



Guided by gaze: Prioritization strategy when navigating through a virtual crowd can be assessed through gaze activity

Meerhoff L.A.^{a,b,*}, Bruneau J.^a, Vu A.^a, Olivier A.-H.^{a,b}, Pettré J.^a

^a Inria, Univ Rennes, CNRS, IRISA - UMR 6074, F-35000 Rennes, France

^b Univ Rennes, Inria, M2S - EA 7470, F-35000 Rennes, France

ARTICLE INFO

Classification codes:

2300 Human experimental psychology

2323 Visual perception

2330 Motor processes

2346 Attention

Keywords:

Crowd navigation

Pareto optimality

Pedestrian

Virtual reality

Locomotion

Interaction neighborhood

ABSTRACT

Modelling crowd behavior is essential for the management of mass events and pedestrian traffic. Current microscopic approaches consider the individual's behavior to predict the effect of individual actions in local interactions on the collective scale of the crowd motion. Recent developments in the use of virtual reality as an experimental tool have offered an opportunity to extend the understanding of these interactions in controlled and repeatable settings. Nevertheless, based on kinematics alone, it remains difficult to tease out how these interactions unfold. Therefore, we tested the hypothesis that gaze activity provides additional information about pedestrian interactions. Using an eye tracker, we recorded the participant's gaze behavior whilst navigating through a virtual crowd. Results revealed that gaze was consistently attracted to virtual walkers with the smallest values of distance at closest approach (*DCA*) and time to closest approach (*TtCA*), indicating a higher risk of collision. Moreover, virtual walkers gazed upon before an avoidance maneuver was initiated had a high risk of collision and were typically avoided in the subsequent avoidance maneuver. We argue that humans navigate through crowds by selecting only few interactions and that gaze reveals how a walker prioritizes these interactions. Moreover, we pose that combining kinematic and gaze data provides new opportunities for studying how interactions are selected by pedestrians walking through crowded dynamic environments.

1. Introduction

As with collective animal behavior (Couzin & Krause, 2003), movements of human crowds emerge from the combination of the local interactions between neighboring pedestrians in the crowd (Moussaïd et al., 2012). The effect of single interactions on the formation of human locomotion trajectories has been extensively studied, for example during collision avoidance (Croft & Panchuk, 2017; Knorr, Willacker, Hermsdörfer, Glasauer, & Krüger, 2016; Olivier, Marin, Crétual, & Pettré, 2012), following (Lemerrier et al., 2012; Rio, Rhea, & Warren, 2014), or grouping (Moussaïd, Perozo, Garnier, Helbing, & Theraulaz, 2010; Rio, Dachner, & Warren, 2018). However, the notion of an interaction neighborhood needs to be developed to fully explain the structure of the collective motion. Interaction neighborhoods in human crowds are typically designed somewhat arbitrarily according to the modeler's beliefs, based on for example distance (Helbing & Molnar, 1995), topology (Van den Berg, Guy, Lin, & Manocha, 2011), or vision (Ondřej, Pettré, Olivier, & Donikian, 2010). These assumptions are necessary as the high number of potential interaction sources make it nearly impossible to infer any causality based on the combined

interactions. An interaction neighborhood is a formalization of which neighbors are likely to have an effect on the walker's trajectory. A formalization solely based on kinematics cannot be achieved without arbitrary hypotheses, therefore additional measurements are required to fully understand the process.

Gaze activity, in addition to kinematics, may provide a good indication of where humans get their (visual) information from for the control of human locomotion (e.g., Patla, 1997; Warren Jr, 1998; Nummenmaa, Hyönä, & Hietanen, 2009). Marigold and Patla (2007) showed that gaze is drawn towards task relevant aspects of the environment, as walkers fixated on locations where they would eventually step arguably to maximize the amount of information available for a safe foot placement. Moreover, gaze behavior changes depending on the risk of collision (Jovancevic-Misic & Hayhoe, 2009). In an experiment where participants came across confederates that would either seek collision or avoid collision, Jovancevic-Misic and Hayhoe (2009) showed that participants adapted their gaze behavior depending on which confederate was approaching them. The risky confederates would draw more attention, whereas the confederate that did not pose a collision risk was gazed at less. Additionally, gaze behavior provided

* Corresponding author.

E-mail address: rensmeerhoff@gmail.com (L.A. Meerhoff).

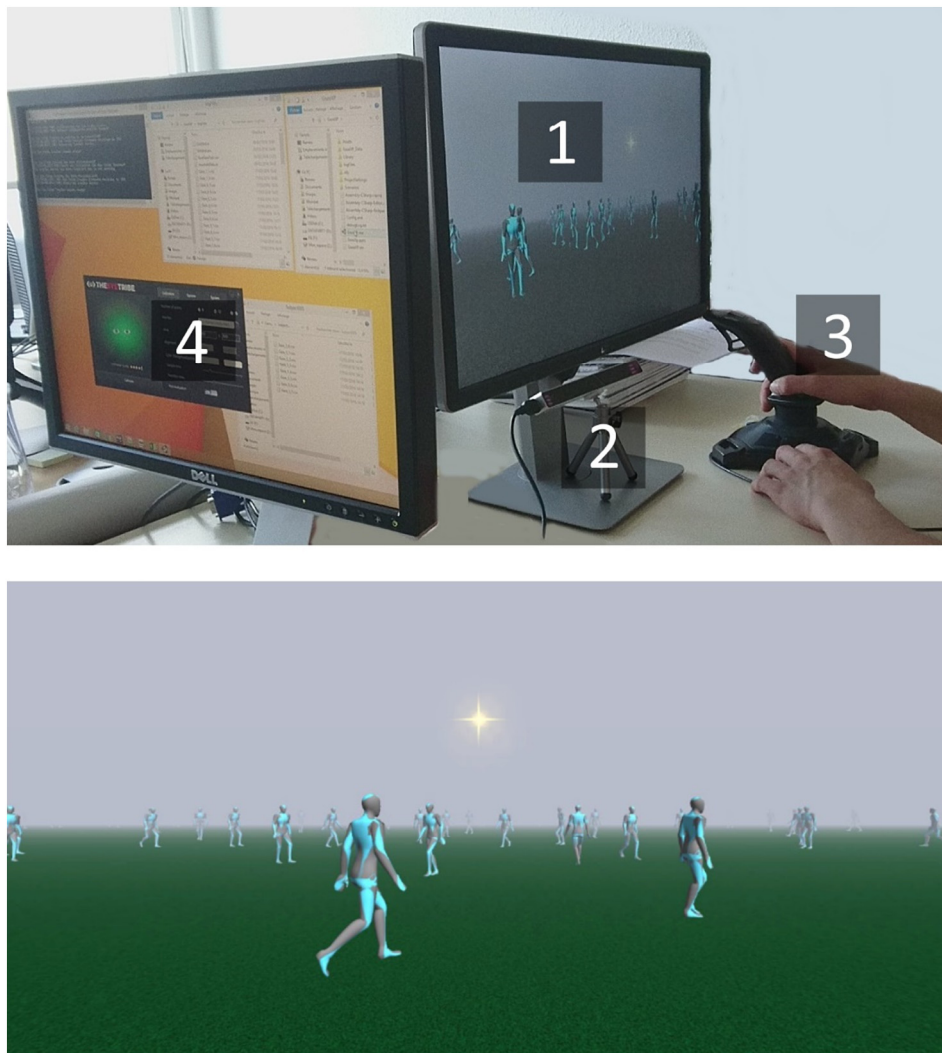


Fig. 1. *Top*) Participants sat in front of a 24" screen (1). An eye tracker below the screen recorded participant's gaze (2). Participants moved through the virtual environment using a joystick (3). An additional screen (4) allowed the researcher to monitor the experiment. *Bottom*) Screenshot of the virtual environment through which participants navigated.

