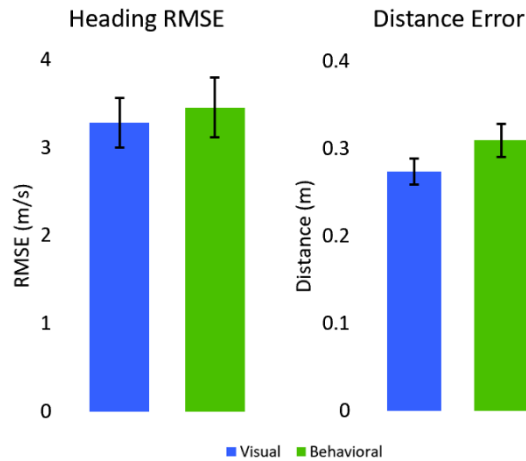


Figure 3.9. Comparison of error measurements between the participant and each model for Experiment 4, demonstrating the visual model did not capture participant data better than the behavioral model.

likely because these neighbors fall within its established distance neighborhood. However, when the near arc is at 4 m (blue dotted curve) or 6 m (red dotted curve), the behavioral model underperforms in capturing participant response, likely because they fall outside the model’s maximal influence distance of 5 m.

Taking the RMS error between the participant’s and each model’s heading for the last two seconds across all trials finds no significant difference between models ( $t(18) = 1.24, p = 0.23$ ; Figure 3.9). Bayesian model comparison yielded a scaled JZS Bayes Factor of  $BF_{10} = 1.48$  for heading, indicating anecdotal evidence for the behavioral model. This tests suggest that the visual model is not capturing human behavior any better than the behavioral model.

### 3.9 Experiment 4 - Discussion

The failure of both models to simulate the human data from Experiment 4, and the differences in the simulation results, clarify why the visual model is not performing well



at simulating crowds. The visual model overshoots the human data when the perturbed arc is beyond the near arc (the mid and far points of the dashed arcs, Figure 3.8). This is caused by an overweighting of farther neighbors. In behavioral results, it is shown that participants are only weakly influenced by neighbors that are within the crowd (beyond the near arc). This suggests that the effect of distance is not just driven by Euclid's law, but some other effect that causes influence to drop within the crowd. Without an additional component to explain this effect, the visual model will not be successful simulate the behavioral results of following in a crowd, as observed in Experiments 3 and 4.

Occlusion is a likely candidate for this component, as it could explain why pedestrians are less influenced by farther neighbors than the visual model along would suggest.

### **3.10 Conclusion**

Chapter 3 attempted to simulate following a crowd using the visual model. Unfortunately, as demonstrated by Experiments 3 and 4, the model does not perform as well as the behavioral model. However, the results from Experiment 4 demonstrate that the effect of distance is gradual to the nearest neighbors and rapid within the crowd beyond. Partial occlusion of far neighbors may explain this observation. Chapter 4 will investigate the role of occlusion in crowds and add it to the visual model, to better simulate human crowd behavior.

## CHAPTER 4

### Visual Occlusion in Crowds

## 4.1 Introduction

Chapter 3 found that following a crowd cannot be successfully simulated using neither the visual model nor the behavioral model. The visual model performed poorly on simulations of empirical data from Experiments 3 and 4, and while the behavioral model reproduced Experiment 3, it also failed at Experiment 4. However, these findings suggest an important principle: the visual model failed because it over-weighted more distant neighbors, neglecting the possible effect of visual occlusion.

Experiment 4 found that influence decays gradually to the nearest neighbor in a crowd and that rapidly within the crowd itself. This work described a doughnut-shaped neighborhood in a crowd having a “double decay” rate with distance (Warren, 2018): gradual to the nearest neighbor and rapid within the crowd itself. This form of double decay cannot be captured by the visual model as implemented thus far. As found in the simulations of Experiment 4, the visual model successfully captured the results when the near arc was perturbed (the first, highest points of each of the three curves in Figure 3.8), regardless of the distance to the near arc. This gradual decay to the nearest neighbor can be explained by Euclid’s law of perspective: as a neighbor’s distance increases, their optical expansion and angular velocity decreases as an arctangent function, thereby reducing their influence on a follower (as described in Figure 2.3). However, this does not explain the more rapid decay of influence within the crowd beyond the nearest neighbor, as observed in Experiment 4, when the mid and far arcs were perturbed.

A possible explanation is the visual occlusion of far neighbors. As the distance to neighbors beyond the nearest neighbor increases, the likelihood of near neighbors visually occluding far neighbors increases for two reasons. First, Euclid’s law applies not just to

optical expansion and angular velocity, but to occlusion too. As the metric distance between a near neighbor and far neighbors increases, more neighbors are likely to be occluded, because the area of occlusion (the near neighbor's "shadow") increases with distance. Putting it the other way, the visibility of far neighbors decreases with their metric distance from near neighbors. Second, Second, as the topological distance of far neighbors increases, there are potentially more near neighbors who might occlude them. In other words, the number of potential occluders increases linearly with ordinal distance ("index"). This is a topological effect, but it combines with the metric effect of Euclid's law.

Thus, the rapid decay within a crowd may be the combined effect of Euclid's law and topological occlusion. Incorporating occlusion into the visual model might account for the results of both Experiment 3 and Experiment 4. Critically, if true, this suggests that occlusion violates the normalized superposition principle for following in crowds, as the interactions between neighbor positions must be considered and are no longer independent from one another.

Chapter 4 will empirically test how visual occlusion factors into following, implement it within the visual model, and re-simulate the results of Experiments 3 and 4.

#### **4.2 Experiment 5 - Effect of occlusion in the crowd**

Experiment 5 empirically tested how visual occlusion of neighbors in a virtual crowd influenced the heading responses of the participant. This was accomplished manipulating the initial positions of far neighbors, so they were either fully occluded by near neighbors, or fully visible, at the beginning of a trial. Degree of occlusion could vary

during trials, as participants or neighbors moved in the virtual environment. As such, the degree of occlusion measured as was participant behavior.

### **Participants**

12 participants were recruited at Brown University, 7 female and 5 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

### **Apparatus**

The experimental apparatus was identical to Experiment 4, with the exception that the head-mounted display was a Samsung Odyssey (1440 x 1600 pixels per eye, 110° field of view, 90 Hz frame rate, 100% binocular overlap) and participant position and rotation was recorded at 90 Hz. The Odyssey has inside-out position and rotation tracking, no longer requiring external systems for this tracking.

### **Procedure**

Participants were instructed to walk with a crowd of 8 virtual human avatars as naturally as possible, to treat the virtual humans as if they were real people, and to “walk with the crowd as if walking down the street with them”. On each trial, the participant walked to the start pole and faced the orienting pole. After 2 seconds, the poles disappeared, and the virtual crowd appeared; a verbal command (“Begin”) was played over headphones and the virtual crowd began walking. The display continued until the participant had walked about 12 m; and a verbal command (“End”) signaled the end of the trial.

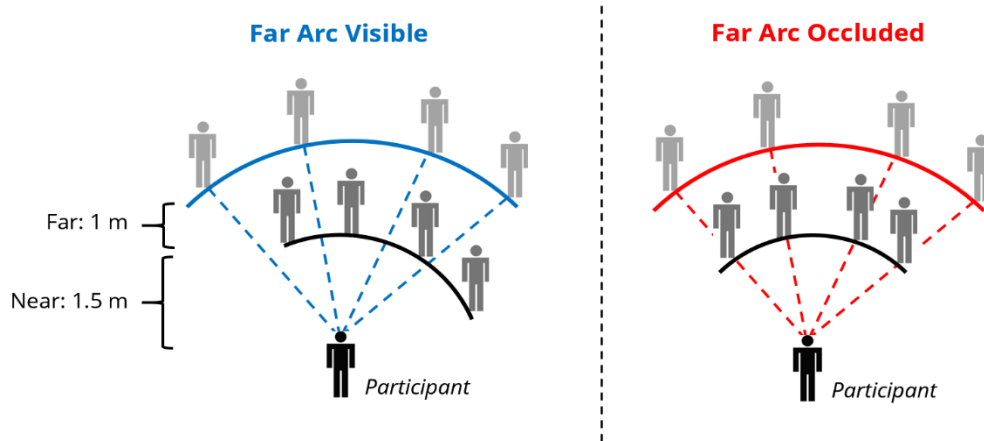


Figure 4.1. Diagram of Experiment 5, showing the far arc visible vs occluded.

## Displays

The 8 virtual neighbors appeared in two arcs: a near arc and a far arc, with 4 neighbors in each arc (near: 1.5 m distant, far: 2.5 m distant; jittering their initial positions by a uniform distribution of  $\pm 0.2$  m). A perturbed subset (0, 2, 4, 6 neighbors) was randomly chosen from one of the two arcs, and those neighbors turned ( $\pm 10^\circ$ , leftward / rightward) after a random interval between 3-4 seconds after trial start. Neighbors moved at a speed of 1.0 m/s in all trials. Because there were 4 neighbors in the near arc and 4 in the far arc, when the subset size was 6, the extra 2 neighbors were randomly selected from the other arc. The initial positions of the neighbors in the far arc were selected so they were either fully occluded by the near arc neighbors, or fully visible, at trial onset (Figure 4.1).

## Design

Experiment 5 had a  $4 \times 3 \times 2 \times 2$  factorial design: 4 subset sizes (0, 2, 4, 6)  $\times$  3 perturbation directions (left, no change, right)  $\times$  2 perturbed row (near, far)  $\times$  2 occlusion positioning (occluded, visible) for a total of 48 conditions. There were 3 repetitions per condition, for a total of 144 trials per participant.

## Data processing

Data was processed using the same procedures as in Experiment 4.

### 4.3 Experiment 5 - Results

While the initial positions of the virtual neighbors in the far arc were controlled to be visible or occluded to the follower, it is possible the proportion of their occlusion was dynamic during a trial. This is because the proportion of occlusion of a given far neighbor is dependent on the motions of that neighbor, closer neighbors, or the participant's own motion. For example, it is possible for the participant to adjust their position in relation to the crowd as to reduce or eliminate any occlusion between themselves and far neighbors. To measure the actual degree of occlusion during each trial, the mean proportion of occlusion across all far arc neighbors by all near arc neighbors was computed over time in the visible and occluded conditions (Figure 4.2). To calculate this, on each time step, two lines of sight are calculated from the participant to the silhouette (rim of the bounding

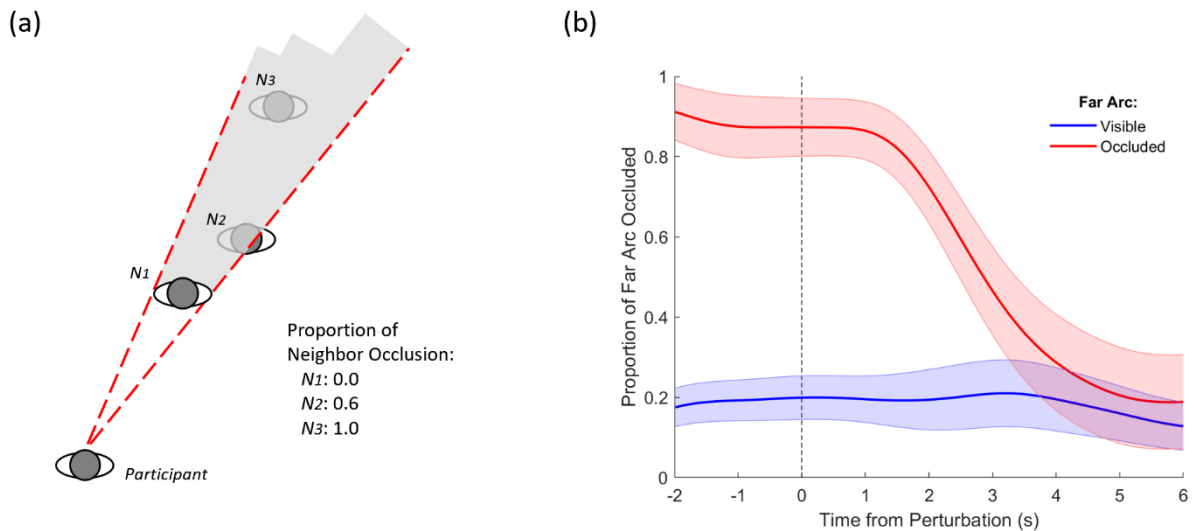


Figure 4.2. Proportion of far arc occluded by the near arc for Experiment 5. Panel (a) shows how the proportion of occlusion is calculated. Panel (b) the mean proportion of occlusion across all far arc neighbors by all near arc neighbors, computed over time for the visible and occluded conditions.

contour) of the nearest neighbor, continuing in depth (red dashed lines on Figure 4.2, panel a). This creates a ‘shadow’ that describes the region in which a near neighbor can occlude any more distant neighbors. For each far neighbor, occlusion is calculated based on the proportion of the farther neighbor's visual angle that lies outside the near neighbor shadow. Any neighbor that falls completely within this shadow is fully occluded and their proportion of occlusion is 1. A neighbor that lies completely outside of the shadow is fully visible, and their proportion is 0. If they are partially occluded, their occlusion is calculated as the proportion of their visual angle (as calculated by Equation 1) that lies outside the shadow. This occlusion shadow is then calculated for the next nearest neighbor and so on, until all neighbors have been accounted for.

Panel (b) of Figure 4.2 plots the results for the far arc's proportion of occlusion calculations for Experiment 5. Before perturbation, the far arc remained occluded in the occluded condition (demonstrating that participants did not deliberately move to a position made the far arc more visible). As neighbors in the occluded condition began to turn (post perturbation), they became visible when they moved out from behind the near arc. In the visible condition, the far arc's occlusion level remained constant and low across all trials. The goal of creating a difference in the proportion of occlusion between conditions was met.

Representative trials from Experiment 5 appear in Figure 4.3. These plot the position, speed, and heading for the participant and eight virtual neighbors, in which 4 neighbors in the near arc turns left, with one trial where the far neighbors are visible and



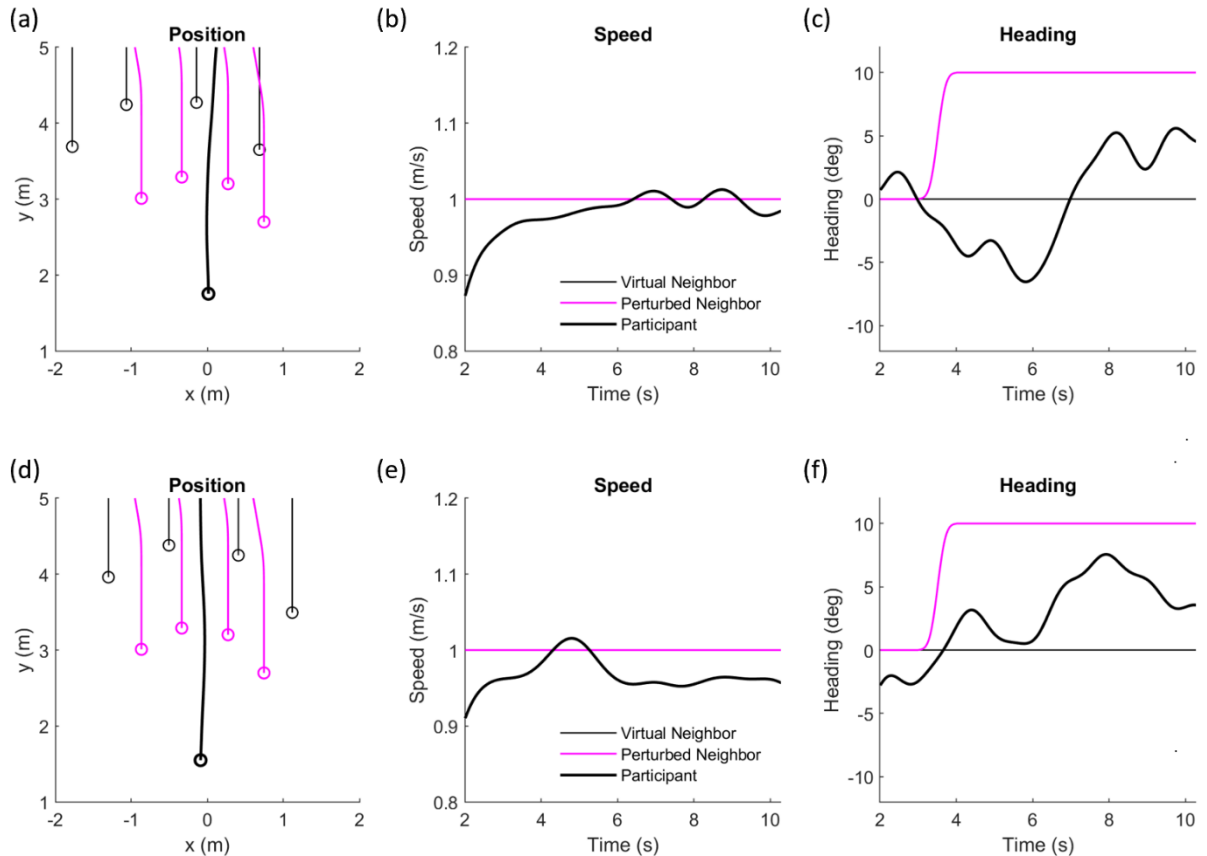


Figure 4.3. Representative trials from Experiment 5, plotting the time series of position, speed, and heading over time for the participants and the virtual neighbors. Both demonstrate 4 near neighbors turning left. Panels (a-c) show a far arc visible condition, while panels (d-f) shows a far arc occluded condition. The position plot is zoomed in to exhibit the differences between neighbor positions by condition.

another where they are occluded. Figure 4.3, panels (a) & (d), present two conditions where the far arc was visible or occlude. In the heading time series (Figure 4.3, panels c & f), the effect of occlusion can be seen. When the far neighbors were visible (Figure 4.3, panels a-c), the participant's response to the perturbation was reduced compared to the occluded condition (Figure 4.3, panels d-f). This occurs because in the occluded condition, the participant is only influenced by the four perturbed near neighbors (as the four far neighbors are occluded), while in the visible condition, the participant is also influenced by the four unperturbed far neighbors. Because the participant's response is presumably controlled by averaging the optical variables of all neighbors, the response is greater when

averaging the four near neighbors in the occluded condition than all eight neighbors in the visible condition.

