Collective Dynamics

Exploring the Dynamic Relationship between Pushing Behavior and Crowd Dynamics

Ezel Üsten[1,2] · Jette Schumann[1] · Anna Sieben[3]

[1] Civil Safety Research, Forschungszentrum Jülich, Jülich, Germany
E-mail: e.uesten@fz-juelich.de, j.schumann@fz-juelich.de
[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany
[3] Department of Cultural and Social Psychology, University of St. Gallen, St. Gallen, Switzerland
E-mail: anna.sieben@unisg.ch

Received: 12 July 2023 / Last revision received: 24 September 2023 / Accepted: 27 September 2023
DOI: 10.17815/CD.2023.143

Abstract Crowds, subjects of considerable complexity, have been extensively studied both as homogeneous entities and as collective sums of individual movements in various studies. However, crowd models, being grounded in physics, are limited in terms of incorporating psychological perspectives on individual behavior. Building upon the premise that crowd behavior is heterogeneous and dynamic, particularly in bottleneck scenarios, this study aims to explore the nuances of forward motion. Adopting the category system proposed by Lügering et al. [1] (consisting of the following categories: strong pushing, mild pushing, just walking, and falling behind), this paper investigates the circumstances and locations where pushing or non-pushing behaviors arise, intensify, or cease within crowds approaching bottlenecks. The study utilized 14 video materials obtained from previous laboratory pedestrian experiments to examine the spatial characteristics of forward motion and pushing behavior in relation to corridor widths and varied motivational instructions. Two trained raters independently annotated these videos, achieving satisfactory inter-rater agreement (KALPHA = .65), and a joint dataset was then created for each video. These videos consisted of both high (7 videos) and low (7 videos) motivation scenarios. The importance of corridor width was also considered: four videos featured a 5.6 m. width, another four featured a 4.5 m. width, and the remaining videos displayed widths of 3.4 m., 2.3 m., and 1.2 m. twice. Our findings suggest a tendency for increased pushing behavior or an increase in the categories as individuals approach the bottleneck, regardless of the width of the corridor or the motivational instruction. Furthermore, non-pushing behaviors were predominantly observed in the areas farther away from the bottleneck. A noticeable trend was observed in high motivation scenarios, which generally exhibited more instances of pushing behavior. The effect of corridor width indicated that, in certain cases, pedestrians who push in wider corridors experience faster access to the bottleneck. However, this effect is less significant in narrower widths. Additionally, temporal analyses indicated that category increases were most prominent in the initial quarter of the experiments, although other peak points were also observed. Calculations of the mean category values for each second revealed three distinct patterns: stability over time, a consistent slow decrease, and an initial increase followed by a decrease.

Keywords Pushing behavior · forward motion · crowd dynamics · rating system · crowd psychology · observation method

1. Introduction

Crowds are a common sight in everyday life, and can be found in public spaces such as streets, parks, and transportation hubs, as well as at events such as concerts, sporting events, and festivals. Dense crowds form when the available space is insufficient to accommodate the flow of pedestrians, leading people to gather in close proximity. However, in the absence of clear guidance on queueing, crowds are susceptible to unfair behaviors such as pushing by pedestrians eager to access exits quickly [2, 3]. Pushing behavior can take many forms, ranging from forceful and aggressive actions to mild and subtle movements. It is particularly common in scenarios where there is only one main entrance or exit, such as event entrance systems. Furthermore, pushing behavior is influenced by social context, including cultural norms and individual motivations. This complexity of factors can affect where and when pushing behavior occurs. Therefore, this paper aims to examine the spatial and temporal properties of pushing behavior in entrance scenarios, using an observation method and a motivational perspective that focuses on individual behavior for early access.

The importance of studying pushing behavior derives from the understanding that it poses a significant safety risk for the purpose of gaining early access in crowded environments. Pushing behavior can be particularly dangerous when it involves pushing others from behind, creating a dense environment [4, 5] that can be hazardous to those being pushed [6]. This can be particularly dangerous in extreme cases, where the felt pressure of the push could be fatal. Moreover, pushing behavior can intensify emergency situations, such as evacuations, where the priority is to get everyone out as quickly and safely as possible [7]. In these scenarios, confusion caused by pushing can hinder the evacuation process. Even in non-threatening situations, such as laboratory experiments with bottleneck setups, pushing behavior is often considered inappropriate and unfair [5]. It can also disrupt the overall speed of the crowd [8].

However, pushing behavior is not a static or consistently present behavior [1], and it can vary in strength, leading to different effects [9]. The likelihood of pushing may vary depending on the social context and norms at play [5]. In situations where the environment promotes unity and helping others, people may be less keen to push and more likely to walk together coherently [10, 11]. Furthermore, pushing behavior can change over time as the motivations of pedestrians shift. It can increase, decrease, disappear, or appear throughout the pursuit of a goal [1], as individual motivations are dynamic and constantly evolving. Therefore, pushing behavior is also a dynamic behavior that is influenced by the motivations of pedestrians.

Previous studies have shown that pushing behavior can be classified by rating individual behavior from crowd videos. Lügering et al. [1] developed a forward motion category system that includes the following categories: strong pushing, mild pushing, just walking, and falling behind. This system includes detailed definitions of pushing and forward motion, including the means by which the behavior is carried out, such as using elbows, protecting oneself, filling gaps, accelerating, and so on. This system allows researchers to rate individual pushing behavior (or the absence of it) from a starting point until the pursuit of a goal is completed. These ratings can be combined with pedestrian trajectories to see the individual effects throughout the sequence. Moreover, the spatial and temporal characteristics of individuals within crowds, as obtained from their trajectories, can be related to their forward motion aspects, allowing for a deeper exploration of their relationship with pushing categories.