information about how a pedestrian engaged an interaction (Croft & Panchuk, 2017). In a collision avoidance task where an interferer crossed a participant's trajectory at 90°, Croft and Panchuk (2017) showed that the gaze behavior revealed whether the participant would cross in front or behind the interferer. Participants tended to pass behind when they looked at the interferer early in the interaction and when the duration of the fixation was long. It has even been shown that people use the gaze behavior of others to adjust their behavior (Dicks, Clashing, O'Reilly, & Mills, 2016; Colombi, Scianna, & Alaia, 2016; Colombi & Scianna, 2017), likely because it informs about the action intentions. Therefore, we focus on understanding how a person interacts with its environment based on its gaze activity. More specifically, it can thus be surmised that gaze activity may shed light onto which walkers prompt collision avoidance when walking through a crowded environment.

The risk of collision with another pedestrian can for example be quantified with distance- or time-based metrics such as the Distance at Closest Approach (*DCA*, also referred to as Minimal Predicted Distance; Olivier et al., 2012; Olivier, Marin, Crétual, Berthoz, & Pettré, 2013) and the Time to Closest Approach (*TtCA*; Dutra, Marques, Cavalcante-Neto, Vidal, & Pettré, 2017). Assuming that, at each time step, both pedestrians maintain their current heading and velocity, the future closest approach can be computed through linear extrapolation of each

walkers' heading and velocity. *DCA* is then the predicted distance between these walkers **at**-, and *TtCA* the time **until**-, the instant of closest approach. As *DCA* and *TtCA* can be computed at every time step and simultaneously incorporate the action of two walkers, these metrics provide an interesting descriptor of the dynamics of an interaction between two walkers. In previous experiments, walkers have been shown to avoid collision when *DCA* is below a threshold of about 1 m in real-world conditions (e.g., Olivier et al., 2012, 2013), which has been replicated in virtual reality (Lynch et al., 2017). To avoid collisions, it is evident that typically lower *TtCA* or *DCA* values correspond to an increased necessity to interact to avoid collision. However, it is challenging to quantify this necessity to interact as it requires combining distance- and time-based metrics. Collision can be avoided with small continuous adjustments early in the interaction, but also with a late abrupt adjustment a short time before the closest approach. As such, in isolation neither *TtCA* nor *DCA* provides a full description of when a walker needs to respond to another walker. If there is not a lot of time left until the closest approach (i.e., low *TtCA*), it may well be that the distance at closest approach will be large (i.e., high *DCA*), and therefore no action would be required. On the other hand, it may be that the distance at closest approach is predicted to be small (i.e., low *DCA*), but with a lot of time left until closest approach (i.e., high *TtCA*) it may not (yet) be necessary to take action (due to the uncertainty of what

happens in the meantime). Therefore, describing the necessity to interact with a walker should incorporate both distance- and time-based metrics.

Most studies on pedestrian interactions are based on kinematics alone, making it difficult to infer what constitutes the interaction neighborhood. As gaze plays an important role in locomotion and pedestrian interactions, we aim to examine the hypothesis that gaze guides the interaction between pedestrians. In the current study we propose to combine the distance- and time-based metrics using Pareto optimality (Keller, 2017), which is a granular method to rank cases based on the combination of two parameters. This Pareto ranking allows us to assess the risk of collision, without making arbitrary assumptions about how distance- and time-based metrics relate. We aim to investigate whether gaze activity informs about how pedestrian interactions unfold. More specifically, we hypothesized that, during locomotion through a crowd, gaze is drawn towards the walkers that pose a risk of collision. Secondly, we hypothesized that an avoidance maneuver was initiated to avoid the walker fixated upon immediately prior to initiating the maneuver. To have full control over the crowd participants were exposed to, as well as kinematics of all neighbors' motion, we performed experiments in a virtual environment, which has been used previously in similar experiments (Schwarzkopf et al., 2013).

2. Methods

2.1. Participants

Nineteen participants (6 women, 13 men) volunteered for this experiment with a mean age of 25 (SD = 4) years. They were naive with respect to the purpose of the experiment. All had normal or corrected-to-normal vision. They gave written and informed consent and the study conformed to the declaration of Helsinki.

2.2. Apparatus

Participants were seated in front of a 24" screen, under which a desktop Eye Tribe eye tracker device was positioned (see Fig. 1, top). The eye tracker accuracy was 0.5–1° and the sampling frequency was 60 Hz. Before the start of the experiment, the eye tracker was calibrated. Participants controlled their virtual motion using a joystick which was previously validated as a locomotion metaphor for pedestrian avoidance behavior in virtual reality (Olivier, Bruneau, Kulpa, & Pettré, 2018). The longitudinal axis of the joystick controlled speed linearly from 0.8 m·s⁻¹ to 2.0 m·s⁻¹. The lateral axis controlled the angular rotation speed linearly from -25° s⁻¹ to 25° s⁻¹. When no action was applied on the joystick, the speed was 1.33 m·s⁻¹ and the angular rotation speed was 0° s⁻¹. That is, without moving the joystick, the participant moved straight and at constant speed through the virtual world. We designed a populated virtual environment using Unity software (see Fig. 1, bottom). The virtual walkers navigated in a straight line and were performing collision avoidance using the RVO2 library (Van den Berg et al., 2011). Parameters were set for a late avoidance by the virtual walkers to ensure active collision avoidance of the participant while preventing any collision that would disturb the participants (between virtual walkers or from a virtual walker coming from a blind spot). The virtual walkers were projected uniformly in terms of color, appearance and speed (1.33 m·s⁻¹).

2.3. Task

Participants had to move through a populated virtual environment towards a distant target (see Fig. 1, bottom) without colliding into the virtual walkers as they would in real conditions. After the participants were habituated to the task with two practice trials, the participants performed two experimental trials; all trials lasted 2 min. Directly afterwards, the participants completed a trial to evaluate the quality of

the eye tracker. During the evaluation trial, participants did not control their movements through the virtual world (and were put on a collision-free trajectory). Additionally, one of the visible virtual walkers was randomly selected to be colored red. Rather than navigating towards a target, the participants had to fixate their gaze upon the walker that was colored red. By matching the eye tracker data with the known red walker, we assessed the accuracy of the algorithm that determined gaze fixations.