The mean time series of heading in the visible and occluded conditions appear in Figure 4.4, when the near arc was perturbed (panel a), and the far arc was perturbed (panel b). A repeated measures ANOVA comparing the mean participant headings across 2-5 seconds after perturbation (shaded yellow region) found a significant difference between near and far arc perturbation ( $F(1, 44) = 249.79, p < 0.001$ ), as well as far arc occlusion ( $F(1, 44) = 11.15, p < 0.002$ ). Critically, there was also an interaction effect between the two ( $F(1, 44) = 92.07, p < 0.001$ ). This demonstrates that the direction of this occlusion

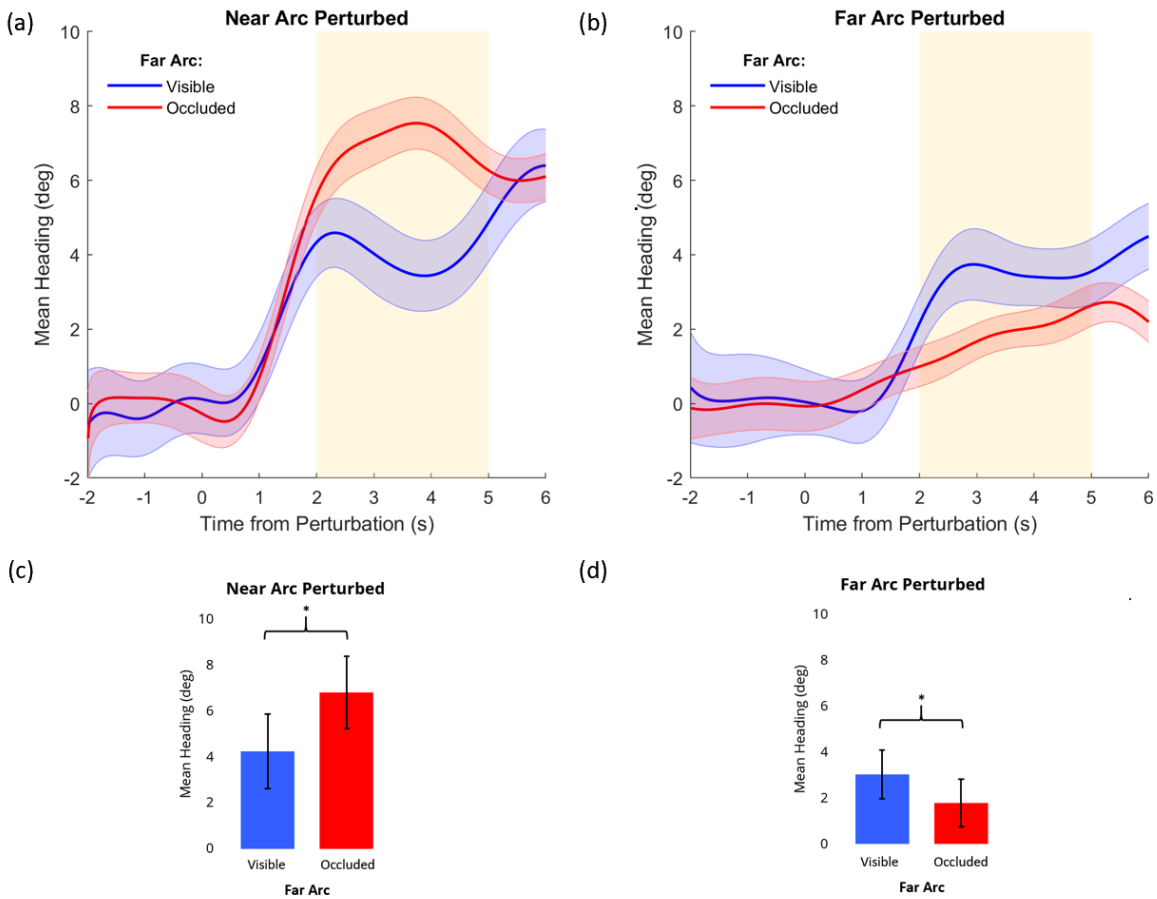


Figure 4.4. Panels (a-b) are the mean participant headings from Experiment 5, over time from perturbation, when the far row is visible or occluded by arc perturbed. Panels (c-d) are the mean headings taken across the shaded regions in panels (a-b) for comparisons.

effect reversed depending on whether the near or far arcs were perturbed. When the near arc was perturbed (Figure 4.4, panel a), there was a larger heading response if the far arc was occluded (red curve), because the participant was primarily influenced by the near arc, which turned (planned comparison;  $t(11) = -7.30, p < 0.001$ ). When the far arc was perturbed (Figure 4.4, panel b), the opposite occurred: the response was weaker if the far arc was occluded (red curve), because the participant was primarily influenced by the near arc, which continued walking straight (planned comparison;  $t(11) = 5.62, p < 0.001$ ). However, by five seconds after perturbation, the perturbed far neighbors emerged from behind the near neighbors, and the heading response converged with the visible condition (blue curves, panels a-b).

#### **4.4 Experiment 5 - Discussion**

Experiment 5 demonstrated two important aspects of occlusion for a pedestrian embedded in a crowd while following. First, occlusion does play a role in the visual control of heading, over and above the previously observed decay with distance. Although the present experiment replicated a distance effect in the visible condition (compare blue curves in panels (a) and (b) of Figure 4.4), there was an additional effect of occluding far neighbors. This can be observed in the reversal of the occlusion effect in Figure 4.4. The role of occlusion potentially explains the doughnut decay hypothesis: decay with distance to the nearest neighbor is due to Euclid's law, while the faster decay within the crowd is due to the added effect of occlusion for farther neighbors in the crowd (which correlates with distance). Second, Experiment 5 also demonstrated that occlusion's role in following a crowd is dynamic. While the far neighbors were initially occluded, as they turned (and

the participants responded) they became more and more visible, causing the participant's response in the visible and occluded conditions to converge. Or to put it another way: the far neighbors in the occluded condition dynamically became visible.

Chapter 3 established that the visual model was insufficient to simulate the influence of multiple neighbors in a crowd. Experiment 5 reveals what was missing: visual occlusion. Having demonstrated that occlusion does play a role in following a crowd, I now turn to developing a model of dynamic occlusion in a crowd.

#### **4.5 Modeling occlusion**

Given the results of Experiment 5, the influence of neighbors in a crowd appears to decay gradually to the nearest neighbor due to Euclid's law, and more rapidly within the crowd due to visual occlusion. The visual model captures the first effect; in order to simulate the second effect, a component of occlusion was added to the visual model. Figure 4.5 provides a simple example of how this component works in the horizontal for 3 timesteps. On each time step, two lines of sight are calculated from the model agent to the silhouette (rim of the bounding contour) of the nearest neighbor, continuing on in depth (red dashed lines on Figure 4.5). This creates a simulated 'shadow' that describes the region in which a near neighbor can occlude a far neighbor. For each far neighbor, a "visibility" weight that ranges from 0 to 1 is calculated based on the proportion of that neighbor's visual angle that lies outside the shadow. Any neighbor that falls completely within this shadow is fully occluded and their weight is set to 0. A neighbor that lies completely outside of the shadow is fully visible, and their weight is 1. If they are partially occluded, their weight is the proportion of their visual angle that lies outside the shadow; any value less than 0.15 is

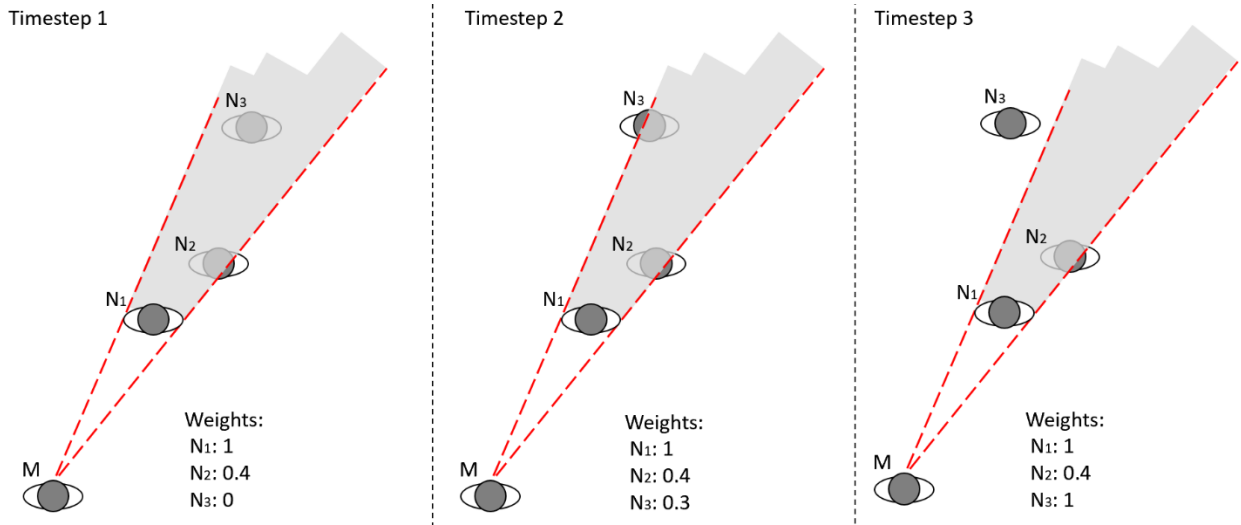


Figure 4.5. Simple example of how occlusion is simulated. On timestep 1, two straight lines (red dashed) are calculated from the position of the model (M), toward points laying on the two outward edges of the nearest neighbor (N1) and then those lines are continued out toward infinity. The area encompassed by the lines beyond the nearest neighbor (the grey shaded region) is the shadow of occlusion. The proportional weight applied to every neighbor depends on how much they are covered by the shadow. On timestep 1, N1 has a weight of 1, as it cannot fall in its own shadow. N2 has a weight of 0.4, as it is 40% visible and 60% occluded by N1. N3 has a weight of 0, as it is fully occluded by N1. After the occlusion shadow for N1 is calculated, similar shadows would be calculated for N2 and N3 (not shown). On timestep 2, N3 begins to move relative to M, becoming increasingly visible from behind N1, increasing the weight to 0.3. By timestep 3, N3 is fully visible and its weight is 1. These weights modify the influence of neighbor optical expansion and angular velocity in the model.

assigned a weight of 0. This occlusion shadow is then calculated for the next nearest neighbor and so on, until all neighbors have been accounted for. Each neighbor's weight is equal to the proportion that remains visible, unoccluded by their nearer neighbors. The influence of those neighbors' optical expansion and angular velocity is then modified by the value of their visibility weight. Of note, this position-dependent interaction between neighbors, causes breaks the principle of normalized superposition when following a crowd, as neighbors are no longer influencing the follower independent from one another.

This weighting allows for near neighbors to fully or partially occlude far neighbors, in order to capture the rapid decline of influence within a crowd as found in Experiment 4, as well as the effect of dynamic occlusion, as found in Experiment 5. With visual occlusion

added to the visual model for following, I simulated Experiment 5 and re-simulated Experiments 3 and 4 to compare the previous model iteration (the visual model) to the new model iteration (visual + occlusion model).

#### 4.6 Experiment 5 - Model Simulations

To evaluate the performance of the models, the stimuli from Experiment 5 were inputted to the visual and visual + occlusion models and the output was compared to the mean time series of heading for the participants. Figure 4.6 shows the mean heading time series when the near arc was perturbed (panel a) and the far arc was perturbed (panel b). Blue curves are conditions in which the far arc was visible and red curves are conditions in which the far arc was occluded. Blue and red shaded areas represent the 95% confidence intervals about the mean human data. For the visual model (dashed curves), the difference between the visible and occluded conditions is negligible in both panels (a) and (b).

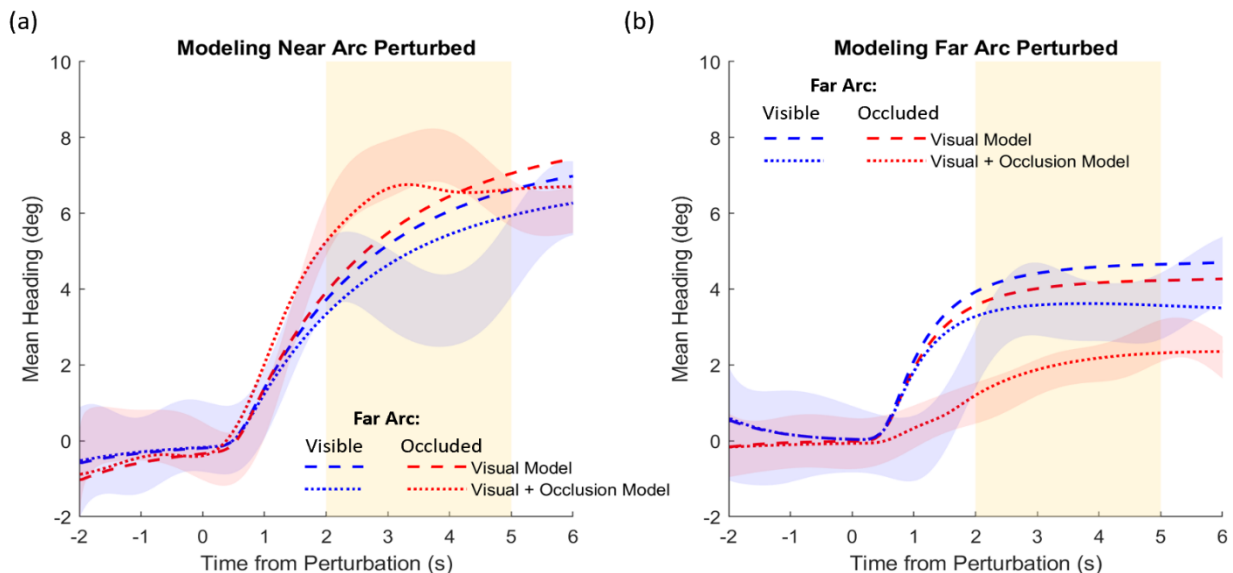


Figure 4.6. Modeling the mean heading from Experiment 5, over time from perturbation, when the far arc was visible or occluded. The blue and red shaded areas are the 95% confidence interval of participant data, and the yellow shaded regions are where comparison means were taken. Dashed curves plot the visual model, dotted curves are the visual + occlusion model.

Without an occlusion component, the model is influenced by both near and far arcs in all conditions, and so the simulations are nearly equivalent. In contrast, with the addition of occlusion, the visual + occlusion model (dotted curves) distinguishes the visible and occluded conditions. When the near arc was perturbed (panel a), the occlusion of the unperturbed far arc caused the simulation to generate a larger turn in the five seconds after perturbation (red dotted curve). When the far arc was perturbed (panel b), the opposite effect was seen. With the far neighbors occluded, the model was primarily influenced by the unperturbed near neighbors, and thus generated a smaller turn (red dotted curve).