The purpose of this paper is to utilize the forward motion category system developed by Lügering et al. [1] to investigate the relationship between pushing behavior and the spatial and temporal properties of the crowds. To achieve this objective, we analyze pushing occurrences from laboratory studies using trajectory data extended by pushing states and link them with crowd properties such as time and distance to the goal. By doing so, we aim to explore the motivational aspects of forward motion behavior, as well as individual pushing occurrences on a large scale, in order to identify collective patterns. While previous studies have shown that pushing behavior can increase the crowd density [4, 5], a more comprehensive investigation has not been conducted to explore the relationship between pushing behavior, pedestrian motivation, and physical crowd properties in detail. The utilization of the pushing category system as a tool will enable us to observe and measure the causal relationship between what pedestrians do to gain faster access and how the crowd changes, or vice versa. This understanding can subsequently be employed to inform the development of crowd management strategies and enhance safety in crowded environments.

2. Method

2.1. Empirical Material

Fourteen videos from laboratory studies (Fig. 1) investigating the movement in bottleneck platforms were selected for the current study from the Pedestrian Dynamics Data Archive [12]. These videos, along with the trajectory data of the pedestrians, were originally recorded for interdisciplinary experiments conducted at the University of Wuppertal [5, 13] and were stored in the Pedestrian Dynamics Data Archive for future research. Out of the original 24 videos available from this experiment series, 14 were selected based on a participant number threshold of n >40, following the approach of the original study [5]. Videos with fewer than 40 participants had shorter durations, which were considered unlikely to offer meaningful results. Additionally, below that participant threshold, only a few instances of pushing behavior were observed. Experimental run numbers and the order were opted to be maintained as they are due to convenience and consistency with the original research.

Figure 1 Exemplary screenshot from the experiments.

All experiments included motivational instructions for the participants to influence their behavior during the experiment. The high motivation instruction consisted of a concert context where participants were told they needed to reach the bottleneck quickly to get good seats. The low motivation instruction told participants that they would have many good seats available, but it was still good to reach the bottleneck earlier. The selected videos were equally divided between these motivation priming conditions, with seven high motivation and seven low motivation videos.

The corridor width was another property taken into consideration. There were five different corridor widths, and it was thought that these different widths would produce different effects or be affected differently in terms of pushing behavior. The widths were as follows: 1.2 m., 2.3 m., 3.4 m., 4.5 m., and 5.6 m. There were two videos each for the 1.2 m., 2.3 m., and 3.4 m. widths, and four videos each for the 4.5 m. and 5.6 m. widths. The number of videos for each width was equal in terms of instructed motivation. Tab. 1 presents a summary of all experimental runs, providing relevant information.

Run Number	Motivation	Corridor Width	Number of Pedestrians	Experiment Time	Flow Time	Crowd Time
030	High	5.6 m.	75	65 s.	0-3 s.	4-65 s.
040	Low	5.6 m.	75	66 s.	0-5 s.	6-66 s.
050	High	4.5 m.	42	38 s.	0-4 s.	5-38 s.
060	Low	4.5 m.	42	41 s.	0-7 s.	8-65 s.
110	High	1.2 m.	63	53 s.	-	0-53 s.
120	Low	1.2 m.	63	65 s.	-	0-65 s.
150	High	5.6 m.	57	57 s.	0-4 s.	5-57 s.
160	Low	5.6 m.	57	56 s.	0-5 s.	6-56 s.
230	High	2.3 m.	42	32 s.	0-5 s.	6-32 s.
240	Low	2.3 m.	42	38 s.	-	0-38 s.
250	High	4.5 m.	42	33 s.	0-3 s.	4-33 s.
260	Low	4.5 m.	42	39 s.	-	0-39 s.
270	High	3.4 m.	67	59 s.	0-7 s.	8-59 s.
280	Low	3.4 m.	67	67 s.	-	0-67 s.

Table 1 Experimental parameters across distinct runs are summarized along with their unique identification numbers.

2.2. Rating Procedure

The rating process was done using the PeTrack software [14]. The trajectories of each pedestrian were first captured using PeTrack during previous experiments, and these trajectories along with their corresponding videos were used in this study to rate the pedestrians' pushing behavior throughout their movement to the bottleneck. The rating was done at specific time points, using frame numbers (1 second is equal to 25 frames). When a participant was selected, their behavior was rated throughout the duration of the behavior. If the behavior changed, the rating was also changed and annotated at the corresponding time point using the software. This process was completed for each participant from the start of the experiment until they reached the bottleneck. After the annotation was finished, a txt file containing all the ratings and the pedestrian coordinates was extracted from the software.

An ordinal four-stage category system developed by Lügering et al. [1] was used to annotate pedestrians in the bottleneck setup experiment videos. This system consists of four inclusive categories for annotating all the behaviors that can be seen throughout the experiments. There are two categories for pushing and two categories for non-pushing behaviors. The pushing categories consist of mild and strong pushing behaviors, with mild pushing including mostly active behaviors such as overtaking and filling gaps without excessive force, and strong pushing including intense pushing behavior. The non-pushing categories are going with the flow and falling behind from the crowd. The full category names are as follows: (1) falling behind, (2) just walking, (3) mild pushing, and (4) strong pushing. All the participants within all frames were rated with these four categories throughout the session of each experiment.

To account for momentary changes in behavior, as suggested by Lügering et al. [1], the two-to-three-second rule (50-to-75 frames) was adopted. If a behavior persists for two to three seconds, it is rated accordingly. However, if a behavior lasts for less time, it is considered unintentional or accidental, as the actor would need at least a couple of seconds to comprehend and act in relation to their environment. The rater's overall comprehension of behavior was also guided by the suggestions of Lügering et al. [1], and the minimum unit of measurement for complex behavior was set to one second (25 frames), instead of one frame, to ensure full comprehension by the rater.

Overall, two trained raters annotated the fourteen videos separately. These videos were then analyzed to assess the reliability assumption and to determine the level of agreement between the two raters. Joint ratings were then created based on the original files. The videos were watched again with a focus on disagreements, and a decision was made on which ratings should be used for the final data.

2.3. Measures

Spatial and temporal measures were employed to accurately capture the distribution of pushing behavior categories in bottleneck scenarios. The actual distance and time data (measured in meters and seconds) along with predefined divided areas were used to conduct the analyses. The category data within these areas were plotted accordingly.