2.4. Data analysis

2.4.1. Pareto ranking

We ranked the risk of collision of each virtual walker by combining a distance- and a time-based metric. More specifically, we used a Pareto ranking (Keller, 2017) as a method to rank each virtual walker based on their combination of *DCA* and *TtCA*. A Pareto rank was allocated to each agent at each timestep following two iterative processes. During the lower order iteration, the virtual walkers on the Pareto frontier were identified. The higher order iteration repeated this process, but only with the remaining walkers. The Pareto rank was determined as the *n*th higher order iteration (i.e., virtual walkers found in the first iteration were ranked 1, virtual walkers found in the second iteration were ranked 2, and so on). The Pareto frontier (see connected data points in Fig. 2, top) contained all virtual walkers that were 'Pareto efficient': they had a potentially optimal trade-off of both *TtCA* and *DCA*. This frontier was determined iteratively by first selecting the virtual walker with the lowest *X*-score (e.g., A46 in Fig. 2) and subsequently identifying the walker with the lowest *X*-score of the remaining walkers and a lower *Y*-score than the previously identified walker (e.g., A35 after A46 in Fig. 2). This process was repeated until no virtual

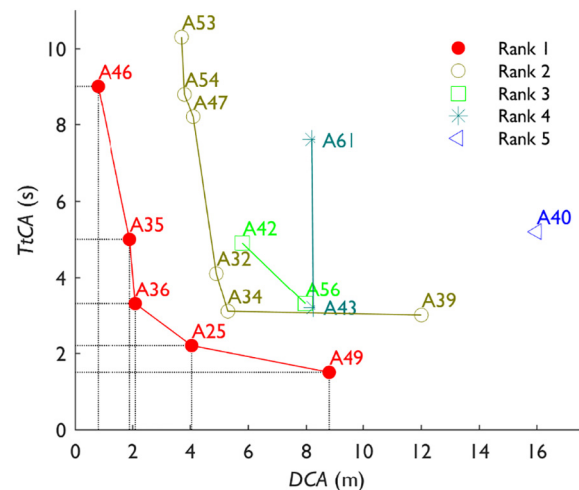


Fig. 2. Top) *DCA* and *TtCA* values of the virtual walkers at one instant during the experiment. The dotted vertical and horizontal lines visualize that no other walker had a more Pareto-efficient combination of *TtCA* and *DCA*. Bottom) Screenshot of the virtual walkers colored according to their corresponding Pareto ranks.

walker remained (i.e., the remaining virtual walkers had a higher Y than the last selected virtual walker, e.g., A49). Note that it was indifferent whether X or Y represents $TtCA$ or DCA . To clarify, the highest rank (i.e., 1) was given to the walkers with the lowest combinations of $TtCA$ and DCA . As an example, we visualized the Pareto ranks of the virtual walkers in Fig. 2 (bottom). A virtual walker with a high rank (i.e., closer to 1) would demand to be interacted with as the risk of collision would be high relative to the walkers with lower rankings.

2.4.2. Fixation

The fixation location of the participant's gaze was derived by interpreting the 2D on-screen gaze location as the line of sight in the 3D virtual world. Additionally, we had to correct for the movements of all the virtual walkers which posed three challenges: 1) gaze location was not fixed because the virtual walkers are moving, 2) other virtual walkers might walk into the line of sight whilst gaze was tracking a virtual walker further back, and 3) gaze behavior alternated between scanning and fixating. To tackle these problems effectively, we build an algorithm to classify the participants' gaze behavior.

First, we computed a distance score for every virtual walker based on the distance of each walker to the line of sight (derived from the 2D on-screen gaze location), similar to determining gaze fixations in static images. However, a virtual walker that was walking past at a different depth but not gazed upon may have been closer to the line of sight than the virtual walker that was being tracked. Therefore, as the second step, we computed a continuity score with which the distance score was boosted whenever a virtual walker was tracked over time. This prioritized virtual walkers that were close to the line of sight over a longer period of time. The virtual walker with the highest score was determined as the virtual walker that the participant fixated upon. The third step was to exclude fixations lasting shorter than 0.2 s to omit scanning and random variations (Salvucci & Goldberg, 2000). We validated the algorithm by computing the percentage correctly identified fixations using an evaluation trial. We excluded two participants as only 1% and 38% of the fixations were correctly identified, indicating the eye tracker data as unreliable. The fixations of the remaining participants were correctly identified with an average score of 84% ($SD = 12\%$).

2.4.3. Relative Fixation Frequency (RFF)

To avoid any bias in our interpretations, we computed the Relative Fixation Frequency (RFF) because the absolute fixation frequency may by definition differ per Pareto rank since the number of walkers per rank can differ. RFF is the ratio between the number of observed (i.e., absolute fixation frequency) and expected fixations in each rank. The number of expected fixations was determined by looking at how many virtual walkers there were in each rank at each time step, multiplied by the total number of fixations. If the participant would randomly look at a virtual walker on-screen, RFF would thus be 1. Alternatively, an RFF of 2 implied that the participant gazed upon a virtual walker in that ranking twice as often as expected by chance; conversely, an RFF of 0.5 only half as often.

2.4.4. Avoided virtual walkers

To establish the onset and end of an adjustment, we used the participant's input on the joystick. First, we computed the participant's linear and rotational speed. Then we identified every non-zero sequence of at least 10 frames (0.17 s). Avoided walkers were defined as walkers that ended with an increased DCA at the end of a participant's adjustment. Lastly, we matched the avoidance behavior with the gaze fixations to identify the ranks of the walkers fixated upon before the onset of an avoidance maneuver, the corresponding fixation duration and the time difference between the end of the fixation and the start of the avoidance maneuver.

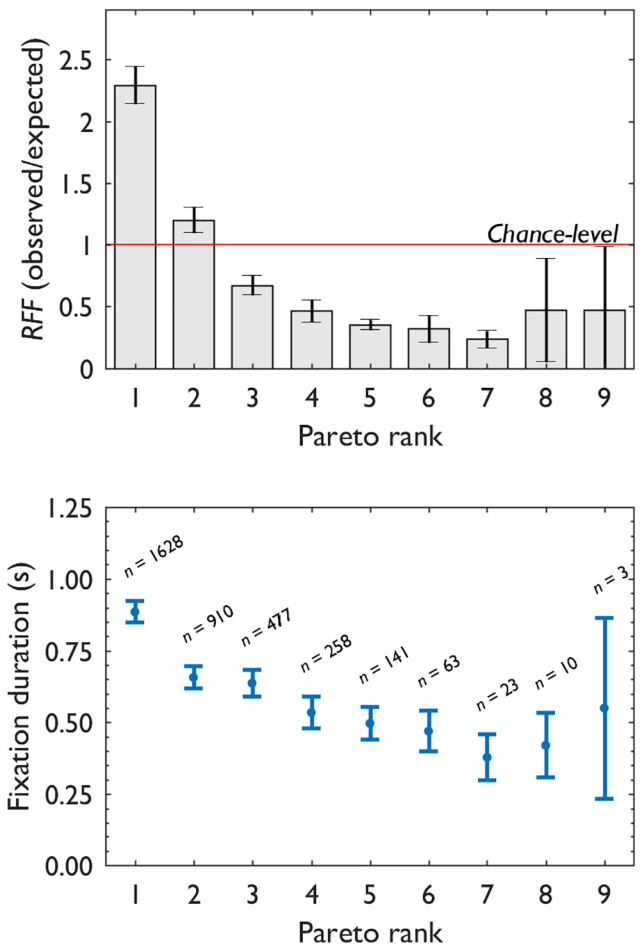


Fig. 3. Top) The average Relative Fixation Frequency (RFF) per participant ($n = 17$) specified per Pareto rank (whiskers indicating the 95%-CI). Values above chance-level (i.e., $RFF > 1$) indicate that the number of fixations was higher than expected by chance (based on the number of walkers visible in each Pareto rank relative to the total number of virtual walkers visible). Bottom) Average duration of a fixation on a walker in each Pareto rank, across all fixations. The number of fixations per Pareto rank is denoted by n .