To compare the visual and the visual + occlusion models (Figure 4.7), the RMSE between each model and participant heading time series was computed between 2-5 seconds after perturbation on each trial (yellow shaded region, Figure 4.6). When the near arc was perturbed, there was no significant difference between models when the far arc was visible ( $t(11) = 1.74, p < 0.11$ ), but there was a significant difference when the far arc

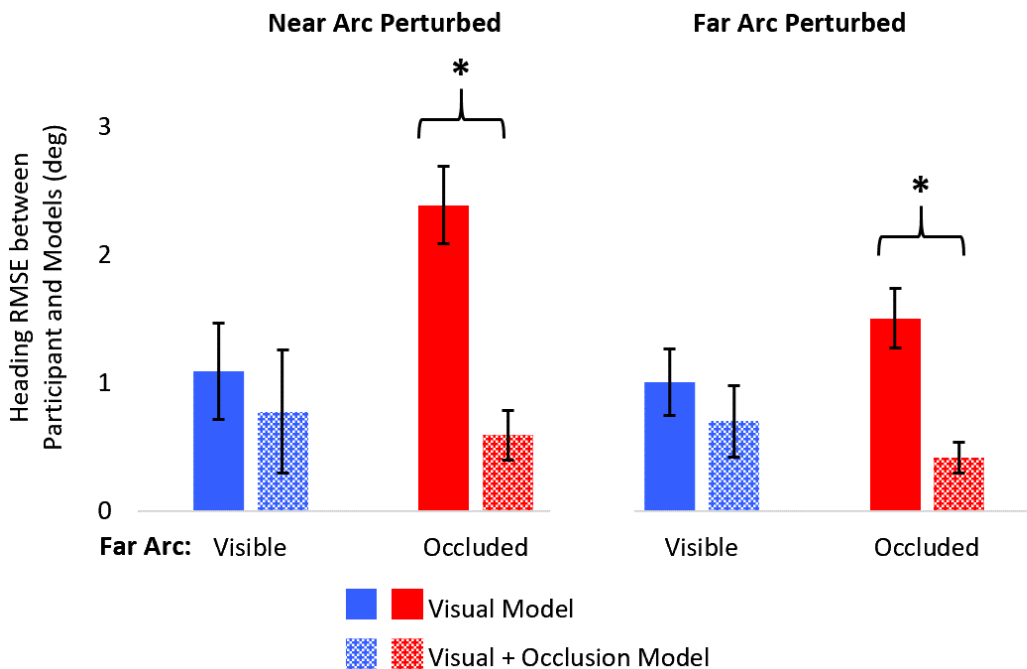


Figure 4.7. Comparisons of RMSE for heading between participant and the visual and visual + occlusion models across conditions for Experiment 5.

was occluded ( $t(11) = 17.33, p < 0.001$ ). When the far arc was perturbed, there was again no significant difference between models when the far arc was visible ( $t(11) = 1.91, p < 0.08$ ), but as with the near arc being perturbed, there was a significant difference between models when the far arc was occluded ( $t(11) = 8.72, p < 0.001$ ). Bayesian model comparison yielded a scaled JZS Bayes Factor of  $BF_{10} = 123190822646$  for the near arc perturbed while the far arc is occluded, and a scaled JZS Bayes Factor of  $BF_{10} = 552091$  for the far arc perturbed while also occluded, each indicating decisive evidence for support for the visual + occlusion model. Taken together, these results demonstrate a critical property of the visual + occlusion model: when occlusion in a crowd is high, a component for occlusion is necessary to simulate following in a crowd, and that it serves as a strong candidate to model previously examined datasets.

#### **4.7 Re-simulating Experiments 3 and 4**

Experiment 5 established that occlusion plays a strong role in following multiple neighbors in a crowd. Adding an occlusion component to the visual model significantly improved simulations of the human data. This finding also demonstrated that occlusion is dynamic and has transient influences on the participant. In this section, I test whether the visual + occlusion model generalizes to Experiments 3 and 4, where dynamic occlusion also occurred, and the visual model performed poorly. All trials in both experiments were simulated with both the visual and visual + occlusion models, and mean RMSEs for the two models were compared.

Figure 4.8 re-plots the mean final speed (panel a) and mean final heading (panel b) of simulations of the visual and visual + occlusion models from Experiment 3. As



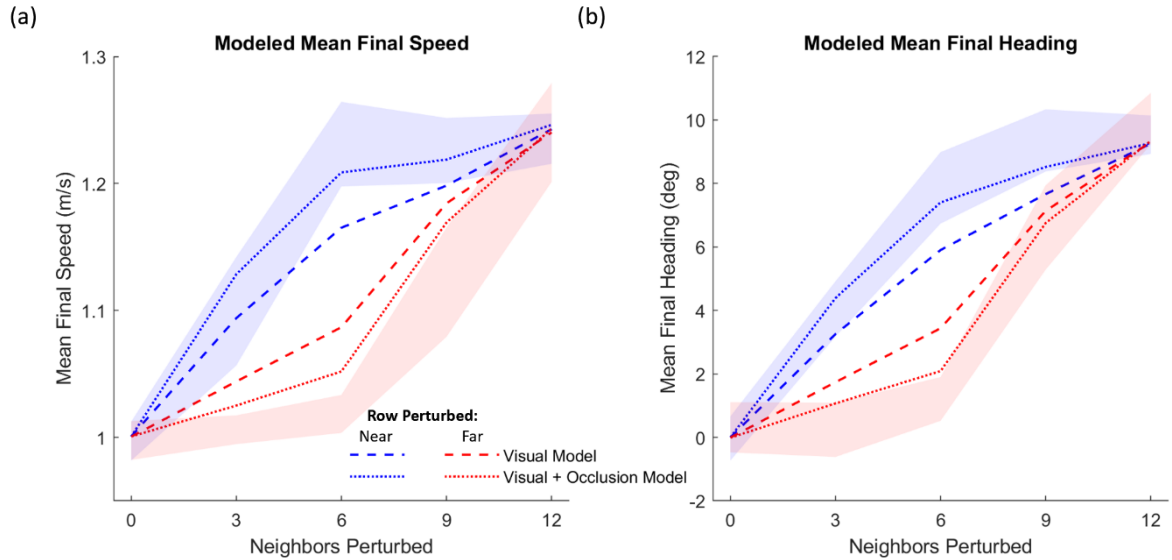


Figure 4.8. Mean final speed (panel a) and heading (panel b) of the last two seconds of participant data from Experiment 3, re-simulated using the visual and visual + occlusion models. The shaded regions are the 95% confidence interval of participant data, which the visual + occlusion model better captures.

previously observed in Figure 3.3, the visual model undershoots the data when the near row is perturbed (blue curves), and overshoots the data when the far row is perturbed (red curves). However, when occlusion is added to the visual model, results closely simulate the data. In the virtual crowd displays, the far neighbors were occluded by the near neighbors on numerous occasions, during which time their influence was reduced, and the participant was primarily influenced by near neighbors. A hypothesis raised in Chapter 3 was that far neighbors fell outside a metric neighborhood, which the visual model failed to capture because it had an unlimited interaction distance. Here it is demonstrated that the visual + occlusion model performs just as well as the behavioral model (found in Figure 3.3).

Tests comparing participant final means to models across all trials is found in Figure 4.9. The results for speed find the visual + occlusion model produces significantly less RMS error compared to the visual model ( $t(18) = -2.639, p = 0.017$ ), and no significant difference in RMS error compared to the behavioral model ( $t(18) = -0.311, p = 0.759$ ). A

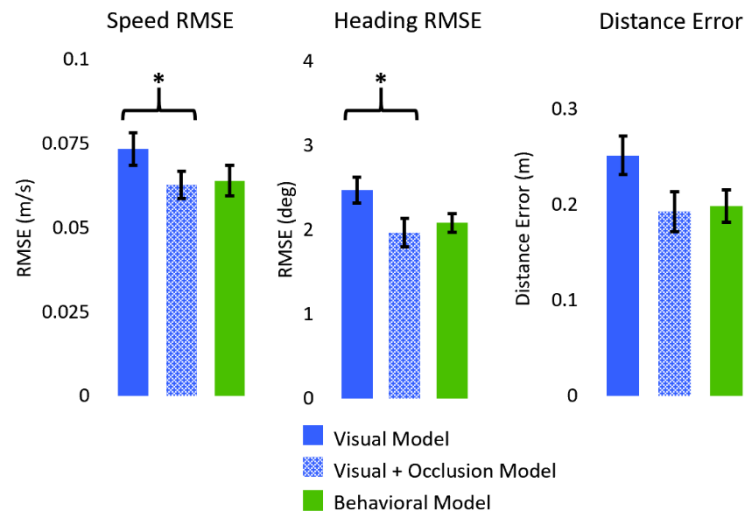


Figure 4.9. Comparisons of RMSE for speed and heading between participant and the visual and visual + occlusion models across conditions for Experiment 3.

Bayesian model comparison for speed between the visual + occlusion model and the visual model yielded a scaled JZS Bayes Factor of  $BF_{10} = 3.60$ , indicating substantial evidence for the visual + occlusion model, and a comparison between the visual + occlusion model and the behavioral model yielded a scaled JZS Bayes Factor of  $BF_{10} = 2.43$ , indicating anecdotal evidence favoring the visual + occlusion model when simulating speed in Experiment 3.

The results for heading find the visual + occlusion model produces significantly less RMS error compared to the visual model ( $t(18) = -3.455, p = 0.003$ ), and no significant difference in RMS error compared to the behavioral model ( $t(18) = -0.891, p = 0.384$ ). A Bayesian model comparison for heading between the visual + occlusion model and the visual model yielded a scaled JZS Bayes Factor of  $BF_{10} = 13.33$ , indicating strong evidence for the visual + occlusion model, and a comparison between the visual + occlusion model and the behavioral model yielded a scaled JZS Bayes Factor of  $BF_{10} = 1.90$ , indicating anecdotal evidence for the visual + occlusion model when simulating heading in Experiment 3.

Taken together, these results show that the visual + occlusion model better captures the results of Experiment 3 compared to the visual model, and equally as well as the behavioral model.

Experiment 4 found neighbor influence decays gradually to the nearest neighbor and then rapidly within the crowd. Figure 4.10 re-plots the mean final heading of simulations of the visual and visual + occlusion models from Experiment 4, to determine whether occlusion explains this pattern of distance decay. As previously observed in Figure 3.6, visual model did not accurately capture the human data. However, the visual + occlusion model is added to the model closely reproduces the data (dotted curves in Figure 4.10). Each mean heading captures the double-decay with distance: gradual to the nearest neighbor and rapid within the crowd itself.

Tests comparing participant final mean heading to models across all trials is found in Figure 4.11. The results find the visual + occlusion model produces significantly less

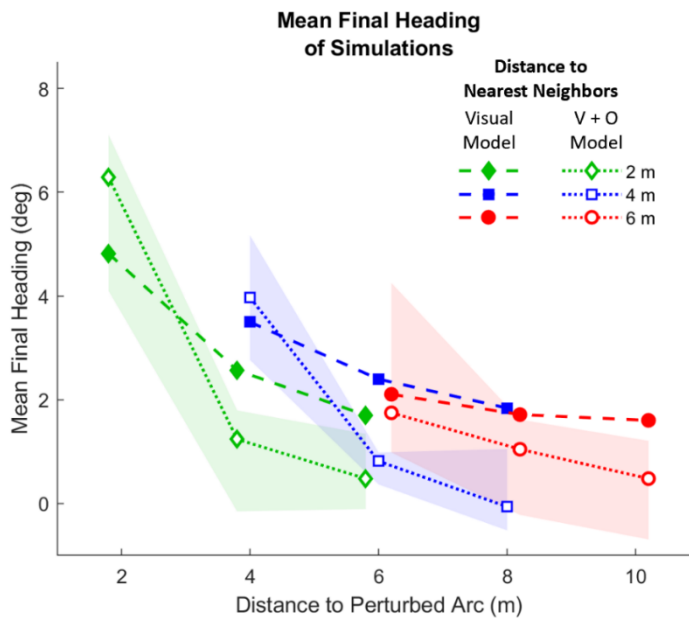


Figure 4.10. Mean final heading of the last two seconds of participant data, and simulations using the visual and visual + occlusion models, taken from Experiment 4. The shaded regions are the 95% confidence interval of participant means.

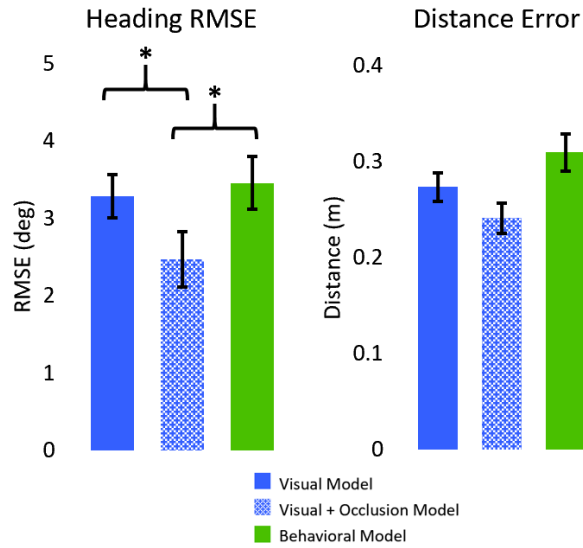


Figure 4.11. Comparisons of RMSE for heading between participant and the visual and visual + occlusion models across conditions for Experiment 4.

RMS error compared to the visual model ( $t(22) = 7.66, p < 0.001$ ), and significantly less in RMS error compared to the behavioral model ( $t(22) = 7.18, p < 0.001$ ). A Bayesian model comparison for heading between the visual + occlusion model and the visual model yielded a scaled JZS Bayes Factor of  $BF_{10} = 75938$ , indicating decisive evidence for the visual + occlusion model, and a comparison between the visual + occlusion model and the behavioral model yielded a scaled JZS Bayes Factor of  $BF_{10} = 5055725$ , also indicating decisive evidence in favor of the visual + occlusion model when simulating heading in Experiment 4.

The results clearly demonstrate that the results from Experiment 4, that describe a double-decay of influence with distance, are driven in part by occlusion, and that the visual + occlusion model successfully captures this behavior.

## 4.8 Conclusion

The results from Chapter 3 indicated that the visual model did not account for a pedestrian following a crowd because it failed to capture the double-decay with distance. One hypothesized explanation for this double-decay pattern was occlusion of far neighbors by nearer neighbors. Experiment 5 demonstrated that occlusion does indeed affect participant responses to neighbors, and provided the groundwork to improve the visual model to account for this optical variable. Re-simulating the human data from Experiments 3 and 4 confirmed that accounting for occlusion is critical for modeling the influence of multiple neighbors and explains the doughnut-shaped neighborhood observed in Experiment 4.

The visual + occlusion model offers a robust model of the visual control of speed and heading for an individual agent following multiple neighbors. In Chapter 5, I will investigate the model in multi-agent simulations to test whether it also generates collective motion.

## CHAPTER 5

### Multi-agent simulations and collective motion

## **5.1 Introduction**

Chapter 4 demonstrated that adding a component for dynamic occlusion greatly improved the visual model at simulating human following a crowd. By including occlusion in the visual + occlusion model, it captured the previous observation of a doughnut-shaped neighborhood in a crowd: influence decays gradually to the nearest neighbor and rapidly within the crowd beyond. Chapter 5 will compare the visual model with the visual + occlusion model in crowd simulations. First, the model will be compared against observational data on crowds of real people, to test whether either version of the visual model can account for individual trajectories in real crowds. Second, we will test whether the models generate collective motion in a series of multi-agent simulations, and evaluate its robustness by varying initial conditions.

## **5.2 Experiment 6 - Modeling human ‘swarms’**

To test the models in a more natural setting, we simulated individual pedestrians in previously collected motion capture data on a human ‘swarm’ scenario. Groups of participants were recorded walking together for periods of 2 minutes, at various initial interpersonal distances (IPD). Segments of trials were simulated using the visual model and the visual + occlusion model to assess the role of occlusion in real crowds.

### **Participants**

Participants were recruited at Brown University. One group of 10 participants (4 female, 6 male), one group of 16 participants (6 female, 10 male) and one group of 20 participants (10 female, 10 male) were tested in separate blocks. None reported any visual or motor impairment. The research protocol was approved by Brown University’s

Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

### **Apparatus**

The groups were tested in a large hall (Brown University's Sayles Hall) with a 14x20 m tracking area marked on the floor with red tape. Each participant wore a bicycle helmet with a unique constellation of five reflective markers on 0.3-0.4 m stalks. Head position was recorded at 60 Hz with a 16-camera infrared motion capture system (Qualisys Oqus).