Four different analysis methods were utilized:

1. Time and distance relationship of pushing categories for each experiment.
2. Showcase of pushing category data in semi-circle areas for each experiment.
3. Showcase of pushing category data in semi-circle areas for each category.
4. Showcase of pushing category data in small (25 cm. x 25 cm.) square areas for each experiment.
5. Number of category increases and decreases, along with the mean pushing category values for each second in all experiments.

These analyses will be presented in the results section, labeled as follows: Time-distance trajectories (1), distance bins (2), category charts (3), heat maps (4), and time analyses (5).

3. Results

3.1. Reliability Analysis

To ensure the reliability of the data, two trained raters independently assessed all 14 videos, which included a total of 776 participants. The data of one rater comprised 1,003,050 frames, equivalent to 40,122 units of measurement (seconds). Due to the manual input of ratings, small time slippages of one or two seconds were prone to occur, so timing corrections were performed up to three seconds between raters. To evaluate the inter-rater reliability, Krippendorf's alpha (KALPHA) was used, as it has been suggested as a suitable method for analyzing the inter-rater reliability of ordinal data [1, 15]. An SPSS macro, developed by Hayes and Krippendorf [16], was utilized to calculate the KALPHA value. The analysis revealed a 75.6 percent overlap between the raters and a KALPHA value of .65, indicating a moderate level of agreement between the raters.

De Swert [17] suggested that a good inter-rater reliability limit would be over .80, but for highly complex data, a minimum level of .60 would be sufficient. Our study involved an extensive dataset comprising an exceptionally high number of analysis units. The categories used in the analysis relied on small, context-dependent behaviors that could be challenging to discern from a limited top-down view. Furthermore, we took into account behavioral shifts over time, and the analysis units were interdependent. For instance, if one observer noticed a shift from a non-pushing category to a pushing category (and e.g., adjusted the rating from 2 to 3), while the other evaluator rated the behavior differently, the rating discrepancy would not be limited to a single second but would extend to multiple seconds. These factors increased the complexity of our rating system and the potential for disagreement between the raters. Due to these factors, we believe that a KALPHA value of .65 is justified in terms of agreement between the raters. Although this value is lower than the ideal value proposed by De Swert, the factors discussed above indicate that a moderate level of agreement was still achieved, supporting the reliability of our data and findings.

3.2. Category Analyses

3.2.1. Time-Distance Trajectories

The first type of analysis is referred to as time and distance trajectories, which illustrates the progression of agents through the bottleneck in terms of time and distance (introduced by Sieben et al. [18]). These plots are divided into four subplots representing different categories, including "strong pushing," "mild pushing," "just walking," and "falling behind," for each run individually.

However, the initial plot of each experimental run, referred to as the bulk plot, consisted of data for all pushing categories. These categories were visually distinguished using different colors: green for falling behind, yellow for just walking, orange for mild pushing (see Fig. 2), and red for strong pushing. The subsequent plots illustrated individual pushing categories across all experimental runs (see App. A.1).

The bulk plots primarily showed that in high motivation runs, there were many participants who chose to engage in pushing behaviors, and these individuals were mostly successful in advancing forward, either by finding gaps or through pushing. This trend was particularly apparent during the initial phase of the experiments. The rapid progress in terms of distance without losing time was particularly evident in these plots for high motivation runs (see Fig. 3a). On the other hand, participants with non-pushing behaviors tended to wait longer and make less progress. In low motivation videos, there were fewer instances of pushing behaviors, resulting in a smaller visible effect, although it was still present (see App. A.1).

Additionally, the width of the corridor appeared to have an impact on the overall flow. Wider corridors, such as those with widths of 5.6 m or 4.5 m, often experienced congestion and high density, resulting in a less smooth flow. Participants using pushing behaviors were able to make faster progress, while those not using pushing means experienced slower advancement. In contrast, corridors with a width of 1.2 m, despite exhibiting many instances of pushing behaviors, showed minimal additional waiting as the flow persisted. It was observed that pushing behavior did not have a prominent effect on gaining faster ac-

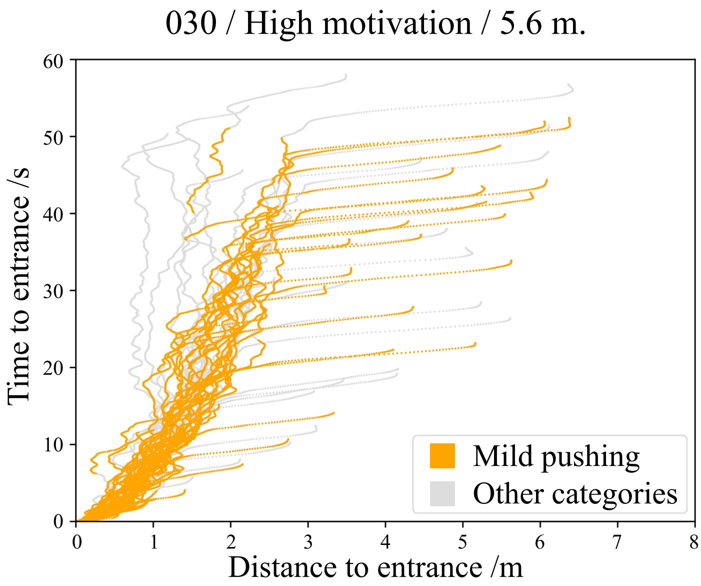


Figure 2 Exemplary time-distance plot for data in the "mild pushing" category. The mild pushing data is highlighted in orange, while the gray lines depict data from all other categories.