2.5. Statistics

Statistics were performed using SPSS 20. We used a χ^2 -test to test whether RFF was randomly distributed across Pareto ranks. Additionally, we computed the 95% confidence intervals to indicate whether the participant average RFF scores were different from chance-level (i.e., $RFF = 1$). We used a Kruskal-Wallis test to assess the difference in fixation duration per Pareto rank. For the average fixation duration during each adjustment, a paired samples t -test was adopted to assess the difference between virtual walkers that were and were not avoided. Cohen's d was computed as a measure of effect size. The significance level was set at $p = 0.05$.

3. Results

3.1. Fixations

In Fig. 3 (top) we provided a descriptive overview of how often each rank of the Pareto front was fixated upon. It stands out that Pareto Rank 1 was fixated upon most often in absolute numbers (as indicated by n in Fig. 3, bottom) as well as when considering the Relative Fixation Frequency (RFF). More specifically, Pareto Rank 1 was fixated upon 2.3 times more than expected by chance (based on the number of virtual walkers in each rank during each fixation). A χ^2 -test revealed that the

distribution of the relative fixation along the Pareto ranks was not random ($\chi^2(8) = 6214.8, p < 0.001$). The 95%-CIs, based on the participants' average *RFF*, with 207 (SD = 41) fixations per participant, revealed that all scores were different from chance-level. Only in Pareto rank 9, with only 3 fixations in total, was *RFF* not significantly different from chance-level. Moreover, there was a statistically significant difference in the fixation duration for the different Pareto ranks ($H(8) = 186.221, p < 0.001$): the higher the Pareto rank, the longer the fixation duration (see Fig. 3, bottom). Note that *RFF* can also be computed based on *DCA* and *TtCA* separately. However, this would complicate a direct comparison with the Pareto ranking, as each Pareto rank may contain multiple agents. See the appendix for further detail.

3.2. Avoidance maneuvers

During all trials combined, a total of 997 maneuvers were registered. Of these maneuvers, 493 were away from the target and therefore considered as avoidance maneuvers; maneuvers towards the target were not necessarily avoidance maneuvers. At the onset of these avoidance maneuvers, there were on average 35 virtual walkers visible and we focus our analysis on the last agent that was fixated upon. The labels of all of these visible virtual walkers at the onset of the avoidance maneuver are presented in Table 1.

When examining specifically the fixations before the participant initiated an avoidance maneuver, it stands out that 88.1% of the avoidance maneuvers were preceded by a fixation on a walker that belonged to the first Pareto rank at the onset of the maneuver (see Fig. 4). Furthermore, Fig. 4 describes whether walkers that were fixated upon prior to an avoidance maneuver were indeed avoided; for example, of all fixations, 74% were upon avoided walkers in rank 1. Across all ranks, 79% of the walkers fixated upon before an avoidance maneuver were indeed avoided during the subsequent avoidance maneuver.

An independent *t*-test comparing the duration of fixations before an avoidance ($M \pm SD = 1.25 \pm 1.94s, 95\% CI[0.75 \ 1.35], n = 493$) and the duration of fixations not before an avoidance ($M \pm SD = 0.72 \pm 0.71s, 95\% CI[0.70 \ 1.16], n = 3521$) revealed that the duration of fixation was significantly higher when a fixation was followed by an avoidance maneuver, ($t(557) = 10.963, p < 0.001, Cohen's d = 0.77$).

Lastly, we described the distribution of three important characteristics of the walker that is assumed to have triggered the avoidance maneuver based on whoever was last fixated upon before an avoidance maneuver was initiated (see Fig. 5). Given the relatively constant velocities, the Distance- and *TtCA*-based histograms are fairly similar. In terms of *DCA*, walkers with lower values were clearly more likely to trigger an avoidance maneuver.

4. Discussion

The objective of our study was to investigate whether gaze underpins the interactions with other pedestrians as a step towards analyzing

Table 1
Categorization of the visible virtual walkers at the onset of all of the 493 avoidance maneuvers.

		Avoided	Not avoided	Totals
Counts	Fixated upon	391	102	493
	Not fixated upon	4,627	3,487	8,114
	Totals	5,018	3,589	17,214
Percentages	Fixated upon	79%	21%	100%
	Not fixated upon	57%	43%	100%

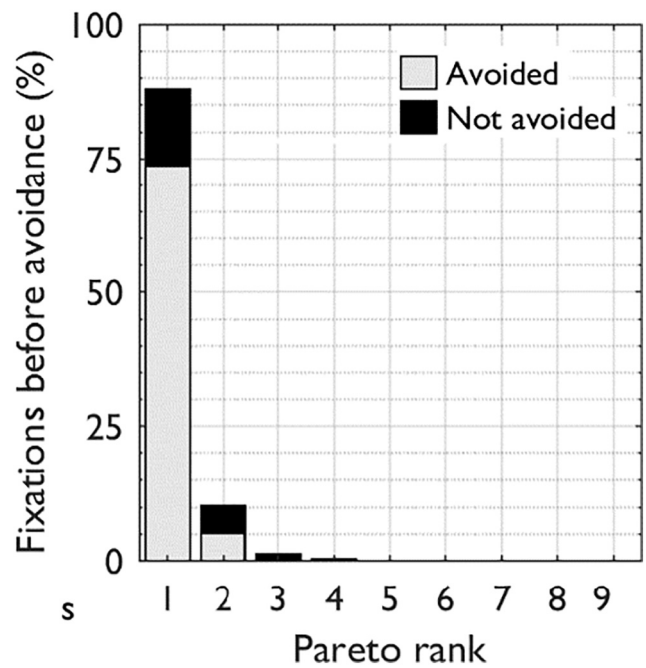


Fig. 4. The percentage of the virtual walkers fixated upon before an avoidance maneuver was initiated per Pareto rank. The shading of the bars indicates the portion of the virtual walkers that was indeed avoided during the subsequent avoidance maneuver.

the interaction neighborhood of a pedestrian walking through a crowd. We examined what can be deduced from gaze activity about pedestrian interactions. We analyzed the gaze behavior when navigating through a dynamic virtual environment (i.e., a crowd of virtual walkers), as well as the characteristics of the walkers fixated upon prior to an avoidance maneuver. We hypothesized that the risk of collision, as determined by a combination of a spatial (*DCA*) and temporal (*TtCA*) measure, would attract the participant's gaze. Moreover, we hypothesized that the fixation before an avoidance maneuver likely triggered that avoidance.