### **Procedure**

Participants were instructed to walk about the tracking area at a normal speed, veering randomly left and right, while staying together as a group. They began each trial in random positions in a 4x4 m (1 m IPD) or 7x7 m (2 m IPD) start box marked on the floor. At a verbal 'go' signal, participants started walking for 2 min, until a verbal 'stop' signal.

### **Design**

Experiment 6 had a 3x2 factorial design: 3 group sizes (10, 16, 20 people) x 2 initial interpersonal distances (1 m, 2 m), for a total of 6 conditions. There were 2 repetitions per condition, for a total of 12 trials.

### **Data processing**

Position and rotation for each participant was retrieved by finding the centroid of the five reflective markers on each participant's headgear at each timestep. The time series of head position in the horizontal plane were filtered using a 4th-order low-pass



Butterworth filter with a cutoff frequency of 0.6 Hz. This was to remove tracker noise and reduce oscillations due to the step cycle. The first two seconds of every trial were removed, so that the participants were in steady-state walking for analysis and modeling. The last second of every trial was removed, to avoid any errors due to filtering at the edges of recorded data (see Howarth & Callaghan, 2009). Speed and heading of each participant were computed from the filtered position data as the displacement between successive time steps. Due to tracking loss of some participants for periods of time during a trial, the data were divided into 10 second segments, in which at least 75% of participants were always recorded for that 10 seconds of time. This yielded 35 segments with good tracking data, which ensured that most of the input to a modeled pedestrian was known. At the beginning of a given segment, the participant furthest back in the group was the one chosen for simulation (the “focal” participant). Of these modeled segments, 5 additional segments were removed due to limitations of the visual and visual + occlusion models. When simulating segments where the front of the crowd turns in a 180° arc (so they are walking toward the focal participant in the rear) it produces a large expansion rate from the viewpoint of that focal participant. Because the model has no function for determining the direction neighbors face, this large expansion causes the model to begin to move backwards (motion opposite the direction they are facing). This does not match human behavior, and produces large error, and so these 5 segments were removed. This leaves 30 segments for analysis (17 segments from the 1 m initial IPD condition, and 13 segments from the 2 m initial IPD condition). RMS error was computed between the focal participant and each model for both speed and heading.

### 5.3 Experiment 6 - Results

Two representative segments from Experiment 6 are seen in Figure 5.1. These plot the position over time, proportion of neighbors occluded, times series of speed, and time series of heading for the participants in the swarm and the visual model with and without occlusion. Segment 1 (Figure 5.1, panels a-d) was a 10 second segment of a 2 m initial IPD trial, with 16 total participants and 13 recorded across the segment. Segment 2 (Figure 5.1, panels e-h) was a 10 second segment of a 1 m initial IPD trial, with 20 total participants and 17 recorded across the segment. Simulations of Segment 1 do not differ between versions of the model (RMSE for speed, visual: 0.20 m/s, visual + occlusion: 0.19 m/s; for heading, visual: 16.7°, visual + occlusion: 17.0°). However, the simulations do differ for

Segment 2 (RMSE for speed, visual: 0.26 m/s, visual + occlusion: 0.13 m/s; for heading, visual: 42.8°, visual + occlusion: 10.1°). In this case, including occlusion in the model

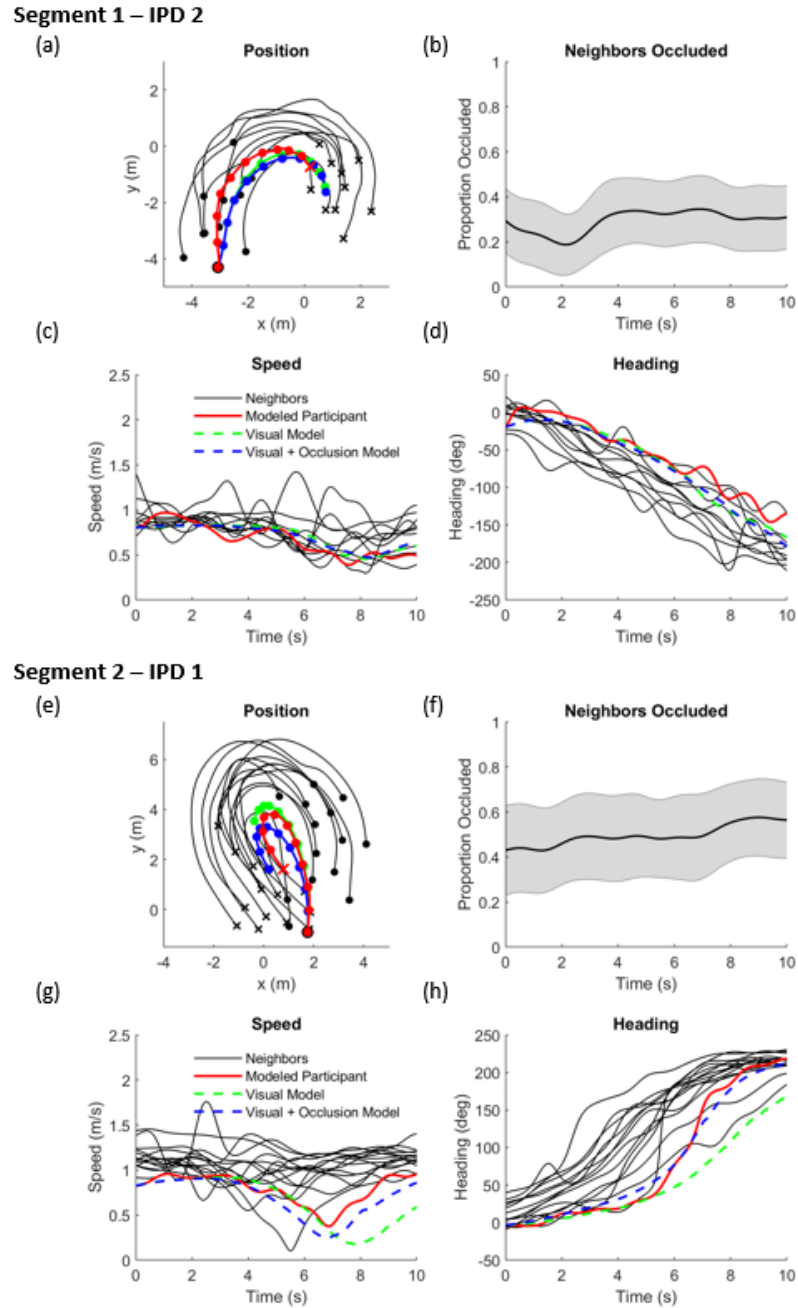


Figure 5.1. Representative trials from Experiment 6, plotting the time series of position, proportion of neighbors occluded, speed, and heading over time for the participants in the swarm, the visual model, and the visual + occlusion model. Segment 1 (panels a-d) is a 10 second segment of a 2 m IPD trial. Segment 2 (panels e-h) is a 10 second segment of a 1 m IPD trial. Comparing the proportion of neighbors occluded between trials (b, f) Segment 2 has more occlusion over the trial.

produced a simulation that more closely matched the focal participant’s locomotor behavior. Comparing the proportion of occlusion over all neighbors between trials (Figure 5.1, panels b & f) Segment 2 had more occlusion over the trial. When occlusion is high in the observed data, the visual model requires an occlusion component to capture the participant's trajectory.

Figure 5.2 compares the performance between the visual model and the visual + occlusion model for all 30 segments from Experiment 6. Across all measures, the visual + occlusion model performs better than the visual model. The results demonstrate that the visual + occlusion model generalizes to real-world data of swarming crowds. Tests comparing the participant each model’s over time shows there is no RMSE difference for speed, likely due to the lack of variation in participant speed across segments ( $t(58) = 0.76$ ,  $p = 0.452$ ), and significantly less RMSE for heading ( $t(58) = 2.81$ ,  $p = 0.007$ ) with the visual + occlusion model. A Bayesian model comparison for speed between the visual +

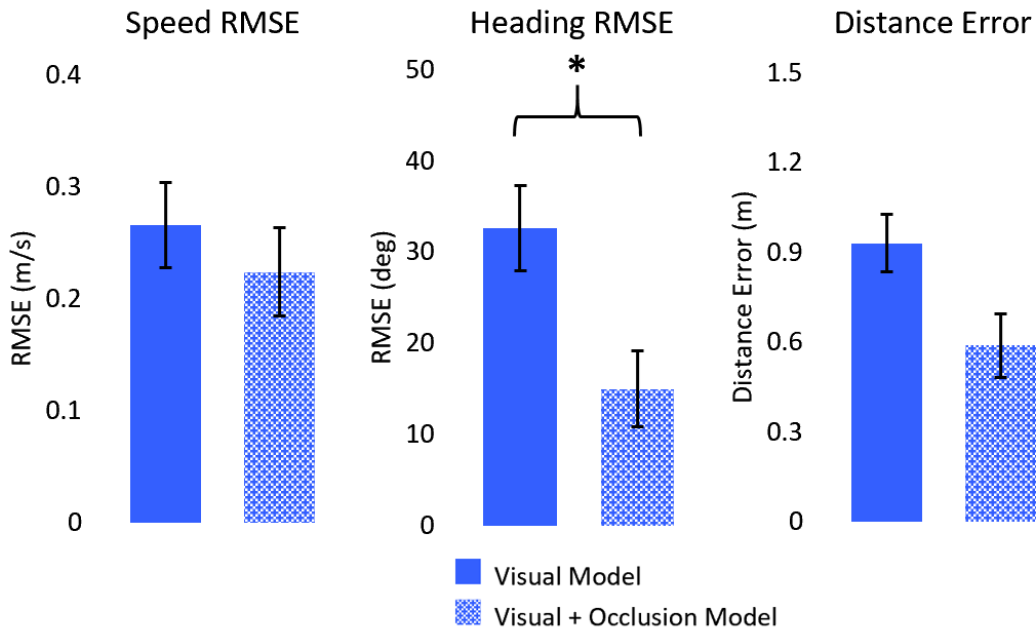


Figure 5.2. Comparison of error measurements for the swarm segments, between the participant and the model with and without occlusion: mean metric distance between participant position and model positions, RMSE for speed, and RMSE for heading.

occlusion model and the visual model yielded a scaled JZS Bayes Factor of  $BF_{10} = 2.98$ , indicating anecdotal evidence for the visual + occlusion model for modeling speed, and for heading yielded a scaled JZS Bayes Factor of  $BF_{10} = 6.45$ , also indicating substantial evidence for support of the visual + occlusion model when simulating heading in Experiment 6.

#### **5.4 Multi-agent simulations**

To this point, I have found that a model of the visual control laws for following a leader scale up to walking with multiple neighbors in a crowd, and can even predict individual trajectories in real crowds. The ultimate test of this bottom-up approach is whether the same "rules of engagement" yield collective motion, as often observed in pedestrian groups. The purpose of this section is to determine whether the visual model can generate self-organized collective motion in a set of mutually interacting agents, and to test the robustness of coherent motion under a range of initial conditions.

##### **Procedure**

A set of initial conditions (initial interpersonal distance, speed, and heading) was generated for 20 simulated agents (0.2 m radius circles) in a 4x5 position grid. These were used as an input to the visual and visual + occlusion models. Each agent's position was then jittered by a uniform distribution equal to  $\pm 25\%$  of that set's initial interpersonal distance. The movement of each agent and their interactions was governed by the models, with synchronous updating of all agents at each timestep. At the end of 30 seconds, the simulation was stopped, and the standard deviation of the final speed and heading was computed across all agents. This was repeated for 20 runs per initial condition set (with

only jitter varying between individual runs). The mean standard deviations of final speed and heading in each initial condition set characterized the degree of convergence for each model over time (the lower the standard deviation, the more the simulated crowd's speed or heading converged across 30 seconds). This performance measure was used to compare the robustness of the visual motion model, the visual + occlusion model, the behavioral model, and the double-decay model (Warren & Dachner, 2018), across a wide range of initial conditions.

## **Design**

Given that varying initial speed primarily influences final speed, and the same is true for heading, these two variables were manipulated and analyzed separately. The initial conditions for the multi-agent simulations had a 10x2x9 factorial design: 10 initial interpersonal distances (ranging from 1 to 10 m, in steps of 1 m) x motion variables (agents were generated with varied initial speed or initial heading) x 9 variable ranges (in increments of  $\pm 0.1$  m/s about 1 m/s, or  $\pm 10^\circ$  about  $0^\circ$ , respectively). Specifically, the mean initial mean speed of agents was 1.0 m/s. When speed was varied, each agent's initial speed was randomly selected from a uniform distribution ranging from 0.9 to 1.1 m/s in the first condition; from 0.8 to 1.2 m/s in the second condition; and so on up to a range of 0.1 to 1.9 m/s in the ninth condition, in increments of  $\pm 0.1$  m/s. The mean initial heading of agents was  $0^\circ$ . When heading was varied, each agent's initial heading was randomly selected from a uniform distribution ranging from  $-10^\circ$  to  $+10^\circ$  in the first condition; from  $-20^\circ$  to  $+20^\circ$  in the second condition; and so on up to a range of  $-90^\circ$  to  $+90^\circ$  in the ninth condition, in increments of  $\pm 10^\circ$ . Initial conditions were generated for each of 20 agents for 20 runs per condition.

## 5.5 Multi-agent Simulations - Speed Results

Three representative runs from the speed simulations appear in Figure 5.3, including paths to final positions (left column), time series of speed (middle column), and time series of heading (right column) for the generated agents in the swarm. The black curves are the visual model, the red curves are the visual + occlusion model, and the blue curve is the agent at the front of the simulated crowd. Because each agent had a FOV restricted to  $\pm 90^\circ$  from the direction of motion an agent was not influenced by any neighbors outside this FOV. For the agent at the front of the crowd, no other agents were in their FOV and so its speed and heading remained constant from its initial conditions. This caused other agents align with the agent at the front (blue curve), making it the defacto leader of the crowd. Occlusion of neighbors had no effect on the lead agent and thus each model produced identical results for that agent only.

In the first row (Figure 5.3, panels a-c), the initial interpersonal distance was 1 m and the initial speed range was  $\pm 0.1$  m/s. Because of the proximity of all neighbors and the low variation in initial speed, the simulations converged quickly and the differences between the models was negligible. Interestingly, position where the models diverged the most was at the rear of the crowd (see panel a). This is a demonstration of information transfer from the front of the crowd to the rear, and how occlusion weakens the influence of far neighbors, which slows information transfer because a pedestrian can't respond as early to far neighbors. In the second row (panels d-f) the initial interpersonal distance was

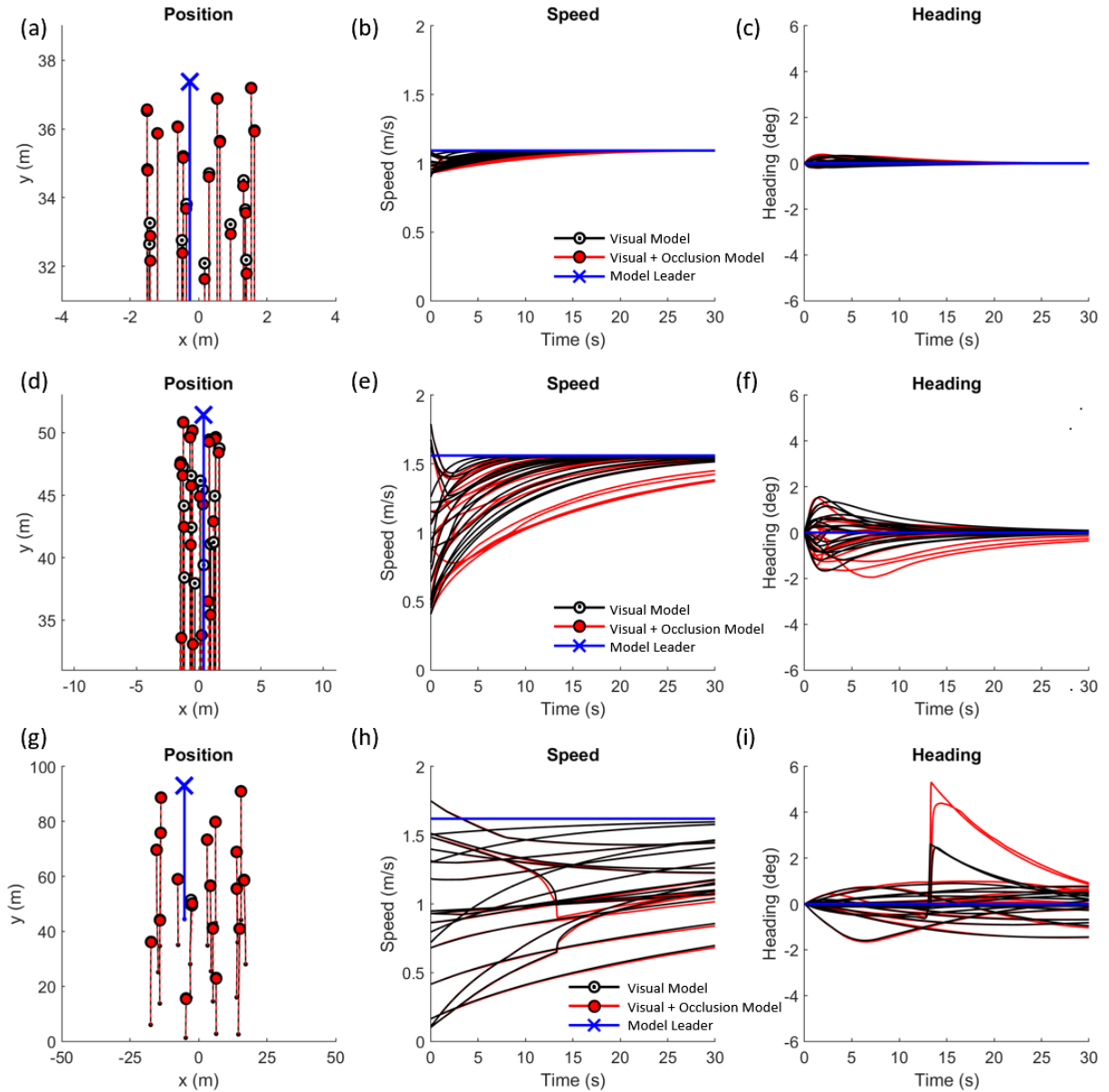


Figure 5.3. Representative runs from the multi-agent simulations for speed. The black curves are the visual model, the red curves are the visual + occlusion model, and the blue line is the agent at the front of the simulated crowd. The black and red ‘o’s are to compare final position across models. Panels (a-c) is an IPD of 1 m and  $\pm 0.1$  m/s variation in speed trial. Panels (d-f) is an IPD of 1 m and  $\pm 0.9$  m/s variation in speed trial. Panels (g-i) is an IPD of 10 m and  $\pm 0.9$  m/s variation in speed trial.