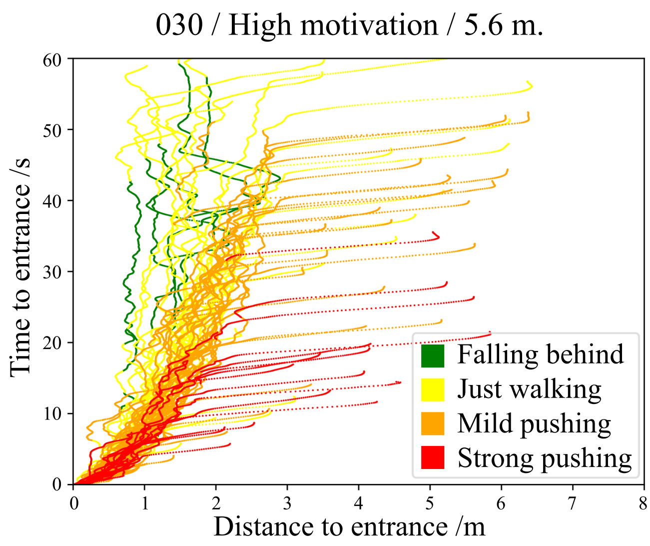
cess in narrow corridors; participants in different categories had similar time and distance periods (see Fig. 3b).

3.2.2. Distance Bins

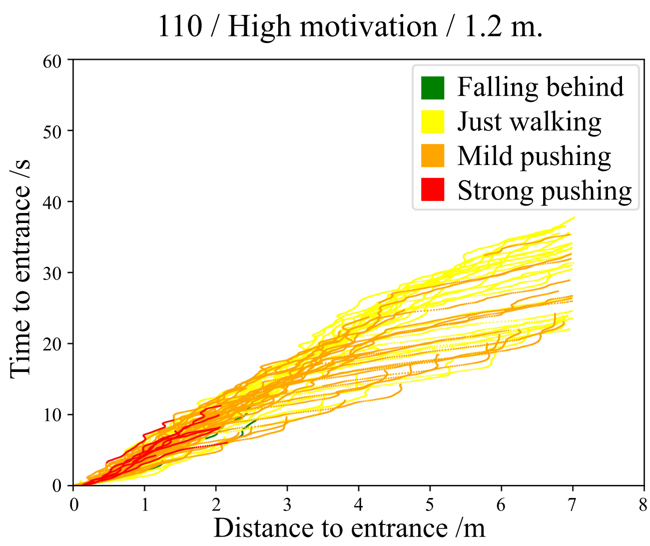
The second type of analysis involved showcasing the category data within predefined areas located in front of the bottleneck. These areas were shaped as semi-circles, positioned at half-meter intervals from the entrance of the bottleneck (see Fig. 4). However, the data from the experiment runs were collected without considering their temporal aspect. All the category data within each second was aggregated and added to the respective areas, without considering the specific frames or seconds in which it was generated. Additionally, some bins differed in size across experiments due to varying corridor widths.

Initially, the objective was to analyze the spatial distribution of the four pushing categories within each bin using colored bars. However, due to the varying sizes of the bin spaces and the unidirectional movement of pedestrians (resulting in some bins containing more data than others, such as those closest to the bottleneck), the data distribution became imbalanced. Consequently, two different types of bin plots were generated to represent the same data: "absolute frequency" plots, which display the raw data without balancing, and "relative frequency" plots, which present the data as a percentage of the total within their respective bins.

Furthermore, we discovered that the relative frequency plots were unintentionally misleading for the bins located further away from the bottleneck, where only a few seconds of data were recorded. During those initial seconds, pedestrians quickly rushed through the bottleneck, creating a high-density environment. In the absolute frequency plots, these bins accurately reflected the insignificance of the data, but in the relative frequency plots,



(a)



(b)

Figure 3 Exemplary plots of all pushing category data from wide and narrow width corridors.

they appeared to contain substantial amounts of pushing or non-pushing data, which was misleading. Upon further investigation, we realized that these problematic bins in the relative frequency plots predominantly consisted of "flow" data rather than "crowd" data, hence they were fundamentally different from each other. In these time periods, pedestrians were free to move further without encountering any junctions or congestion, at least for a few seconds before the crowd formed. Therefore, we categorized these time periods as "flow" and "crowd" data in the plots. The "flow" data included only the first few seconds of the respective videos, spanning from 4 to 9 seconds.

Our time and distance plots, as showcased in Fig. 2 and Fig. 3, noticeably display

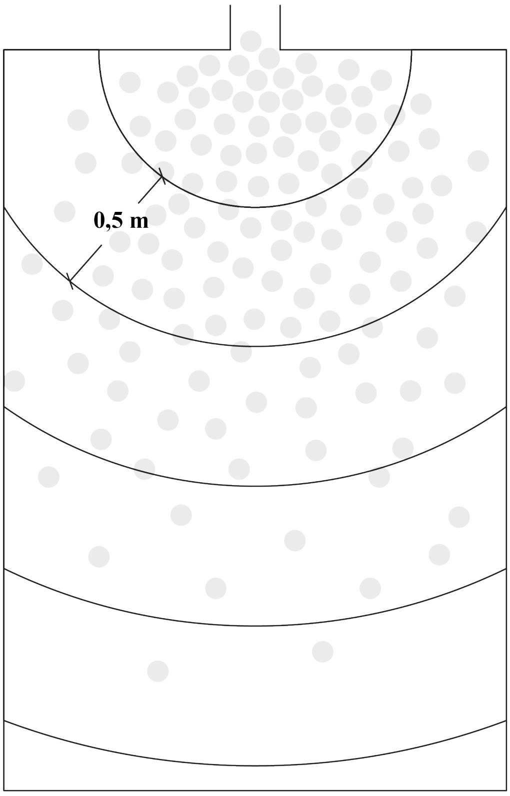


Figure 4 Predefined areas for the bins.

the distinction between "flow" and "crowd" data. Initially, many pedestrians are able to progress rapidly in the first few seconds, representing the "flow" phase. Subsequently, pedestrians begin to advance in a less robust phase, primarily due to the dense environment caused by the intersecting pathways of all pedestrians toward the bottleneck. This later phase, referred to as the "crowd" phase, can be observed by the clustering of the trajectories after the conclusion of the flow phase.

Regarding the data itself, the flow phase, which represents the first few seconds before the crowd formation, showed a balanced distribution of pushing and non-pushing categories, although there was a tendency for more pushing categories among pedestrians in the first few meters rather than non-pushing categories. This presence of pushing categories was observed in almost half of the runs, with a balanced distribution between low and high motivation conditions, and between the different widths. However, because the data exhibits a balanced distribution among the categories and conditions, and considering that we have data for only the first few seconds of the experiments, we have opted not to present any additional plots for the flow phase.