4.1. Fixations

Our results clearly revealed that most of the gaze fixations were directed towards the virtual walkers with the highest risk of collision with the participant (see Fig. 3, top). Even when correcting for the number of walkers per Pareto rank, the virtual walkers with Pareto rank 1 were gazed upon 2.3 times more often than expected by chance. Moreover, the average fixation duration was significantly higher for walkers with Pareto rank 1 compared to walkers with a lower ranking (except for rank 9, but it had only 3 fixations; see Fig. 3, bottom). These results are in accordance with the study of Jovancevic & Hayhoe (2009) where participants, in a constrained oval path, looked longer (also sooner) to potentially dangerous pedestrians (i.e., walking on a collision path with the observer). It was previously shown in static environments that walkers fixate on locations that maximize the information that can be integrated by the nervous system to ensure a safe foot placement (Marigold & Patla, 2007). Therefore, we suggest that risk of collision is salient and influenced the gaze behavior of the participants.

Previously, Andersen and Kim (2001) showed that observers can accurately detect upcoming collisions. In their experimental design, participants had to press a button to indicate that they perceived a collision. In our study, we showed that observers looked at virtual walkers that showed a risk of collision but also adapt their trajectories to these virtual walkers. However, it is important to note that we have displayed to participants relatively uniform crowds of walkers, walking

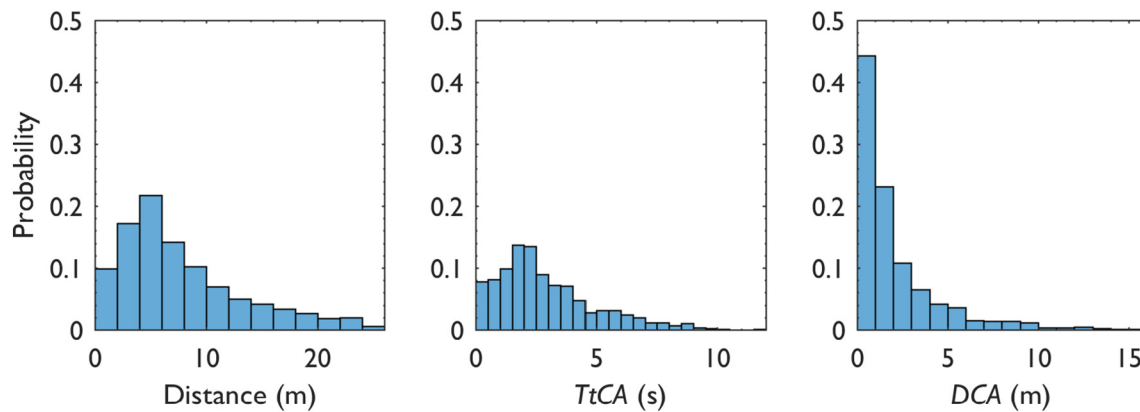


Fig. 5. The probability distributions for Distance to the participant (left), $TtCA$ (mid), and DCA (right) of the walkers that were fixated upon prior to an avoidance maneuver. The distributions were normalized between 0 and 1, with the sum of all bars of one subplot being equal to 1.

at the same speed, performing linear trajectories, with a similar level of density along the path. Future work is required to investigate the influence of additional features such as relative speed, angle of approach or the geometry of the path performed.

Although it is difficult to objectively quantify a walker's risk of collision with another walker, we suggest that it should be based on the notions of space and time, defined in our study by DCA and $TtCA$. We combined space and time features and were able to rank characters by the risk of collision using Pareto fronts (see appendix for a more detailed analysis of ranking virtual walker based on DCA and $TtCA$ alone). The advantage of the Pareto ranking is that no assumptions have to be made about the relative importance of DCA and $TtCA$ with regard to the risk of collision. However, one of the disadvantages of this technique is that several walkers will have the same rank. Additional research is required to provide a more accurate definition of risk of collision and how to enable more precise evaluation of this risk with respect to the relative position and motion of pedestrians. Dedicated experiments could exploit the relation between collision avoidance and gaze to further tease apart the temporal and spatial components underpinning risk of collision. Future research could for example refine the definition of the walker that triggered an avoidance maneuver. Currently, we investigated the hypothesis that whoever was gazed at immediately before an adaptation was the one that triggered an adaptation. Additionally, one could explore what the influence may have been of the n last walkers that were fixated upon. Furthermore, although it is difficult to establish concrete thresholds, it may also be possible to better identify who triggered an adaptation by excluding any walkers of which it is safe to assume that it could not have triggered an avoidance maneuver (e.g., $DCA > 3$ m).

4.2. Avoidance maneuvers

We checked the existence of a relation between gaze behaviors and avoidance maneuvers. To this end, we decomposed virtual walkers in sets of avoided and not-avoided virtual walkers, as well as sets of gazed and non-gazed virtual walkers. Of the last fixated upon walkers, a vast majority (79%) was indeed avoided, see Table 1. This indicates that the walker last fixated upon was likely to be the walker that triggered an avoidance maneuver. In 21% of the avoidance maneuvers on the other hand, the last fixated upon walker was not avoided. This could be because of the simplistic definition of 'avoided' (i.e., increased DCA at the end compared to the start of an avoidance). Alternatively, it is possible that the fixation that actually triggered the avoidance was not long enough to be included as a fixation (only fixations longer than 0.2 s were considered). Croft and Panchuk (2017) also reported avoidance behavior without fixation, which may be explained as the result of parafoveal gaze (Croft, Button, & Dicks, 2010) Parafoveal gaze would

not typically be registered as a fixation, but may indeed allow trajectory adjustments. Alternatively, it is also possible that these virtual walkers were indeed avoided without being gazed upon as a result of an adjustment to one of the other virtual walkers. Not only was the last fixated upon virtual walker indeed avoided most of the time, this last avoided virtual walker also belonged to Pareto rank 1 in most cases. As can be seen in see Fig. 4, virtually all fixated upon virtual walkers belong to Pareto rank 1 and 2. This highlights that combining the distance- and time-based metric (DCA and $TtCA$) using a Pareto ranking yields a meaningful categorization of risk of collision.

Although the specific definition of interaction neighborhood needs to be more closely examined, the histograms in Fig. 5 reveal that the walker that triggered a reaction is not that well described only based on distance. It seems that DCA would be the best classifier for determining which walkers may elicit an avoidance maneuver. Future work on interaction neighborhoods should examine how the combination of the proposed measures can inform which walkers influence how interactions between pedestrians emerge.

Additionally, the results revealed that the fixation duration of the fixations before an avoidance was initiated were significantly longer than fixations not prior to an avoidance maneuver. This provides further support that avoidance maneuvers are in fact guided by gaze. During the avoidance maneuver the participant attuned its adjustments to the information provided by fixating its gaze upon the walker that is being avoided. In our study, fixating on the walkers who have the highest risk of collision within a crowded environment may maximize the information available to ensure a safe navigation. By combining gaze and kinematic analyses, one can better assess the underlying process of these individual-level interactions. As such, any model for understanding how pedestrian interactions emerge could benefit from the information provided by a walker's gaze fixations.