1 m and the speed range was  $\pm 0.9$  m/s. The high variation in initial speed caused convergence to take longer, but the proximity allowed it to still occur. Note that the visual + occlusion model took longer to cohere (red curves in panel e). Information transfer again flows from the front of the crowd to the rear, and is slowed by occlusion (panel d). In the



third row (panels g-i) the initial interpersonal distance was 10 m and the initial speed range was  $\pm 0.9$  m/s. The large distance between agents and high variation in initial speed prevented either model from converging within 30 seconds. However, because the distance between agents was so great, occlusion was low across the run, causing both models to produce nearly identical results. Note the spike in heading between 10-15 seconds. This is caused by an agent moving fast enough to move into another agent's FOV. For the visual model, eccentricity determines how the trade-off in optical expansion and angular velocity influence both speed and heading. By suddenly appearing in the FOV of another agent, it caused a spike in expansion that caused the agents to turn away from each other.

Heat maps of the standard deviation of final speed as a function of initial interpersonal distance by the range and the initial speed rate appear in Figure 5.4. The temperature (color) of each cell represents the standard deviation of the final speed across all agents generated, averaged across the set of 20 runs for each initial condition. The heat maps estimate the cohesion of the simulated swarm, where low standard deviations (warm colors) signify high speed cohesion and high standard deviations (cool colors) signify low speed cohesion. Figure 5.4 panel (a) is the heatmap of the speed cohesion for the behavioral model with its exponential decay with distance. Large initial interpersonal distances and initial speed ranges led to the lowest cohesion across all models. Panel (b) plots the heatmap for the behavioral model with the double decay with distance (Warren & Dachner, 2018). This flexible doughnut-shaped neighborhood extended the range of interaction, yielding increased cohesion at greater IPDs and speed ranges. Panel (c) presents the heatmap for the visual model (without occlusion). Its pattern closely resembles that of the doughnut model, but with more overall cohesion. Panel (d) plots the heatmap for the visual + occlusion

model. While generally cohesion is as high as the visual model, low values of initial interpersonal distance (1-3 m, which causes more occlusion) and high ranges of initial speed caused a notable decrease in cohesion. The patterns of these heatmaps will be discussed in more depth below.

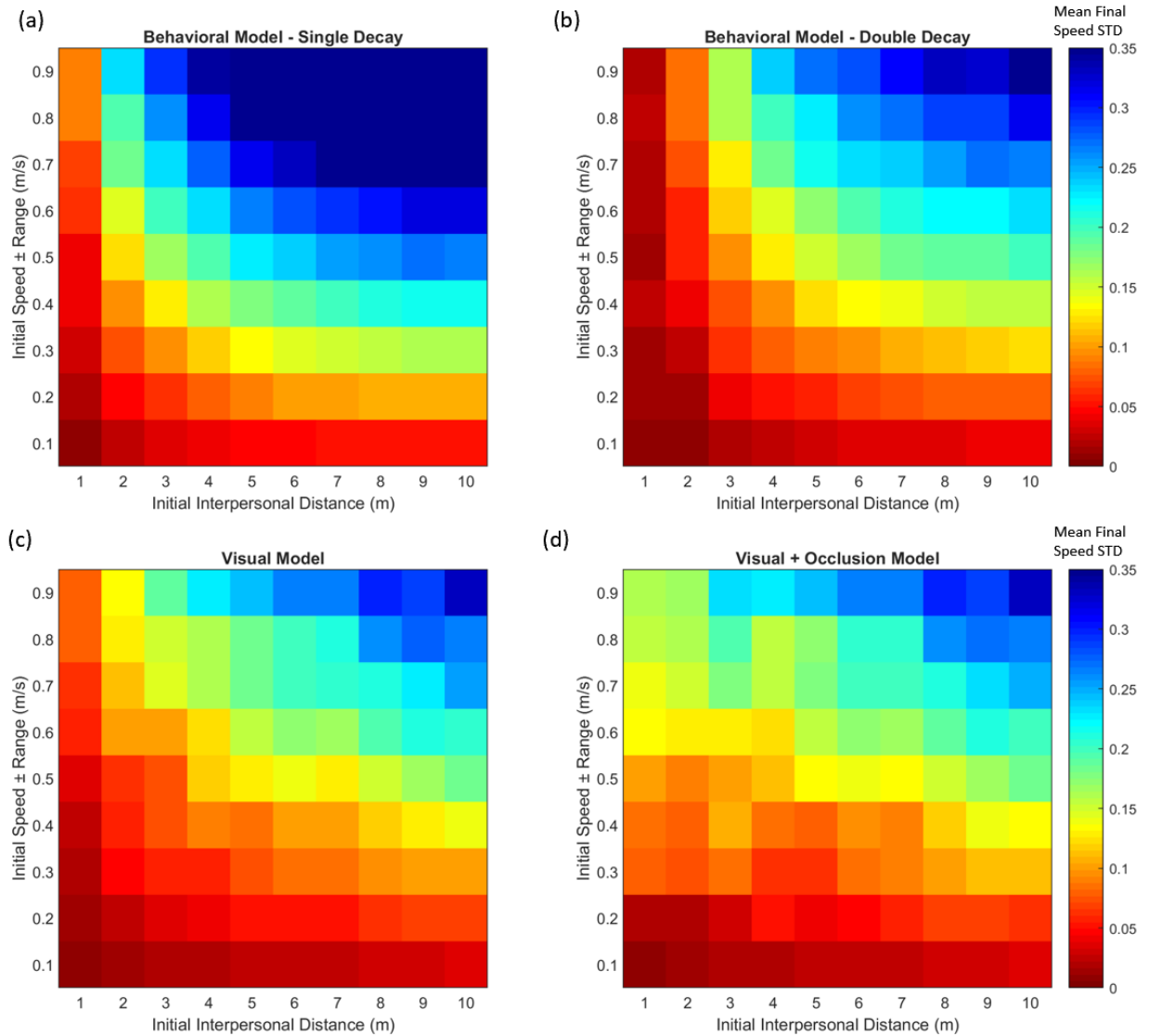


Figure 5.4. Heatmaps for the cohesion of speed across bulk simulations of crowds, plotting initial speed range by initial interpersonal distance, where each cell is the mean final standard deviation for speed across all generated agents across all trials in the set. Warm means that combination of initial conditions leads to higher cohesion (lower standard deviation) and cool means lower cohesion (higher standard deviation).

## 5.6 Multi-agent Simulations - Heading Results

Three representative trials from the heading simulations appear in Figure 5.5, including paths to final positions (left column), time series of speed (middle column), and time series of heading (right column) for the generated agents in the swarm. The black

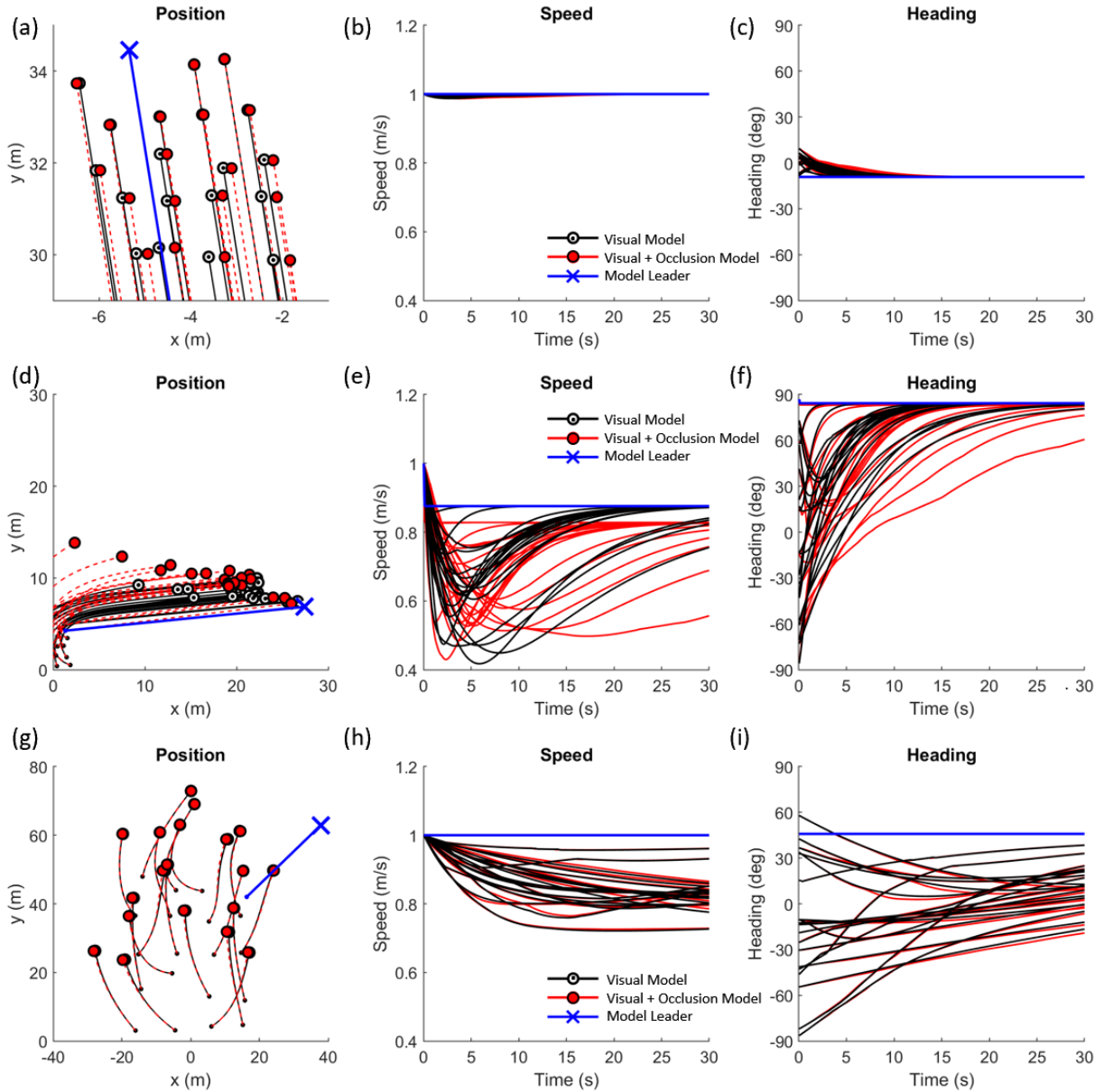


Figure 5.5. Representative trials from the multi-agent simulations for heading. The black curves are the visual model, the red curves are the visual + occlusion model, and the blue line is the agent at the front of the simulated crowd. The black and red 'o's are to compare final position across models. Panels (a-c) is a generated crowd with an IPD of 1 m and  $\pm 10^\circ$  variation in heading. Panels (d-f) is a generated crowd with an IPD of 1 m and  $\pm 90^\circ$  variation in heading. Panels (g-i) is a generated crowd with an IPD of 10 m and  $\pm 90^\circ$  variation in heading.

curves are the visual model, the red curves are the visual + occlusion model, and the blue curve is the agent at the front of the simulated crowd.

In the first row (Figure 5.5, panels a-c), the initial interpersonal distance was 1 m and the initial heading range was  $\pm 10^\circ$ . In the second row (Figure 5.5, panels d-f) the initial interpersonal distance was 1 m and the heading range was  $\pm 90^\circ$ . In the third row (Figure 5.5, panels g-i) the initial interpersonal distance was 10 m and the initial heading range was  $\pm 90^\circ$ . These trials produce similar patterns as seen with multi-agent simulations of varied initial speeds (Figure 5.3). Simulated crowd cohesion was highest when interpersonal distance and variation were low; high variation and low interpersonal distance caused high occlusion and thus the most differences between the models; and high variation and high interpersonal distance caused little to no cohesion. In addition, similar to varied speed simulations (Figure 5.3), varied heading simulations produce the largest differences between models when both variation is high and initial interpersonal distance is low, with information transferring from front to rear of the simulated crowd.

Heat maps of the standard deviation of final heading as a function of initial interpersonal distance and the initial heading rate appear in Figure 5.6. The temperature (color) of each cell is the standard deviation of the final heading across all agents run, averaged across the set of 20 runs for each initial condition. The heat maps estimate the cohesion of the simulated swarm, where low standard deviations (warm colors) signify high heading cohesion and high standard deviations (cool colors) signify low heading cohesion. The heading variation on cohesion heatmaps have a similar pattern to those found for speed in Figure 5.4: the behavioral model with a single decay with distance struggled to reach coherence, the behavioral model with a double decay with distance and the visual

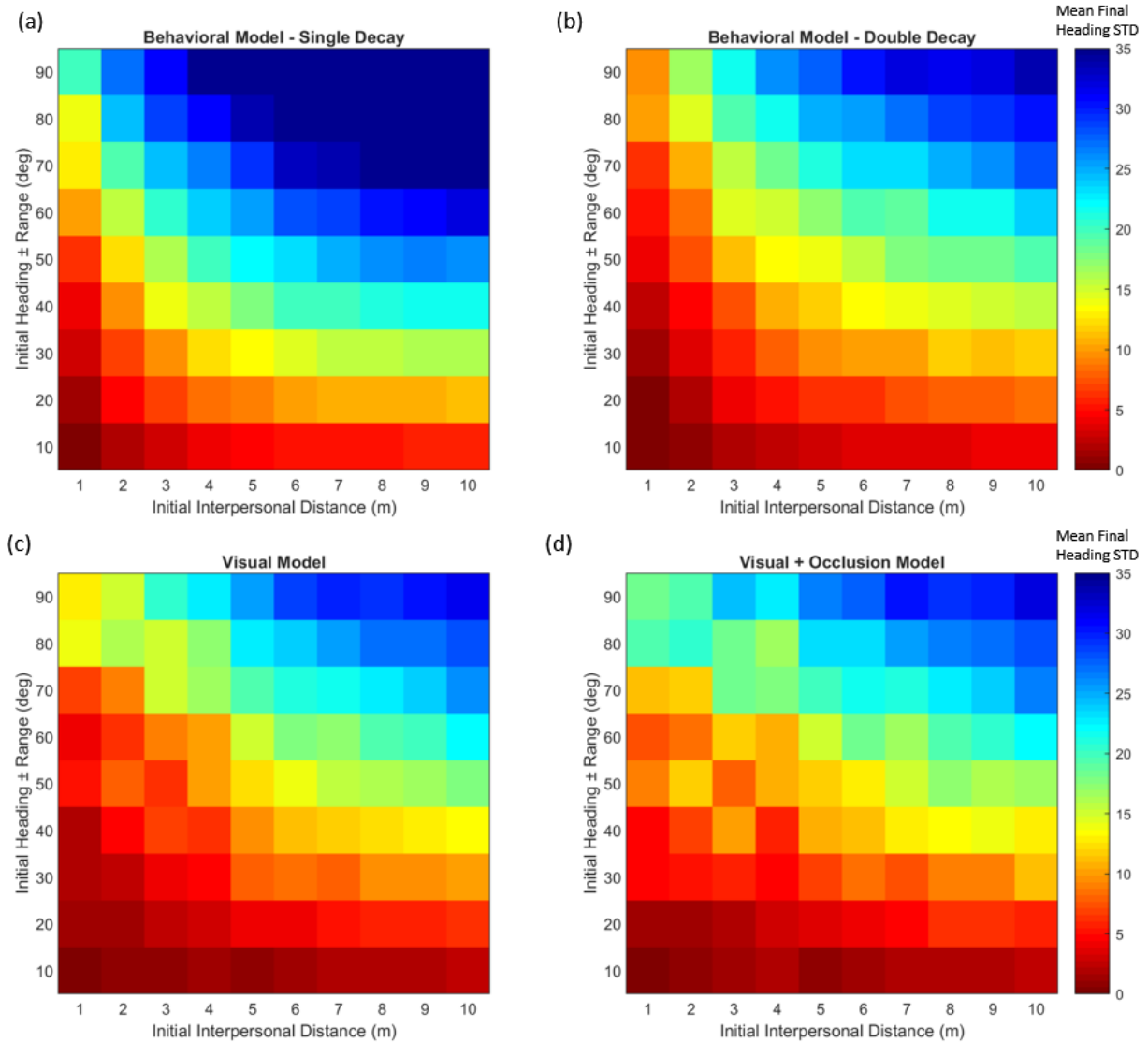


Figure 5.6. Heatmaps for the cohesion of heading across bulk simulations of crowds, plotting initial heading range by interpersonal distance, where each cell is the mean final standard deviation for heading across all generated agents across all trials in the set. Warm means that combination of initial conditions leads to higher cohesion (lower standard deviation) and cool means lower cohesion (higher standard deviation).

model were similar and cohere, and the visual + occlusion model did not cohere at low initial interpersonal distances and high heading variations.