The absolute frequency crowd data provides a clear picture of the presence of pushing categories as pedestrians get closer to the bottleneck, and this presence decreases as they move further away. However, due to the unequal spacing of the bins and the unidirectional flow of the crowd, there is more data in the largest semi-circle space when the crowd is formed, typically between 1 m and 2 m. The decrease between 0 m and 1 m is solely due to the smaller space of the first semi-circle. Percentage-wise, the increase in pushing categories is evident throughout the bottleneck, as seen in the relative frequency plots (see App. A.2). Mild pushing is the most prominent category, but strong pushing also shows a significant increase. The presence of just walking, along with the less visible falling behind, decreases as pedestrians approach the bottleneck area, particularly in the high motivation runs. In the low motivation runs, the same observations hold true if there are instances of pushing behavior. However, if there is little or no pushing behavior, an increase in non-pushing categories can be observed (from "falling behind" to "just walking"). These patterns hold true for the different corridor widths, and are clearly observable in almost all of the runs. Fig. 5 aims to demonstrate the difference between absolute (Fig. 5a) and relative (Fig. 5b) frequency plots, along with the aforementioned observations. For clarity, the appendix section exclusively includes the relative frequency (crowd) plots, as they offer better interpretation (see App. A.2).

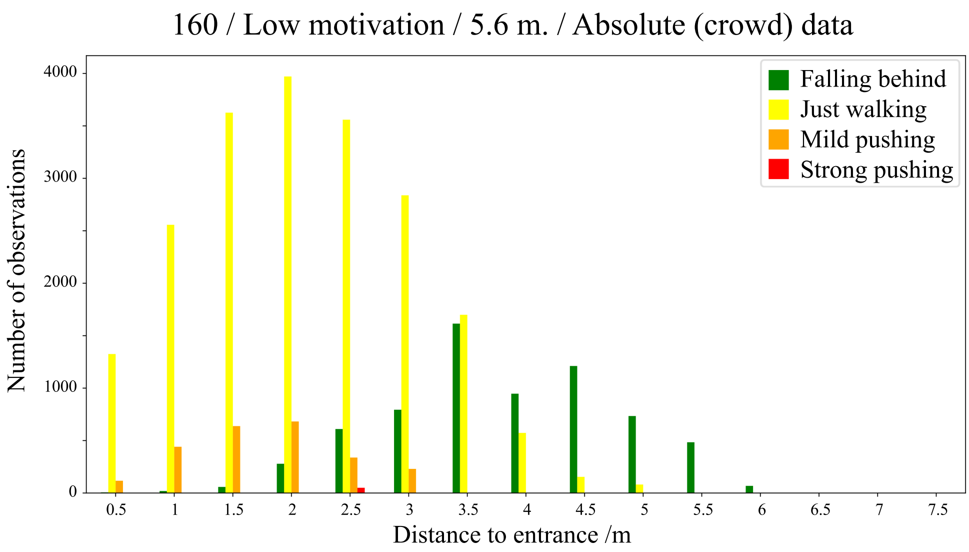
3.2.3. Category Charts

The "crowd" data, which consists of the number of observed categories within each distance bin (absolute frequency), was also utilized in the category charts, but this time grouped by pushing categories. Although the addition of flow data wouldn't lead to a misleading interpretation due to the use of absolute frequency data, it was opted not to be included to maintain consistency across the spatial analysis sections. The objective was to gain insights into the distance distribution to the bottleneck for each individual pushing category. Additionally, the runs were further divided into high motivation and low motivation categories to examine potential differences in the distance distribution for each pushing category between these two types of runs.

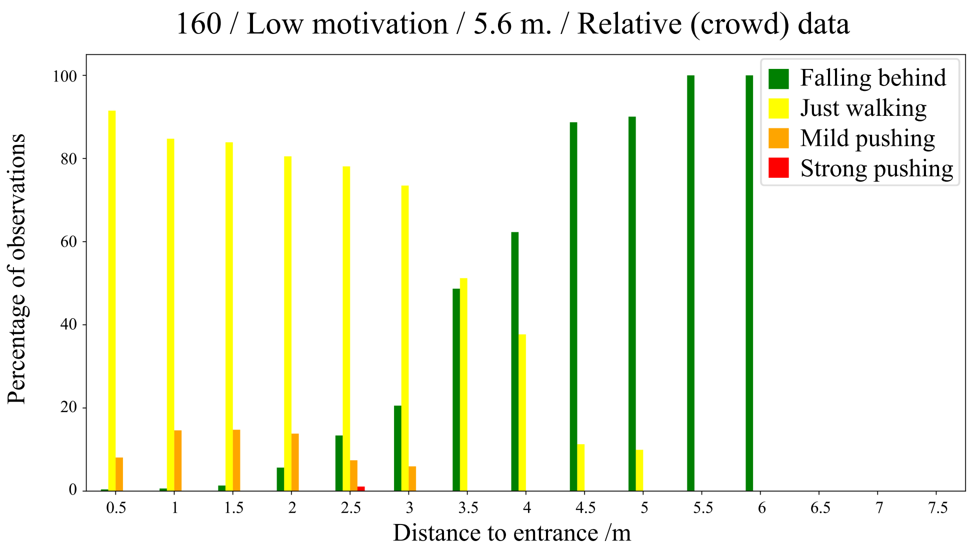
In all of the analyses, it is evident that the categories are predominantly clustered within the first three meters, which is likely the threshold where the crowd formation begins. Prior to reaching three meters, there is a slight increase in all categories across different runs, but this increase becomes more pronounced after the threshold is reached (see App. A.3).