4.3. Conclusion

While previous studies showed the importance of visual information in the control of locomotion (Patla, 1997; Warren, 1998) and the relation between gaze behavior and trajectory adjustments to another walker (Croft & Panchuk, 2017), our study showed for the first time that the combined analysis of gaze and trajectory data provides new insights on interaction neighborhoods and collision avoidance. The notion of interaction neighborhood is important in the design of crowd simulation algorithms, and the hypotheses made about it should be revisited in light of our current results. It would be interesting to explore more deeply the relations between gaze and avoidance behaviors, and for example, examine whether agents are fixated upon to evaluate the risk of collision associated with them, or to use the perceived information to control the avoidance maneuver. While performing a

navigation task in a highly dynamic environment, humans need to adapt their motion to avoid collision. The gaze activity is clearly oriented to serve this task and informs about the strategy of humans in their selection of most threatening interactions. Such information has many applications, such as the design of crowd simulation algorithms, the importance of which is crucial today for urbanism, safety, and architecture.

Our work opens several avenues for future research. One direction is to better understand what makes the risk of collision salient in complex settings such as navigating through a crowd of pedestrians. Specifically using more ecologically valid methods such as eye tracking in a Computer Assisted Virtual Environment (CAVE) could enhance the generalizability of the findings. Visually speaking, the features of the optical flow can indicate to the walker the risk of collision with all the moving objects around him. One question to determine is whether those features are salient to the human vision. Another direction is to explore the nature of the information gazed upon in the environment. Our scene was uniform in terms of density and virtual walkers' motion. Probably, the gaze behavior would be affected by changing scene features, such as density: would we be able to observe a change in the strategy from such data? A final direction would be to search for new relations between

gaze behaviors and the kinematics of avoidance maneuvers, especially in the case of multiple simultaneous interactions. Combining gaze tracking with motion data has the potential to enable exploring those difficult questions, previously left unanswered by considering only one of these two methods separately.

Declarations of interest

None.

Acknowledgements

We acknowledge the voluntary contribution of our participants.

Funding

This research has received funding from the University of Rennes 2 (Scholarship), France; the French National Research Agency, project Percolation (ANR-13-JS02-0008), France; and Brittany region, project SAD 2015-INTERACT (9297), France.

Appendix

The following section provides some detailed analyses of our paper on the role of gaze in locomotor collision avoidance. In our paper, we use a Pareto ranking to support our claim that gaze reveals important insights on how collision is avoided.

Most of these claims are based on the idea that it is necessary to interact with certain virtual walkers that yield a high risk of collision. To quantify this necessity to interact, both *DCA* and *TtCA* play a role, but it is unclear what the importance is of *DCA* and *TtCA* with respect to each other. Therefore, we used the Pareto ranking to combine the two measures without making a clear choice about their relative importance.

Rather than using the combination of *DCA* and *TtCA* in a Pareto ranking, it is of course also possible to compute the results from the paper based on *DCA* and *TtCA* alone. Although it was not the main focus of the paper, it is worth checking if one of these two criteria has more importance than the other, and if one of these two would be sufficient to support the same conclusion. Here, we present more detailed results to discuss these two questions concerning fixations (Figs. A1 and A2) and concerning fixations in relation to avoidance maneuvers (Fig. A3).

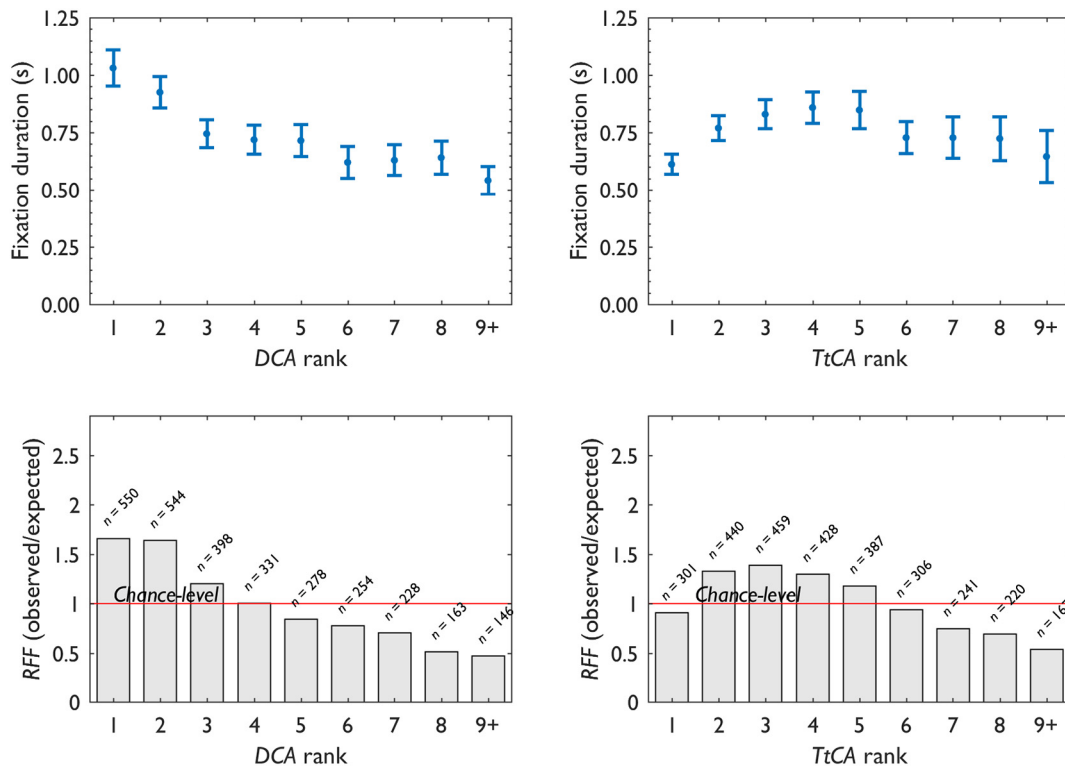


Fig. A1. The Fixation Duration (top) and RFF (bottom) based on the *DCA* (left) and *TtCA* (right) rank at the onset of the fixation with one walker allocated to each rank (NB, at each instant, there was only one walker in each rank). Whiskers indicate the 95%CI.

Fig. A1 shows that neither *DCA* nor *TtCA* alone provide a clear insight to why a walker may have been fixated upon. The *RFF* of the *DCA* ranking shows a trend that the higher ranks (i.e., closer to 1) were more fixated upon, but it is not as clear-cut as the Pareto ranking (see Fig. A2, top left). For *TtCA* it stands out that walkers with the lowest *TtCA* were certainly not fixated upon more often. This was likely because walkers with a very low *TtCA* have a high *DCA* and are therefore not posing a risk of collision.