## 5.7 Multi-agent Simulations - Discussion

Taken together, the heatmaps for speed (Figure 5.4) and heading (Figure 5.6) reveal how distance and occlusion interact to govern self-organized collective motion. The behavioral model's single exponential decay with distance eliminated neighbor influence at distances greater than 5 m. Beyond that distance, the model could not cohere beyond the initial speed range of  $\pm 0.3$  m/s or the initial heading range of  $\pm 30^\circ$ . The behavioral model with an exponential double decay with distance and the visual model both extended the range of interaction, and produced very similar results for both speed and heading. However, the reason is likely different for each model. In the behavioral model's case, it captured the previous observation that influence decays gradually to the nearest neighbor and then rapidly within the crowd. For the visual model, the model did not have any maximal distance: it responded to neighbors regardless of distance, although far neighbors had a very small influence. Both these conditions produced similar patterns of cohesion across initial conditions.

A new pattern was observed for the visual motion + occlusion model, however. The component for occlusion only impacted the model when initial interpersonal distance was low and variation in initial speed or heading was high. In these circumstances, cohesion was less compared to mid-levels of interpersonal distance for a given initial speed or heading range. This makes a novel predication about the visual + occlusion model. Low interpersonal distances drive higher rates of occlusion within a given trial. When near neighbors occlude far neighbors in a crowd, a given agent can only align with their near neighbors instead of aligning with all members of the crowd. This serves to dampen far neighbor influence, restricting information transfer from the front to the rear of the crowd,

and increasing the time for a crowd coherence. However, low ranges of initial speeds or headings are too similar to reveal this effect, and once interpersonal distance increases, occlusion becomes less likely, and the crowd is faster to reach cohesion.

## **5.8 Conclusion**

Chapter 5 found that the visual model not only reproduced trajectories at the local level, but also generated collective motion at the global level. It demonstrated that the model accurately predicted individual trajectories in real crowds. Experiment 6 showed that individual trajectories are better simulated with a component for dynamic occlusion when interpersonal distances are low. Using these concepts, multi-agent simulations of crowds were conducted, to understand how self-organized collective motion depends on initial conditions. One important takeaway from these multi-agent simulations was that the visual model for following a crowd with dynamic occlusion produces a unique pattern of crowd cohesion, making a unique prediction about cohesion in crowds when interpersonal distance is low. These demonstrate that the rules of engagement for following a single leader not only explain individual trajectories in a crowd, but also generate collective motion in a set of interacting agents, further establishing the strength of a bottom-up approach to following a crowd. These findings encourage future work to empirically test if human behavior matches these predictions.

Thus far, this research has treated the visual variables of optical expansion and angular velocity as the only information governing collective motion. However, there are other sources of visual information that a follower could be taking advantage of. Chapter

6 will examine these other information sources and how they may play a role in following a crowd.



## CHAPTER 6

### The Role of Visual Information for Depth

## **6.1 Introduction**

In Chapter 2, I reported that the speed and heading of a pedestrian follower are controlled by nulling the optical expansion and angular velocity of a single leader. In Chapters 3 and 4, I found that the same visual information plus occlusion closely simulate a pedestrian walking with a virtual crowd. However, the previous experimental work isolated 2D optical variables independent of other information that pedestrians may be using to guide their locomotion. Although Chapter 5 demonstrated that the visual + occlusion model reproduces individual trajectories in real crowds, additional visual information is available to humans in crowds. In this chapter, I will investigate the role of depth information and experimentally test the contribution of binocular and perspective information to pedestrian following. Two experiments were conducted. Experiment 7 tested following a single neighbor and Experiment 8 tested following a in a crowd with multiple neighbors. In each experiment, the previous sources of information (optical expansion, angular velocity, occlusion) were put in conflict with standard sources of depth information (binocular disparity, vergance, declination angle from the horizon to the ground plane). This dissociated the contribution of traditional depth information and motion information in the control of speed and heading alignment. I will close with a discussion of how the visual + occlusion model may be expanded to include these sources of depth information.

## **6.2 Experiment 7 - Depth information to following a single neighbor**

To understand what visual information guides collective motion in crowds, the bottom-up approach begins by analyzing the information used to follow a single neighbor.

Once these visual "rules of engagement" are defined, they can be scaled up to crowd behavior. Experiment 2 demonstrated that optical expansion and angular velocity are sufficient visual information to control speed and heading alignment. However, in real scenes traditional depth information is also available, providing redundant visual information that a pedestrian could use to follow a leader. Experiment 2 eliminated binocular disparity and vergence by using synoptic displays, in which the same image is presented to each eye, corresponding to an optically infinite distance. In addition, perspective information was eliminated by removing a ground plane. In particular, this eliminated the declination angle from the horizon to the base of the target on the ground plane, which specifies egocentric distance. These sources of depth information would have been in conflict with the manipulated optical expansion, and may have led to different locomotor behaviors in participants. For example, when the target artificially expanded, this specified that the target slowed down relative to the participant. However, if binocular and perspective information were also available, the participant may have perceived the target as not changing speed, but as changing size at a constant distance. This may have elicited no speed change from the participant. In most circumstances in the real world, a followed neighbor does not change size in this way, but it is important to understand whether 3D depth information and 2D motion information both contribute to locomotor control, and if so, how they interact. The underlying issue is how the visual system uses information to control action. If the visual system helps build an internal 3D model of the world to control action, then "traditional" depth cues (e.g. binocular disparity, vergence, and perspective) are required so as to build a robust internal model to control action (Loomis & Beall, 2004). However, if the visual system maps any and all optical variables

to actions directly, then traditional depth cues are redundant (but not necessarily unused) and evidence against an internal 3D model of the world (Warren, 1998).

To empirically test whether depth information is used to control following, Experiment 7 was conducted with an identical structure to Experiment 2, but with binocular disparity, vergence, and a ground plane added. These results were then compared to Experiment 2 to evaluate three hypotheses: (1) if depth information plays no role in following a single neighbor, results of Experiment 7 should be identical to Experiment 2, (2) if depth information completely overrides optical expansion and angular velocity, following behavior should be the same in conditions where optical expansion is in conflict with depth information compared to control conditions where these information sources are in agreement, or (3) all sources of information are combined, yielding an intermediate response in the conflict condition.

### **Participants**

12 undergraduate students were recruited at Brown University, 6 female and 6 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

### **Apparatus**

The experiment was conducted in the Virtual Environment Navigation Laboratory (VENLab) at Brown University. Participants walked in a 12 x 14 m tracking area while wearing a stereoscopic head-mounted display (HMD, Samsung Odyssey). Participant head position and orientation were recorded with the Odyssey's inside-out tracking at 90 Hz.

## **Procedure**

The procedure was identical to Experiment 2. Participants began each trial by facing toward orientation pole floating in space. After 3 seconds, the orientation pole disappeared and the target appeared within their field of view, moving away from the participant. Participants were instructed to “walk with the target as if walking down the street with it, while maintaining a consistent distance and orientation to the target”.

## **Display**

The target was a virtual pole (0.4 m diameter, 2 m tall), sitting on a ground plane. The target appeared at five initial eccentricities relative to the participant:  $-60^\circ$  and  $-30^\circ$  (to left of the participant),  $0^\circ$  (directly in front of the participant), and  $+30^\circ$  and  $+60^\circ$  (to the right of the participant) and at 2 distances (1 m, 4 m). Unlike Experiment 2 (which presented the pole at an infinite distance as to be synoptic), the target in Experiment 7 was presented at a defined location, providing distance information from disparity and vergence. The target pole would move away from the participant in a straight line at 0.8 m/s. 4.5 seconds into the trial, the pole changed its rate of expansion (specifying a speed change of  $-0.2$  m/s, no change,  $+0.2$  m/s), angular velocity (specifying a turn  $-30^\circ$  to the left, no change, turning  $+30^\circ$  to the right), or both. Importantly, the target never actually changed its speed, only its size. While optical expansion in the absence of other information normally specifies changes in speed, in Experiment 7 it conflicted with depth information, which specified that the target remained at a constant distance and changed size. There was no such conflict in the control condition, when the target did not change size. After the participant traveled the length of the room, the trial would end, the target would disappear, and a new orientation pole would appear, signifying the beginning of the next trial.

## **Design**

The experiment had a 5x2x3x3 factorial design: 5 initial target eccentricities (-60°, -30°, 0°, +30°, +60°), 2 initial distances (1 m, 4 m), 3 optical expansion rates (-0.2 m/s, no change, +0.2 m/s), 3 changes in angular velocity (+30°, no change, -30°), for 90 trials. These was replicated twice over two blocks, for a total of 180 trials.

## **Data processing**

2160 trials were recorded across participants. Of those, 105 had to be discarded due to tracking errors with the Odyssey, leaving 2055 usable trials (95% of trials). Otherwise, the data processing was identical to Experiment 2.

Because the target could change size while remaining at a fixed distance from the participant, two target trajectories were calculated to compare with participant data. The first was the synoptic trajectory determined by the optical expansion, calculated by taking the distance specified by the visual angle of a 0.4 m object at the target's eccentricity from the participant's known position at every time step. The second was calculated by taking the distance specified by depth information (binocular disparity, vergence, and declination angle to the ground plane) and the target's eccentricity.

## **6.3 Experiment 7 - Results**

Representative trials from Experiment 7 are found in Figure 6.1. These plot position, speed, and heading of the participant (black curves) and the two target trajectories specified by optical expansion (red curves) and depth information (blue curves). In these trials, participants' responses appear to be somewhere in between the two trajectories. To summarize this over all trials, Figure 6.2 plots the mean time series of participant speed

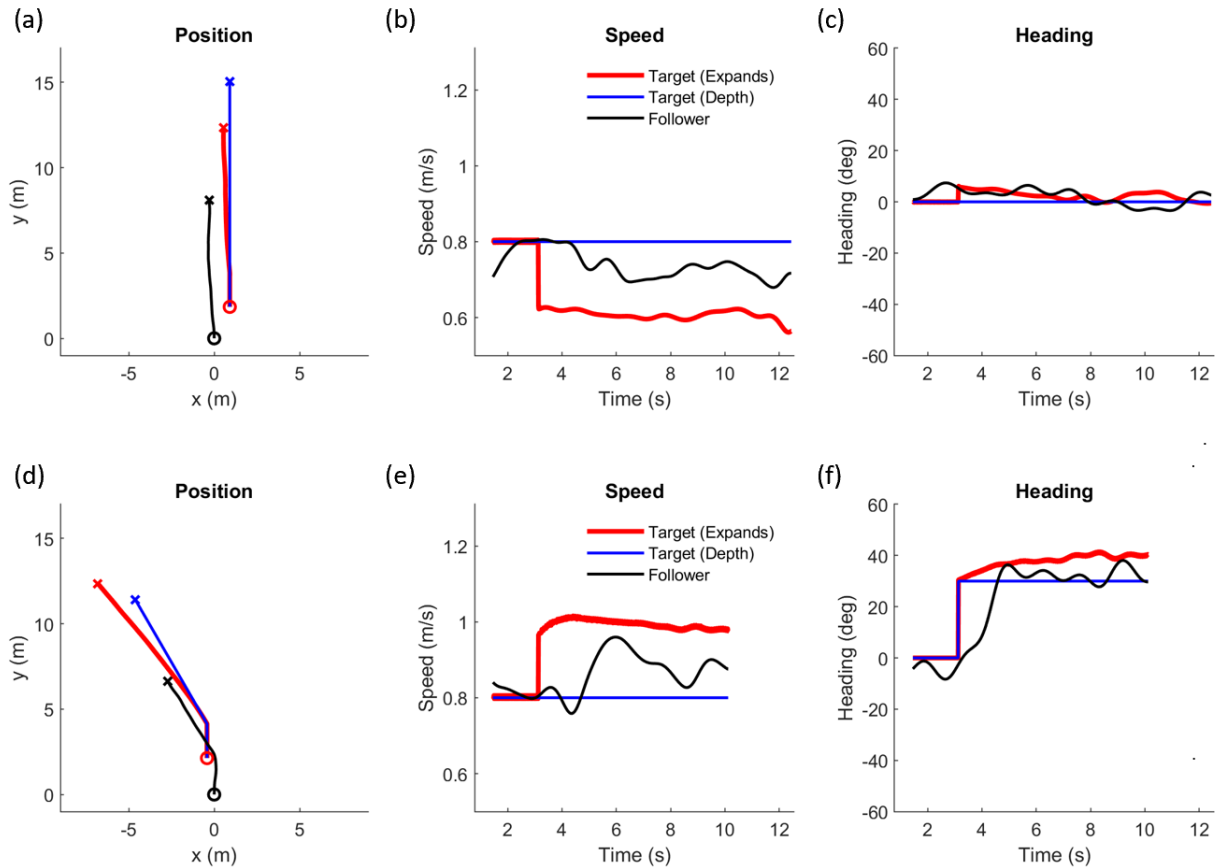


Figure 6.1. Representative trials from Experiment 7, plotting position, speed, and heading of the participant follower and the target. The target's timeseries was separated into two sets: (1) the target's synoptic motion specified by artificial optical expansion (red curves) and (2) the information specified by depth (blue curves). Panels (a-c) plots a 1 m distant at  $+60^\circ$  eccentricity trial with added expansion (specifying slowing down). Panels (d-f) plots a 1 m distant at  $-30^\circ$  eccentricity trial where the target and turns left while contracting (specifying speeding up).

from Experiment 2 (solid curves) and Experiment 7 (dashed curves) to compare the magnitude of speed change. Experiment 2 showed expansion rate had a large influence on participant speed. When optical expansion conflicted with depth information in Experiment 7, participants still responded to the expansion, but that response was reduced. This suggests that participants combined conflicting information to control their walking speed.