In the first meter, which contains less data, there appears to be a decrease in all categories. However, it is important to note that this should not be interpreted as an actual decrease, as discussed earlier and disproved in the relative frequency distance bins. Additionally, it is worth mentioning that just like the pushing categories, the non-pushing categories also exhibit a significant increase in data as pedestrians move closer to the bottleneck. However, we need to interpret this with caution since their proportion decreases overall, as shown in the relative frequency distance plots. Nevertheless, the absolute frequency data of all categories demonstrates a similar trend across all cases.

Furthermore, in the high motivation runs, all categories exhibit a peak point that gradually approaches as the intensity of the pushing behavior increases. For instance, the peak point of "mild pushing" (see Fig. 6a) is slightly further away from the peak point of "strong pushing" (potentially peaking at the first bin) (see Fig. 6b), but still much closer compared to the peak point of non-pushing categories. On the other hand, the low motivation runs, which have fewer instances of pushing behavior, primarily show peak clusters



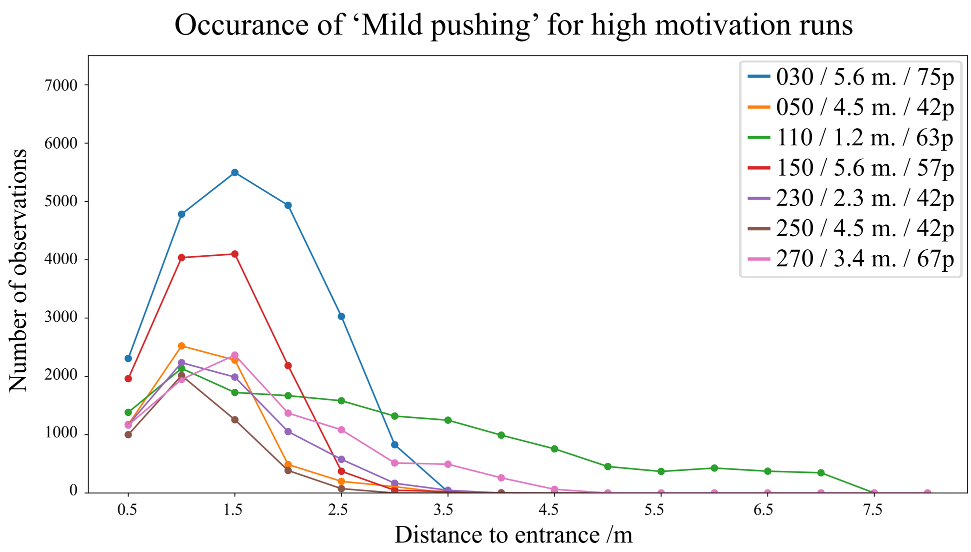
(a)



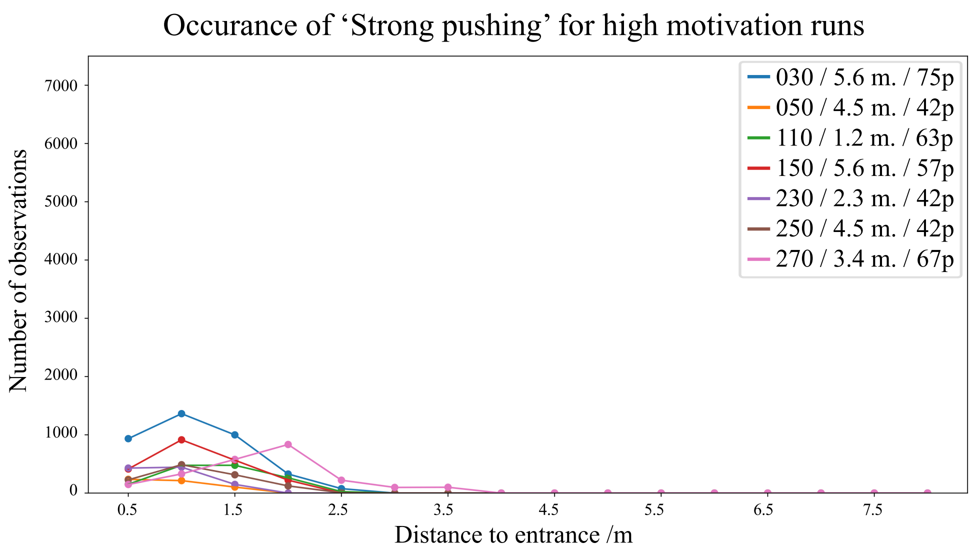
(b)

Figure 5 Exemplary absolute (a) and relative (b) frequency bin plots (crowd data). The X-axis of each plot represents ”distance to entrance” in meters (0 to 7.5 m., with each bin placed at 0.5 m. intervals), and the Y-axis represents the collected category data from all frames (a) and the percentage of category data (b) collected within each semi-circle, presented as bins. Each bin is color-coded according to category labels. Titles provide the necessary information about the experimental runs from which the data was selected.

for non-pushing categories. However, a gradual proximity for the active categories is still observable (see App. A.3 and Fig. 13).



(a)



(b)

Figure 6 Category charts for the "mild pushing" (a) and "strong pushing" (b) categories (absolute frequency data) across all seven high motivation runs are presented. The X-axis of each plot represents "distance to entrance" in meters (ranging from 0 to 7.5 m., with each bin set at 0.5 m. intervals), while the Y-axis represents the collected category data from all frames within each semi-circle. Distinct colors correspond to various experimental runs. The labels provide insight into the experimental setups, detailing corridor width and pedestrian count for each run.

3.2.4. Heat Maps

For the next analysis, a heat map-like visualization was utilized to provide a clearer representation of the spatial distribution of the pushing categories. Instead of using semi-circles, the bottleneck platform was divided into equal square sections, with each cell measuring 25cm by 25cm. The category data was used as ordinal numbers, with a coding of 4 for strong pushing, 3 for mild pushing, 2 for just walking, and 1 for falling behind, as used during the actual rating process. The color of each square in the heat map was determined based on the mean values of the pushing category data within that cell, using a color scale ranging from 1 (falling behind) to 4 (strong pushing). Once again, the temporal aspect of the data was disregarded, and all the “crowd” phase data within each second or frame were collected for analysis.