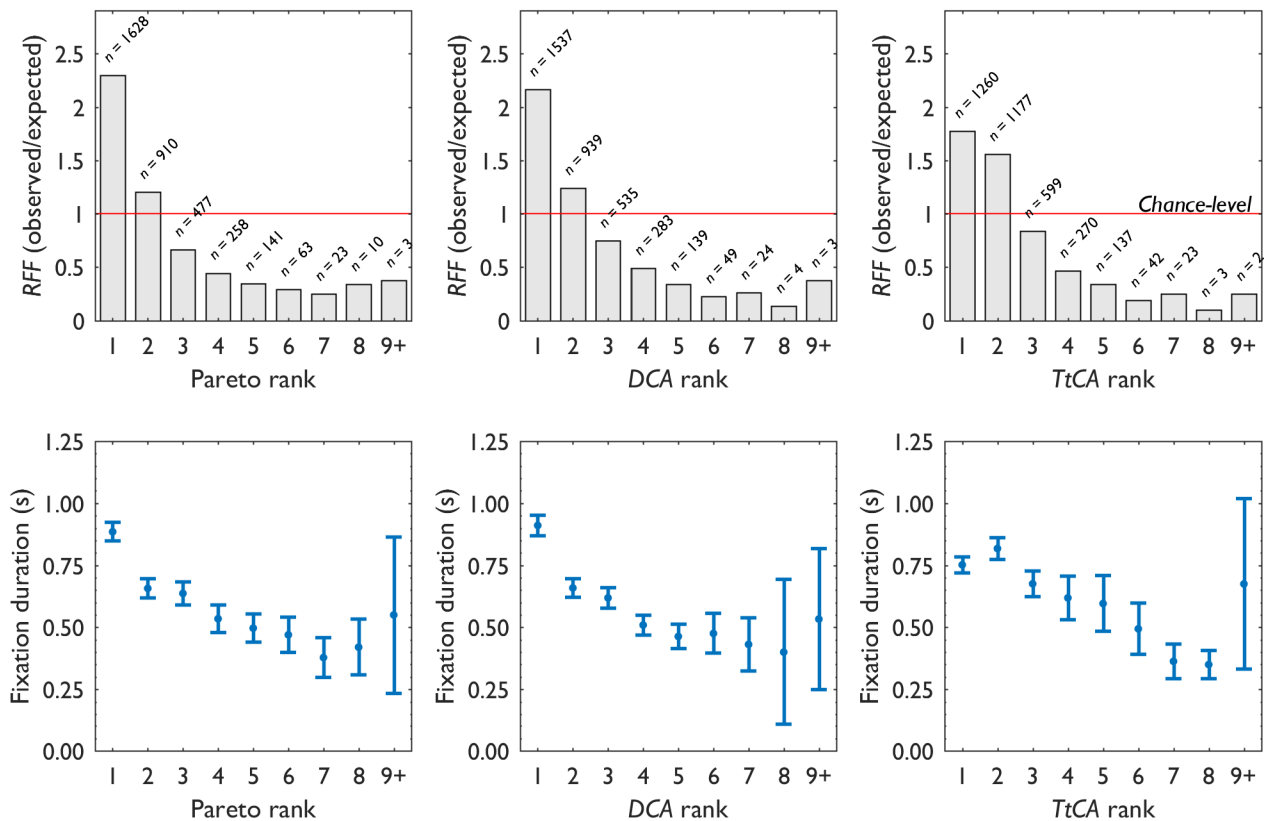


Fig. A2. The *RFF* (top) and Fixation Duration (bottom) based on the Pareto rank (left), *DCA* rank (mid) and *TtCA* rank (right) at the onset of the fixation. The number of walkers allocated to each *DCA* and *TtCA* rank is based on the number of walkers allocated to the Pareto ranks (i.e., the number varies per rank and per fixation). The figures on the left are similar to the figures presented in the main paper. Whiskers indicate the 95%CI.

The combined Pareto ranking (Fig. A2, left) clearly reveals that the participants' gaze was attracted to the walker that scored high on both *DCA* and *TtCA*. However, a completely honest approach would need to correct for the increased number of walkers allocated to each rank in a Pareto ranking compared to the *DCA* and *TtCA* ranking of Fig. A1. In Fig. A2 (middle and right column), the results of the main paper are presented for the *DCA* and *TtCA* rankings. It stands out that when the rankings are corrected for the number of virtual walkers, the difference between the Pareto and other rankings decreases.

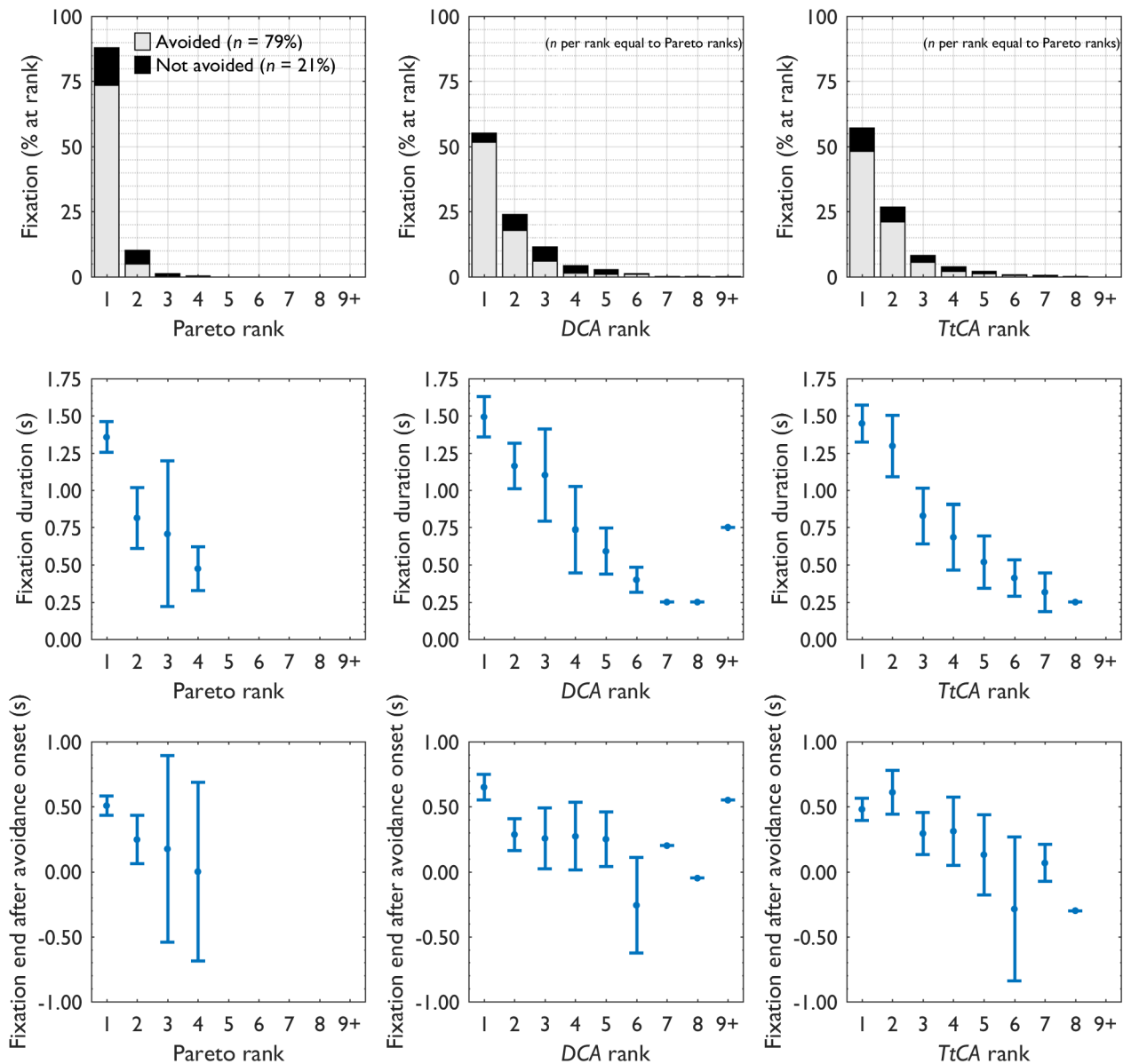


Fig. A3. The extended analysis of the fixations preceding an avoidance maneuver. The top row shows the fixation percentage at each rank (where the number of walkers in each rank is equal to the number of walkers allocated to the Pareto ranks). The middle row shows the duration of the fixation to each of the ranks. The bottom row shows the difference (in seconds) between the end of the fixation and the start of the avoidance maneuver (where positive values indicate that the fixation lasted until after the avoidance maneuver was initiated). The left column shows the Pareto ranking, the middle column the *DCA* ranking and the right column shows the *TtCA* ranking. Whiskers indicate the 95%CI.