Figure 6.3 plots the influence of optical expansion (collapsed across expansion and contraction) on mean speed for comparisons. A two-way ANOVA found a significant

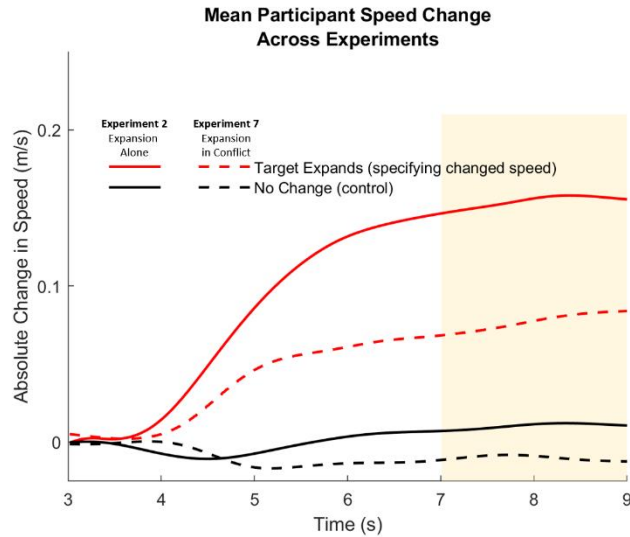


Figure 6.2. Mean participant speed across Experiments 2 and 7. In the absence of any other information, the target expanding (red solid curve; Experiment 2) caused large changes in participant speed. When information conflicted (red dashed curve; Experiment 7), participants still changed speed, but to a lesser degree.

difference between Experiment 2 and Experiment 7 ( $F(1, 44) = 16.53, p < 0.001$ ), as well as between the presence of expansion and control conditions ( $F(1, 44) = 42.5, p < 0.001$ ). Critically, there was also an interaction effect between experiments and conditions ( $F(1, 44) = 8.31, p = 0.006$ ). Post-hoc tests further found a significant difference between condition means for expansion alone and control for Experiment 2 ( $t(22) = 7.22, p < 0.001$ ),

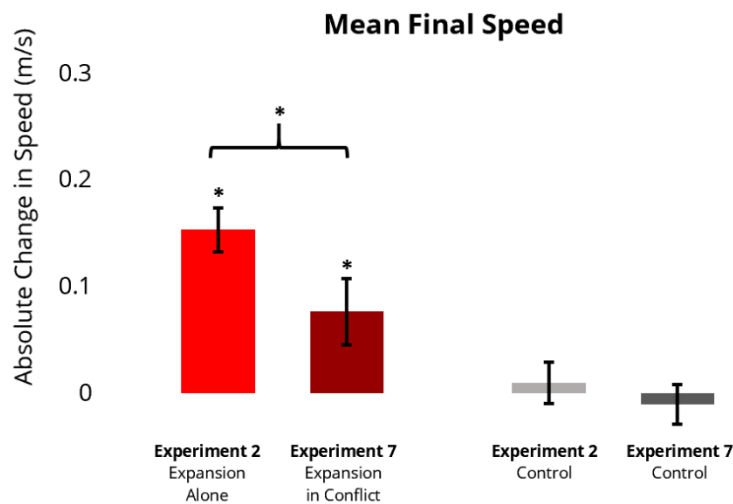


Figure 6.3 Collapsed mean final speed for Experiments 2 and 7, suggesting that visual information is combined for following a single leader.



reiterating the findings from Chapter 2 that expansion alone is sufficient to control following. Critically, there was a significant difference between condition means for expansion in conflict with depth and control for Experiment 7 ( $t(22) = 2.40, p = 0.026$ ), indicating that depth information does not completely override optical expansion. In addition, there was a significant difference between condition means for expansion alone from Experiment 2 and expansion in conflict with depth from Experiment 7 ( $t(22) = 4.90, p < 0.001$ ), suggesting that expansion alone does not control following. These results together provide evidence for the hypothesis that visual information combines for following a single leader.

#### **6.4 Experiment 7 - Discussion**

Comparing Experiment 7 to Experiment 2 demonstrates two important findings: 1) expansion influences following behavior and this occurs even when there are other sources of visual information that are in conflict, and 2) when information is conflicting, followers combine these sources of information. This shows that “traditional” depth sources (i.e. binocular disparity, vergence, and perspective to the ground plane) do play a role in following a single leader.

However, the results of Experiment 7 are not evidence that the visual system is building an internal 3D model for action. Experiment 2 provides evidence that expansion alone is sufficient for following. What Experiment 7 does provide is evidence that depth sources are used. Notably, these depth sources are formally related to optical expansion (Warren 1998; e.g. change in declination angle to the ground plane is just expansion of a contour in the vertical dimension). A control law could cancel change in declination angle, change in vergence, and change in disparity much like canceling optical expansion. This

suggests that these additional optical variables could be incorporated into the control laws that govern the visual model, allowing it to best account for the full range of optical information used to following a single leader.

An important question remains however: how these sources of depth information interact at the crowd level and if previously found principles of superposition hold true when information is in conflict.

### **6.5 Experiment 8 - Information Conflict in Crowds**

Crowd behavior emerges from the pairwise interactions among all members in the crowd. The bottom-up approach argues that the interaction rules for leader-follower pairs should scale up to all the interactions in a crowd. The purpose of Experiment 8 was to determine whether optical expansion is combined with depth information, as observed for a single neighbor (Experiment 7), when walking with multiple neighbors in a crowd. The present experiment thus included conditions in which optical expansion was isolated, was placed in conflict with binocular and/or ground plane information, was consistent with that information, as well as a no-change control condition. In short, the question was whether the combination of motion and depth information would generalize from a single leader to a crowd.

#### **Participants**

12 undergraduate students were recruited at Brown University, 8 female and 4 male. None reported any visual or motor impairment. The research protocol was approved by Brown University's Institutional Review Board, in accordance with the principles

expressed in the Declaration of Helsinki. Informed consent was obtained from all participants, who were paid for their time.

### **Apparatus**

The experiment was conducted in the Virtual Environment Navigation Laboratory (VENLab) at Brown University. The apparatus was the same as in the previous experiment. Participant head position and orientation were recorded with the Odyssey's inside-out tracking at 90 Hz.

### **Procedure**

Participants began each trial by orientating toward a target sitting on the ground plane. After 3 seconds, the orientation target disappeared and a group of 12 target poles to be followed appeared within the participant's field of view. Participants were instructed to "walk with the poles as if walking down the street with them, while maintaining a consistent distance and orientation to them".

### **Displays**

Each target was a virtual pole (0.4 m in width, 2 m in height). The targets appeared on two arcs: near and far. The near arc was 2 m distant, with 5 poles equally spaced along a 100° arc. The far arc was 3.5 m distant, with 7 poles along a 120° arc. Poles were randomly jittered in every trial by a uniform distribution of  $\pm 0.25$  m. The target poles initially moved away from the participant in a straight line at 1.0 m/s. After a random interval of 3-5 seconds nine of twelve or all twelve poles changed their speed (-0.2 m/s, no change, +0.2 m/s), or rate of expansion (specifying -0.2 m/s, no change, specifying +0.2 m/s). After a 3 second perturbation, the target poles returned to 1.0 m/s speed or ceased to have artificial expansion, dependent on condition. In half of all trials, the poles could be

presented synoptically or with disparity to manipulate binocular information about the distance of the poles. This was cross-balanced with a ground plane present or removed, to vary declination angle information about the distance of the targets, and to find if this information interacted with binocular information. After the participant traveled the length of the room, the trial ended, the target poles disappeared, the ground plane reappeared if it had been removed, and a new orientation target appeared, signifying the beginning of the next trial.

### **Design**

The experiment had a  $2 \times 2 \times 2 \times 2 \times 2 + 4$  factorial design: 2 information conditions (optical expansion, speed change),  $\times 2$  specified speed perturbations (-0.2, +0.2 m/s),  $\times 2$  perturbed subsets (9, 12 targets),  $\times 2$  ground plane conditions (ground, no ground),  $\times 2$  binocular conditions (synoptic, disparity), plus 4 no-change control trials (one in each ground plane  $\times$  binocular condition). This was replicated 4 times, presented in random order, for a total of 144 trials.

### **Data processing**

1728 trials were recorded across participants. Of those, 41 had to be discarded due to tracking errors with the Odyssey, leaving 1687 usable trials (97.6% of trials). Trials were processed using the same methods from Experiment 7.

## **6.6 Experiment 8 - Results**

Representative trials from Experiment 8 appear in Figure 6.4. The top row (panels a-c) presents a trial with binocular disparity and a ground plane, in which 9 of 12 targets increased their speed, so motion and depth information were consistent. The participant

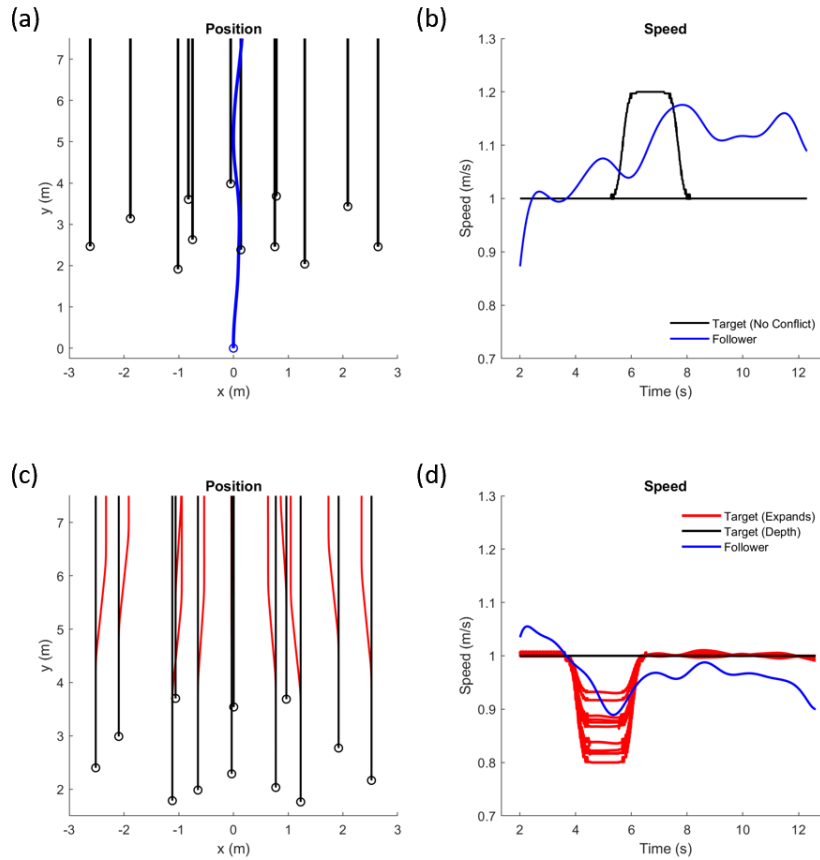


Figure 6.4. Representative trials from Experiment 8. Panels (a-b) is a trial in where 9 of 12 targets sped up (so there was no information conflict), with a ground plane and binocular disparity. Panels (c-d) is a trial where all 12 targets expanded, causing a conflict between information sources.

sped up in response. The second row (panels d-e) presents a trial with disparity and a ground plane in which all 12 targets optically expanded, creating a conflict between expansion information (red curves) and depth information (black curves). The participant slowed down in response. Because the targets expanded, those with higher eccentricities also appeared to turn toward the participant.

To find how conflicting information influences participant speed, Figure 6.5 plots expansion change trials for two condition sets: no conflicting information (e.g. targets presented synoptically and without a ground plane) and conflicting information (e.g. targets

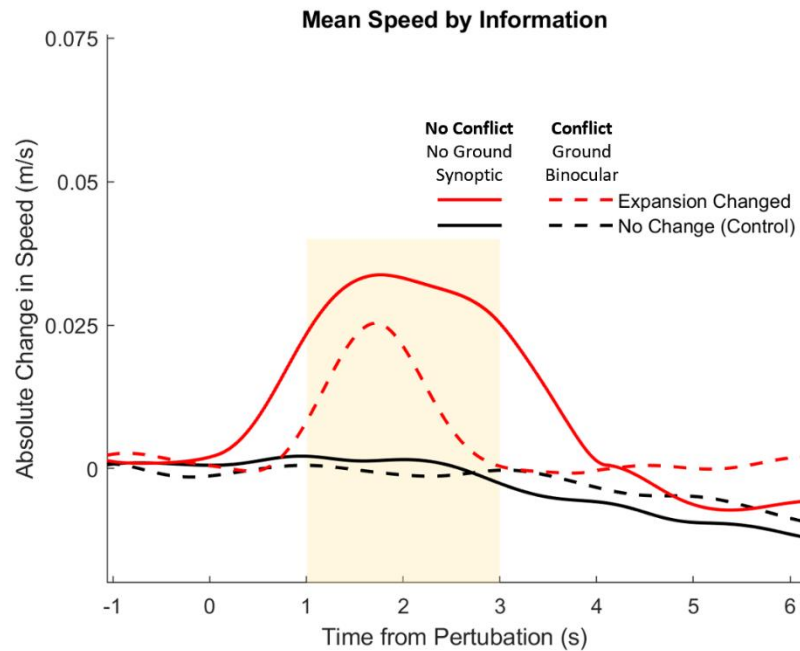


Figure 6.5. Mean participant speed from Experiment 8 by whether the information was conflicting or not conflicting.

presented binocularly and with a ground plane). Walking with a crowd showed effects similar to walking with a single neighbor (Experiment 7), where non-conflicting information (solid red curve) causes a large change in participant speed, while conflicting information (dashed red curve) causes a change in speed, but to a reduced effect. In these circumstances, participants used both sources of information.

Figure 6.6 plots the influence of optical expansion on mean speed across 1-3 seconds after perturbation (yellow shaded region in Figure 6.5), by information conflict, for comparisons. A two-way ANOVA found a significant difference between expansion change and control ( $F(1, 44) = 9.97, p = 0.003$ ), but no significant difference between no conflict and conflict conditions ( $F(1, 44) = 0.47, p = 0.495$ ), and no interaction ( $F(1, 44) = 0.1, p = 0.749$ ). Post-hoc tests found a significant difference between expansion changed

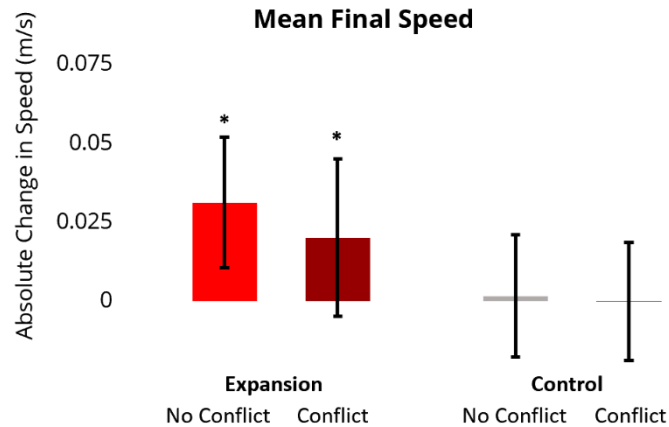


Figure 6.6. Mean participant speed from Experiment 8, across 1-2 seconds after perturbation, by whether the information was conflicting or not conflicting.

and control for the no conflict conditions ( $t(22) = 2.10, p = 0.047$ ) and for the conflict conditions ( $t(22) = 2.32, p = 0.030$ ), but no difference between no conflict and conflict when expansion changed ( $t(22) = 0.23, p = 0.821$ ).

## 6.7 Experiment 8 - Discussion

These results provide evidence that expansion change is used for the control of speed not just with following a leader, but following multiple neighbors. This provides further evidence for the superposition hypothesis for following in a crowd. The results also demonstrate that expansion's influence on a pedestrian is equivalent, regardless if the expansion information is conflicting with depth cues or not, when following a crowd. Experiment 7 found this not to be the case, that no conflicting information is more influential than conflicting. Potential explanations are that when increasing the number of neighbors to follow, it not only reduces the influence of expansion information alone, but reduces the influence of depth cues as well (that are also formally related to optical expansion). In addition, expansion specifies different motion changes in neighbors,

dependent on eccentricity. When following a single leader, those changes can specify speed or heading change. With multiple neighbors, given their structure in Experiment 8, it provides conflicting information about the crowd. This is, if the two neighbors at farthest and opposite sides of the crowd expand, it specifies they are both turning toward the participant (as seen in Figure 6.4, panel c). This conflict of information may also be contributing to the reduced influence of expansion seen in Experiment 8. Further study is necessary to find why this difference occurs in following a crowd.

## **6.8 Conclusion**

The findings from Experiment 7 and Experiment 8 demonstrated that multiple sources of visual information are used for following a single neighbor and following a crowd. When these sources were in conflict (for example, when a neighbor was changing size while also maintaining disparity, Experiment 7; or when multiple neighbors are exhibiting conflicting expansion, Experiment 8) the information is combined, reducing the effect of optical expansion.

While the visual model for following a crowd currently only uses optical expansion, angular velocity, and occlusion information for simulation purposes, an improved model could account for these other sources as well. Change in declination angle to where a neighbor rests on the ground is functionally identical to change in optical expansion along the vertical dimension. While the visual model currently only calculates change in the horizontal dimension (as it is normally locked to change in the vertical), adding a component to the model for vertical change would function as capturing change in the declination angle along the ground plane. Disparity could receive the same form of



treatment: treat the angle of each eye to a neighbor as different model inputs and null that change in angle difference.