The data presents a clear distribution pattern of the categories, with the pushing categories predominantly observed in the proximity of the bottleneck. This observation is consistent across all high motivation runs (see Fig. 7a), as the closer cells tend to contain a higher concentration of pushing data. Similar results are observed in the low motivation runs (see Fig. 7b), although the visibility of this pattern varies due to the lower occurrence of pushing categories. However, there is still an increase within the non-pushing categories from “falling behind” to “just walking”. Thus, regardless of whether it involves pushing or non-pushing behavior, there is an overall increase in the categories as pedestrians approach the bottleneck. Some anomalies can be observed in the distant cells, showing an increase in pushing; however, this occurrence is likely due to the smaller amount of data available for those cells. Fig. 7 displays the general tendency of these observations in all the heat maps (see App. A.4 for all heat maps).

3.2.5. Time Analyses

Lastly, frequency-time analyses were conducted. These analyses primarily considered temporal aspects, disregarding spatial information, as the previous spatial analyses did not account for time. Two distinct types of plots were created for this analysis: Frequency charts illustrating category increases and decreases, and mean category value charts.

For the first plot type, all category increases (e.g., from 2 to 3: from just walking to mild pushing) and decreases were counted for each experiment individually, with consideration of the second at which they occurred. Spatial aspects, as well as the specific type of increase (e.g., from 1 to 2, 2 to 3, 2 to 4, etc.), were disregarded. Each increase and decrease was tallied and presented in a frequency chart over the course of the experiment period, with data aggregated in three-second intervals. The goal was to identify the time points at which these changes occurred and examine potential patterns that might correspond to psychological or social crowd dynamics. Fig. 8 illustrates the variations in increase and decrease counts for selected high and low motivation runs (refer App. A.5 for all the plots).

In general, the figures reveal a tendency for an increase in the pushing category after the first three-second interval, corresponding to seconds between 3 and 6. This increase may be attributed to the previously mentioned distinction between the “flow” and “crowd” phases. During the “flow” phase, pedestrians move relatively quickly and freely, but as the initial seconds pass, they may have a tendency to increase their category as they seek to take advantage of the situation. While this outcome was reasonable and expected, it’s worth noting that this wasn’t the only peak in category increases observed throughout the

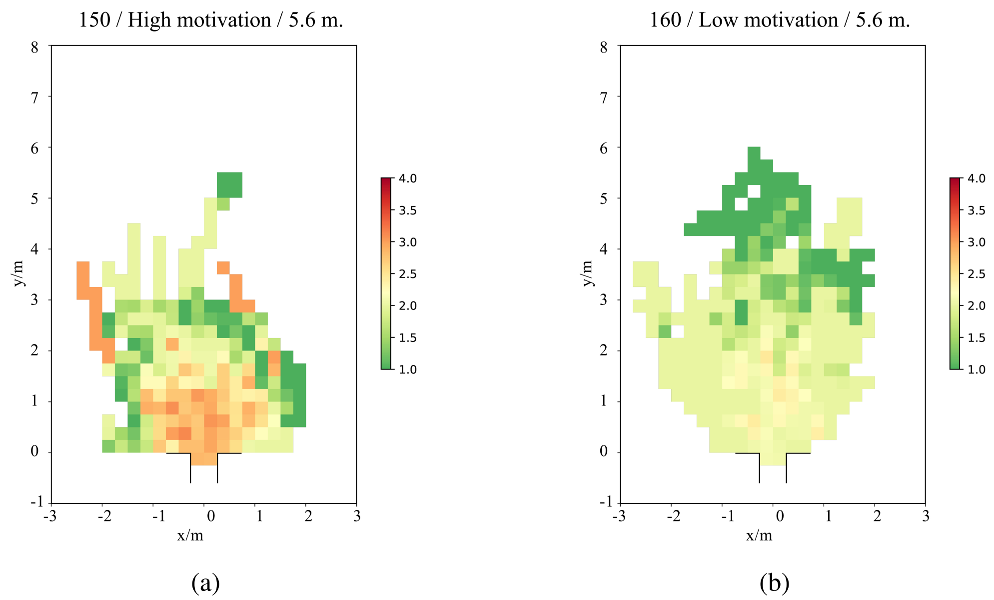


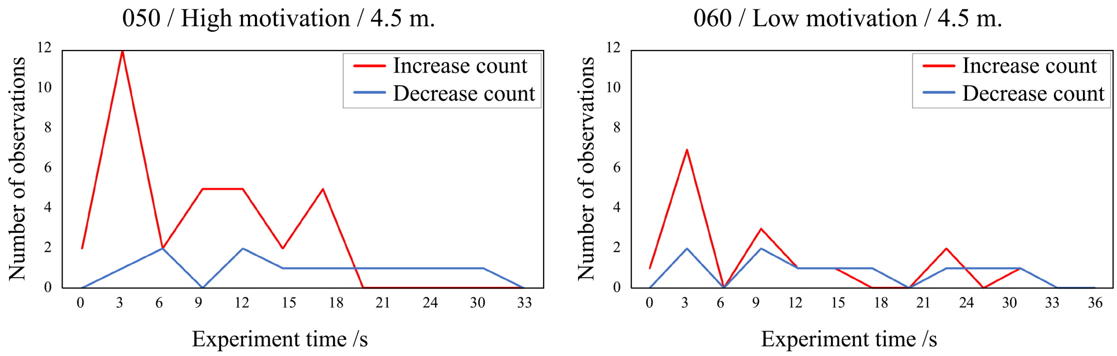
Figure 7 Exemplary heat maps from high motivation and low motivation runs. Figures illustrate the mean category value collected in 25 cm. to 25 cm. cells. The X-axis represents the width of a hypothetical bottleneck platform (-3 m. to 3 m., with a total width of 6 m.), while the Y-axis represents the length of the bottleneck platform (8 m.). The data flow, or pedestrian flow, is depicted from top to bottom. The collected data in each cell are color-coded based on the color scale derived from the category labels. Red represents "Strong pushing," orange represents "Mild pushing," yellow represents "Just walking," and green represents "Falling behind." Intermediate colors from the color scale are also used.

runs. For most runs, the second peak point in category increases occurred during later periods, often in the middle of the experiment, which hints at a pattern throughout the runs (see App. A.5, additionally, for a percentage-based version, refer to App. A.6).