Finally, **Fig. A3** provides more in-depth information about the fixation immediately before an avoidance maneuver was initiated. In this case, the effect of ranking method clearly holds. Note that in this analysis the number of walkers per rank was adjusted to the number of walkers per Pareto rank (i.e., the number varies per rank and per fixation). The Fixation durations show that – typically – higher ranked virtual walkers are fixated upon longer. Together with the trend that fixations end sooner after the maneuver was initiated in lower ranks, it seems that the fixation upon the highly ranked virtual walkers before an adaptation was initiated was really to guide the avoidance maneuver.

References

Andersen, G. J., & Kim, R. D. (2001). Perceptual information and attentional constraints in visual search of collision events. *Journal of Experimental Psychology: Human Perception and Performance*, 27(5), 1039.

Colombi, A., & Scianna, M. (2017). Modelling human perception processes in pedestrian dynamics: A hybrid approach. *Royal Society Open Science*, 4, 160561. <https://doi.org/10.1098/rsos.160561>.

Colombi, A., Scianna, M., & Alaia, A. (2016). A discrete mathematical model for the dynamics of a crowd of gazing pedestrians with and without an evolving environmental awareness. *Computational and Applied Mathematics*, 36, 1113–1141. <https://doi.org/10.1007/s40314-016-0316-x>.

Couzin, I. D., & Krause, J. (2003). Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior*, 32, 1–75.

Croft, J. L., Button, C., & Dicks, M. (2010). Visual strategies of sub-elite cricket batsman in response to different ball velocities. *Human Movement Science*, 29(5), 751–763.

Croft, J. L., & Panchuk, D. (2017). Watch where you're going? Interferer velocity and visual behaviour predicts avoidance strategy during pedestrian encounters. *Journal of Motor Behavior*, 1–11.

Dicks, M., Clashing, C., O'Reilly, L., & Mills, C. (2016). Perceptual-motor behavior during a simulated pedestrian crossing. *Gait & Posture*, 49, 241–245.

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. *Computer*

- Graphics Forum, 36(2), 337–348.
- Helbing, D., & Molnar, P. (1995). Social force model for pedestrian dynamics. *Physical Review E*, 51(5), 4282.
- Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *The Journal of Neuroscience*, 29(19), 6234–6238.
- Keller, A. A. (2017). Pareto optimality. *Multi-objective optimization in theory and practice I: Classical methods* (pp. 46–66). Sharjah, UAE: Bentham Science Publishers.
- Knorr, A. G., Willacker, L., Hermsdörfer, J., Glasauer, S., & Krüger, M. (2016). Influence of person-and situation-specific characteristics on collision avoidance behavior in human locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, 42(9), 1332.
- Lemercier, S., Jelic, A., Kulpa, R., Hua, J., Fehrenbach, J., Degond, P., ... Pettré, J. (2012). Realistic following behaviors for crowd simulation. *Computer Graphics Forum*, 31(2), 489–498.
- Lynch, S. D., Kulpa, R., Meerhoff, L. A., Pettré, J., Crétual, A., & Olivier, A.-H. (2017). Collision avoidance behavior between walkers: Global and local motion cues. *IEEE Transactions on Visualization and Computer Graphics*, 1–11 PP(99).
- Marigold, D. S., & Patla, A. E. (2007). Gaze fixation patterns for negotiating complex ground terrain. *Neuroscience*, 144, 302–313.
- Moussaïd, M., Guillot, E. G., Moreau, M., Fehrenbach, J., Chabiron, O., Lemercier, S., ... Theraulaz, G. (2012). Traffic instabilities in self-organized pedestrian crowds. *PLoS Computational Biology*, 8(3), e1002442. <https://doi.org/10.1371/journal.pcbi.1002442>.
- Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., & Theraulaz, G. (2010). The walking behavior of pedestrian social groups and its impact on crowd dynamics. *PLoS One*, 5(4), e10047.
- Mummenma, L., Hyönä, J., & Hietanen, J. K. (2009). I'll walk this way: Eyes reveal the direction of locomotion and make passersby look and go the other way. *Psychological Science*, 20(12), 1454–1458.
- Olivier, A.-H., Bruneau, J., Kulpa, R., & Pettré, J. (2018). Walking with virtual people: Evaluation of locomotion interfaces in dynamic environments. *IEEE Transactions on Visualization and Computer Graphics*, 24(7), 2251–2263.
- Olivier, A.-H., Marin, A., Crétual, A., Berthoz, A., & Pettré, J. (2013). Collision avoidance between two walkers: Role-dependent strategies. *Gait and Posture*, 38(4), 751–756.
- Olivier, A.-H., Marin, A., Crétual, A., & Pettré, J. (2012). Minimal predicted distance: A common metric for collision avoidance during pairwise interactions between walkers. *Gait & Posture*, 36(3), 399–404.
- Ondřej, J., Pettré, J., Olivier, A.-H., & Donikian, S. (2010, July). A synthetic-vision based steering approach for crowd simulation. *ACM Transactions on Graphics*, 29(4), 123.
- Patla, A. E. (1997). Understanding the roles of vision in the control of human locomotion. *Gait & Posture*, 5(1), 54–69.
- Rio, K., Dachner, G., & Warren, W. H. (2018). The local neighbourhood underlying collective motion in human crowds. *Proceedings of the Royal Society B*, 285, 20180611. <https://doi.org/10.1098/rspb.2018.0611>.
- 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.
- Salvucci, D. S., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the 2000 symposium on eye tracking research & applications* (pp. 71–78). ACM.
- Schwarzkopf, S., von Stülpnagel, R., Büchner, S. J., Konieczny, L., Kallert, G., & Hölscher, C. (2013). What lab eye-tracking tells us about wayfinding: A comparison of stationary and mobile eye-tracking in a large building scenario. *Proceedings of the 1st international workshop*. (pp. 31–36).
- Van den Berg, J., Guy, S., Lin, M., & Manocha, D. (2011). Reciprocal n-body collision avoidance. *Robotics research*. 70. *Robotics research* (pp. 3–19).
- Warren, W. H., Jr. (1998). Visually controlled locomotion: 40 years later. *Ecological Psychology*, 10(3–4), 177–219.