By definition, the model's control laws nulls any changes in visual information. Separating out changes in declination angle or disparity into its component changes in visual angles provides a means to model these within the visual model framework. This would allow the model to continue to function without an explicit recovery of distance, while accounting for circumstances where expansion and depth information are in conflict (as rare as they may be).

## CHAPTER 7

### Conclusion

## 7.1 Discussion

Crowd collective motion is inherently complex. The presented work does address how pedestrians locomote in all manner of crowds (i.e. walking to a goal or collision avoidance), nor the final answers on the information they use, nor the ultimate method by which to model them. These are questions that will be asked and examined for decades to come, as the layers of collective motion and visual information are peeled away.

What this research does provide is a strong foundational work on the role of visual information in collective motion and a powerful method by which to model it, that relies on empirically grounded studies of how information controls following. Chapter 2 explored the visual information provided by optical expansion and angular velocity, as control variables for following a single neighbor. Based on the laws of optics for these variables, I derived the visual model, a visual control law that describes the "rules of engagement" for a leader-follower pair. Model parameters were fit to data from Experiment 1, and successfully predicted human responses in Experiment 2. The results demonstrated that these optical variables are sources of information used to control walking speed and heading when following a leader. Chapter 3 attempted to scale up this model to following in crowds. I applied the visual control law for following a single neighbor to walking with multiple neighbors, on the assumption that these pairwise interactions could be linearly combined (the superposition principle). However, simulations of previous experimental data on a participant walking with a virtual crowd found that the visual model failed to generalize to some conditions, and did not always perform as well as the previous behavioral model. I hypothesized that dynamic occlusion of some far neighbors might reduce their influence on the participant, and a visual occlusion component was missing

from the model. Chapter 4 experimentally tested the effect of visual occlusion when following a virtual crowd. The results showed that occlusion reduced the influence of far neighbors. Occlusion was added to the visual model, instantiating the visual + occlusion model, by computing each neighbor's occlusion and weights their influence by the proportion of visibility. Re-simulating the experimental data showed that adding the occlusion component significantly improved the fit of the visual model, and was even better than the behavioral model for dense crowds at farther distances. Chapter 5 took the visual model out of the laboratory, to test whether it generalized to crowds in more naturalistic settings. Results showed that the model with dynamic occlusion accounted for individual trajectories of participants in a "swarm". Moreover, multi-agent simulations showed that the model converges to collective motion over a wide range of initial conditions, demonstrating that visual control laws for alignment are sufficient to generate robust collective motion. Chapter 6 recognized that optical expansion, angular velocity, and dynamic occlusion may not be the only visual information that governs collective motion. Experiments on walking with a single neighbor and a crowd found that declination angle to the ground plane and binocular disparity also played a role as control variables for walking speed. However, change in declination angle and change in disparity are analogous to optical expansion. Thus, the visual model might be expanded, in principle, by adding components that null these variables as well.

Taken together, the combined work of this dissertation provides the first experimental evidence on how visual information governs local interactions when walking with a crowd, giving rise to human collective motion. This research derived, tested,

explored, and experimentally validated the first model of coherent motion in crowds in which locomotion is controlled solely by visual information.

## **7.2 Theoretical Conclusions**

Examining the body of this research raises broader theoretical conclusions about crowd behavior and modeling approaches to it, that deserve discussion. A few of these are the following: that alignment-based neighbor interactions are sufficient for collective motion, that visual information provides a more parsimonious account compared to previous modeling frameworks, that bottom-up approaches successfully model collective motion, and that locomotor control can be explained by a linear combination of visual control laws.

My work has shown that the alignment of locomotion behaviors between neighbors is by itself sufficient to make an account of collective motion. However, alternative models have supplemented alignment with additional components of interaction (i.e. the attraction and repulsion between neighbors; Couzin et al., 2002). So, while my work has shown alignment is sufficient, it does not show it is necessary. And conversely, does demonstrate that attraction and repulsion are not necessary, but it does not test if they are sufficient as well. It has been shown that attraction / repulsion models can generate collective motion by means of attraction to neighbors ahead and repulsion from neighbors behind, as in locust swarms (Romanuczk, et al., 2009). However, this work presumes agents with a panoramic field of view, whereas humans have a 180° field of view and only respond to neighbors ahead of them (Rio, Dachner, & Warren, 2018). This implies that attraction/repulsion is insufficient to generate collective motion in human crowds.

An objection could be made that humans do not ordinarily follow each other or walk together in groups, but only walk to goals while avoiding collisions with moving neighbors or static obstacles. However, there is no evidence that pedestrians do not walk together in collective motion. By instructing participants to “follow” a leader or “walk with” a crowd, my goal was to test whether following and alignment of locomotor behaviors are sufficient to generate and explain collective motion. This work has demonstrated this to be the case.

Nearly all previous models of collective motion have been “behavioral” models, based on a bird’s-eye view of physical variables (e.g. the input speed, heading, and distance directly into the model). The major contribution of my thesis demonstrates that a model that uses visual information, based on an embedded view and optical variables accomplishes the following functions: it accounts for a wide range of collective motion data, it accounts for some aspects of the data more accurately compared to behavioral models (e.g. asymmetries of relative direction of motion and distance), it is more parsimonious (by accounting for behavioral asymmetries by using asymmetries in information), it functions without an explicit distance term (instead relying on Euclid’s laws), and, finally, it provides a causal/mechanistic explanation of locomotor control (as opposed to a phenomenological description).

Phenomenological models only describe behavior, and assume that physical variables somehow influence the agent without showing how. Mechanistic models, like the visual model I have presented here, provide a causal mechanism by showing how visual information influences the agent’s behavior.

Next, my work demonstrates that the bottom-up approach works: by beginning with the study and modeling of interactions in leader-follower pairs, I could scale up to model interactions with multiple neighbors, and show that multi-agent simulations of the model generate collective motion based on the same interaction rules that govern pairwise interactions.

There is one critical point to be made on this however: scaling up by using the principle of superposition (that interactions between multiple neighbors are the sum or average of pairwise interactions) has limits: although the behavioral model and visual model obey superposition, the visual + occlusion model does not. This form of the visual model violates a primary rule of superposition when it considers the position of an occluded neighbor in reference to other neighbors. This breaks superposition because it forces the consideration of interactions between neighbors, instead only considering the interactions between a follower and their neighbors. This violation of superposition has implications for other senses, too, e.g. in fish, pressure waves sensed by the lateral line organ will be affected by the bodies of “occluding” fish; in bats, returning echoes will suffer interference from the emissions of other bats, etc.

To close out theoretical conclusions, thus far, it should be noted that my research has demonstrated that locomotor control can be explained by the linear combination of visual control laws (e.g. acceleration of speed or heading is calculated by nulling change in angular velocity, horizontal expansion, declination angle (vertical expansion), and disparity. This is different from the linear combination of visual depth cues to form a consistent 3D model (or percept) of the environment, which is then used to control locomotion, for example, by keeping a constant distance from neighbors, or a constant

position in the 3D configuration of neighbors. Instead, I have demonstrated the information for alignment in collective motion comes from the nulling and combination of visual motion.

### **7.3 Future Directions**

Further work remains to be done to develop the visual model of collective motion. It would serve to compare this model to other contemporary models of crowd behavior. The multi-agent simulations in Chapter 5 make novel predictions about the conditions for collective motion. The experimental results of Chapter 6 imply that additional sources of information need to be incorporated into the model. Finally, a successful visual model of collective motion open questions about visual information transfer in complex systems such as human crowds.

With the inclusion of occlusion into the visual model, the visual + occlusion model is robust at simulating following a single neighbor and following a crowd. However, as noted in Chapter 1, there are many other contemporary models of crowd behavior. A very important next step for the visual + occlusion model would be to compare it to these competing models in the literature. Importantly, to test if the same data of following in human crowds presented here can be explained by: models with only attraction and repulsion components from neighbors (e.g. Romanczuk, et al., 2009), models in which each agent is attracted to its own goal and avoids collisions with neighbors (e.g. Moussaid, et al., 2011), or the Social Force Model (e.g., Helbing, et al., 1995). Further, model comparisons need to be done to find which models best simulate following in a crowd.



Chapter 5 explored how the visual model generates collective motion at larger scales in multi-agent simulations. One important find was that the visual model + occlusion model produces a unique pattern of crowd cohesion compared to previous physical and behavioral models, and makes a unique prediction about the conditions for collective motion. Other models predict that when crowd density is high, the crowd will quickly converge to cohesive motion regardless of the variation in initial conditions. However, the visual model with dynamic occlusion predicts that at high densities (between 1-3 m IPD), increased occlusion results in slower convergence and less cohesive motion when the variation initial conditions is large. These findings encourage future work to empirical test if human behavior matches this prediction. One means to test this is by using the framework from Experiment 6 and expanding it through variations in crowd density and higher magnitudes of variations in neighbor behavior.

As found in Chapter 6, in addition to optical expansion, angular velocity, and dynamic occlusion, two other sources of information contribute to the control of walking speed: binocular disparity and declination angle. However, the model does not currently take these variables into account. One potential option to improve the model is for it to treat horizontal optical expansion and vertical change in declination angle as separate variables, and add a term that nulls change in the declination angle. With objects in the real world, however, horizontal and vertical expansion are highly correlated, as rigid objects optically expand in all directions when they are approached. Thus, another option would be to treat the change in declination angle as the vertical component of expansion. This question might be investigated experimentally by isolating vertical expansion and declination angle and testing whether they have a similar influence on speed control.

Binocular information might be incorporated into the model by adding terms that null the change in disparity and the change in vergence angle.

Taking the model further, it can be understood that angular velocity and optical expansion and contraction are composed of local optical flow vectors. Future versions of the visual model could incorporate a "front end" that senses the local motion of points on the surface of objects, and derive higher-level properties such as region velocity and expansion rate to drive the model. This is doubly important when it is considered that individual neighbors within crowds in the real world are complex-shaped objects. Changing the visual model to use optic flow as and binocular disparity input could help generalize the model to real-world crowds.

Finally, crowds are complex dynamic systems. Collective motion in any system poses the question of the means and rate at which state information is transferred among the individuals within the group. For example, bird flocks have been shown to have a rapid rate of transfer, allowing for individual birds to move in response to a predator so that the whole flock is cohesive (Bialek et al., 2014). Large human crowds with many neighbors may be studied and modeled to investigate information transfer. In particular, multi-agent simulations can be used to test the effect of neighborhood size and time delays on the rate of information propagation through the crowd from front to rear. Occlusion and field of view are also likely to constrain these properties, further analysis is needed to understanding information transfer.

This future work taken together would lead to a more accurate, robust, and ecologically valid model of the visual control of locomotion in crowds. The model has significant applications to areas of society including urban planning, evacuations, and

architectural design, in addition to improving our basic understandings of human vision and crowd behavior.

## References

- Bastien, R., & Romanczuk, P. (2020). A model of collective behavior based purely on vision. *Science Advances*, 6(6), eaay0792.
- Bialek, W., Cavagna, A., Giardina, I., Mora, T., Pohl, O., Silvestri, E., Viale, M., & Walczak, A. M. (2014). Social interactions dominate speed control in poising natural flocks near criticality. *Proceedings of the National Academy of Sciences*, 111(20), 7212-7217.
- Bonneaud, S., & Warren, W. H. (2014). An empirically-grounded emergent approach to modeling pedestrian behavior. In *Pedestrian and Evacuation Dynamics 2012* (pp. 625-638). Springer International Publishing.
- Breder Jr, C. M. (1954). Equations descriptive of fish schools and other animal aggregations. *Ecology*, 35(3), 361-370.
- Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985–2010. *Vision Research*, 51(13), 1431-1456.
- Chen, Z. (2012). Object-based attention: A tutorial review. *Attention, Perception, & Psychophysics*, 74(5), 784-802.
- Couzin, I. D., Krause, J., James, R., Ruxton, G. D., & Franks, N. R. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, 218(1), 1-11.
- Couzin, I. D., & Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior*, 32, 1-75.

- Dachner, G. C., & Warren, W. H. (2014). Behavioral dynamics of heading alignment in pedestrian following. *Transportation Research Procedia*, 2, 69-76.
- Dachner, G. C., & Warren, W. H. (2017). Vision-based model for the joint control of speed and heading in pedestrian following. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.
- Dutra, T. B., Marques, R., Cavalcante-Neto, J. B., Vidal, C. A., & Pettré, J. (2017, May). Gradient-based steering for vision-based crowd simulation algorithms. In *Computer Graphics Forum* (Vol. 36, No. 2, pp. 337-348).
- Fajen, B. R., & Warren, W. H. (2003). Behavioral dynamics of steering, obstacle avoidance, and route selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 343.
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.*, 59, 167-192.
- Gibson, J. J. (1979). The ecological approach to visual perception. *Psychology Press*.
- Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical Review E*, 51(5), 4282.
- Hoffmann, E. R. (1994). Estimation of time to vehicle arrival—Effects of age on use of available visual information. *Perception*, 23(8), 947-955.
- Howarth, S. J., & Callaghan, J. P. (2009). The rule of 1 s for padding kinematic data prior to digital filtering: Influence of sampling and filter cutoff frequencies. *Journal of Electromyography and Kinesiology*, 19(5), 875-881.
- Loomis, J. M., & Beall, A. C. (2004). Model-based control of perception/action. In *Optic Flow and Beyond* (pp. 421-441). Springer, Dordrecht.

- Moussaïd, M., Helbing, D., & Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17), 6884-6888.
- Nuijten, M. B., Wetzels, R., Matzke, D., Dolan, C. V., & Wagenmakers, E. J. (2015). A default Bayesian hypothesis test for mediation. *Behavior Research Methods*, 47(1), 85-97.
- Ondřej, J., Pettré, J., Olivier, A. H., & Donikian, S. (2010). A synthetic-vision based steering approach for crowd simulation. *In ACM Transactions on Graphics (TOG)* (Vol. 29, No. 4, p. 123). ACM.
- Regan, D., & Beverley, K. I. (1973). The dissociation of sideways movements from movements in depth: psychophysics. *Vision Research*, 13(12), 2403-2415.
- Regan, D., & Beverley, K. I. (1979). Binocular and monocular stimuli for motion in depth: changing-disparity and changing-size feed the same motion-in-depth stage. *Vision Research*, 19(12), 1331-1342.
- Regan, D., & Hamstra, S. J. (1993). Dissociation of discrimination thresholds for time to contact and for rate of angular expansion. *Vision Research*, 33(4), 447-462.
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21(4), 25-34.
- Rio, K. W., Dachner, G. C., & Warren, W. H. (2018). Local interactions underlying collective motion in human crowds. *Proc. R. Soc. B*, 285(1878), 20180611.
- Rio, K. W., Rhea, C. K., & Warren, W. H. (2014). Follow the leader: Visual control of speed in pedestrian following. *Journal of Vision*, 14(2), 4-4.

- Romanczuk, P., Couzin, I. D., & Schimansky-Geier, L. (2009). Collective motion due to individual escape and pursuit response. *Physical Review Letters*, *102*(1), 010602.
- Schellinck, J., & White, T. (2011). A review of attraction and repulsion models of aggregation: Methods, findings and a discussion of model validation. *Ecological Modelling*, *222*(11), 1897-1911.
- Shanno, D. F., 1985. On Broyden-Fletcher-Goldfarb-Shanno method. *Journal of Optimization Theory and Applications* *46*(1), 87–94.
- Strandburg-Peshkin, A., Twomey, C. R., Bode, N. W., Kao, A. B., Katz, Y., Ioannou, C. C., Rosenthal, S. B., Torney, C. J., Wu, H. S., Levin, S. A., & Couzin, I. D. (2013). Visual sensory networks and effective information transfer in animal groups. *Current Biology*, *23*(17), R709-R711.
- Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., & Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, *75*(6), 1226.
- Vicsek, T., & Zafeiris, A. (2012). Collective motion. *Physics Reports*, *517*(3), 71-140.
- Wann, J. P., Poulter, D. R., & Purcell, C. (2011). Reduced sensitivity to visual looming inflates the risk posed by speeding vehicles when children try to cross the road. *Psychological Science*, *22*(4), 429-434.
- Warren, W. H. (1998). Visually controlled locomotion: 40 years later. *Ecological Psychology*, *10*(3-4), 177-219.
- Warren, W. H. (2018). Collective motion in human crowds. *Current Directions in Psychological Science*, *27*(4), 232-240.

Warren, W. H., Dachner, G. C. (2018). Comparing Disk and Doughnut Models of Collective Crowd Motion. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.

Wirth, T. D., & Warren, W. (2017). Recruitment of Pedestrians into Collective Crowd Motion. Poster, *Vision Sciences Society Conference (VSS)*; St. Pete Beach, FL.