Regarding the differences between categories, there was a general tendency for high motivation runs to exhibit higher frequency in category increases when compared to low motivation runs. Concerning decreases, on the other hand, while they were less frequent than increases, the main peak point for decreases also occurred in the earlier seconds, although not necessarily in the initial intervals. Lastly, corridor width does not appear to have a significant effect on category increases and decreases.

The second chart type involved tracking the mean category value throughout the experiments in three-second intervals. To accomplish this, all the category data were recorded for each second, the mean category value was calculated for each experiment individually, and then the results were combined into a single plot. The goal was to identify any patterns where the mean category value was higher or lower at specific time points, providing additional insights into temporal crowd dynamics. Fig. 9 illustrates this analysis by combining results from both high motivation (a) and low motivation (b) scenarios.

The figures reveal that, in general, high motivation runs exhibited higher average pushing category numbers compared to low motivation runs, and there were no distinct dif-

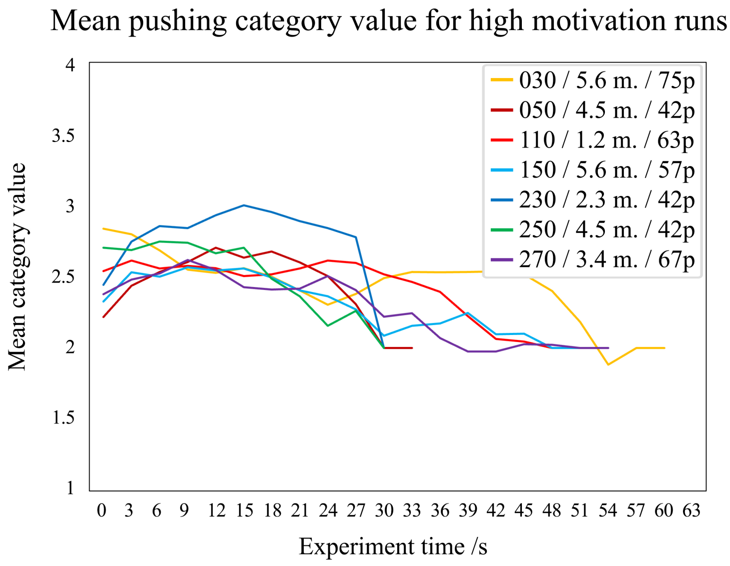


(a) (b)

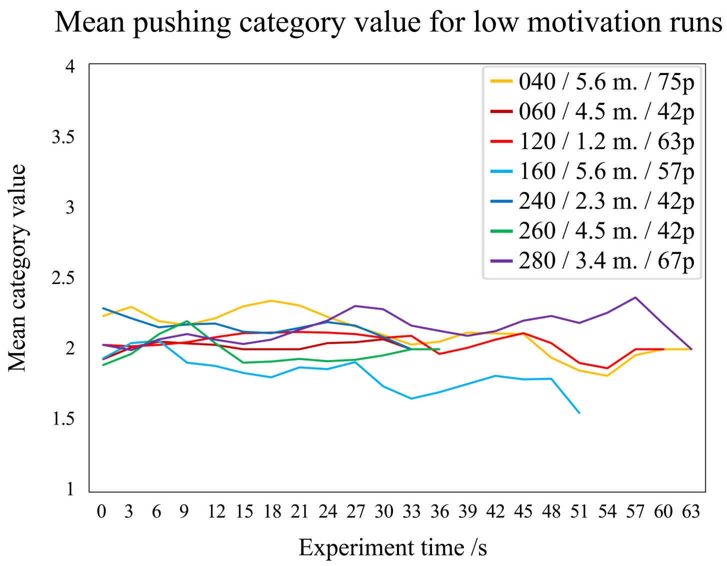
Figure 8 Exemplary time/frequency charts for pushing category increases and decreases in one high motivation (a) and one low motivation (b) run are displayed. These figures illustrate the category increases and decreases collected in each second, aggregated in three-second intervals. The X-axis of the plots represents the overall experiment time period in seconds, while the Y-axis represents the frequency of category increases (in red) and decreases (in blue).

ferences between different corridor widths. Within high motivation runs, two distinct patterns emerge: firstly, a curved trend with an initial increase followed by a decrease, and secondly, a consistent slow decrease over time. However, it's notable that the category increases discussed previously did not substantially impact the overall mean value data. The averages peaked around 10 to 20 seconds, even though the most prominent category increases occurred between 3 to 6 seconds. However, it's important to consider that the overall pedestrian number gradually decreased in each second due to pedestrians exiting the bottleneck over time. It's possible that pedestrians with increased categories exited when they were near the bottleneck, leading to their data not being counted after they left. This interpretation aligns with the previous spatial analyses, which showed that the categories increased as pedestrians got closer to the bottleneck.

For both types of plots, the initial intention was to combine all the data from the experiments. However, this approach had to be revised because the experimental runs had varying durations and different numbers of pedestrians in each second. To address the issue of varying pedestrian counts, we created an alternative version of the category increase and decrease charts. This version includes ratios calculated from increase and decrease frequency divided by the total pedestrian count on the platform for that specific second, resulting in a percentage representation (e.g., if there are 40 pedestrians in a specific second with 10 increases and 5 decreases, it is represented as .25 increase and .12 decrease, as shown in App. A.6). However, due to the differences in experimental durations and fluctuations in the data, this plot type was presented individually for each experiment. In contrast, the mean category value charts were shown collectively since they demonstrate clear and continuous averages of the means. While we believe this shouldn't pose an issue for interpreting the plots, it's important to consider that mean category values are calculated from different numbers of pedestrians in each second (e.g., second 1 with 60 participants, second 40 with 10 participants), with variations across all experiments.



(a)



(b)

Figure 9 Mean category value charts for high motivation (a) and low motivation (b) runs, including all the experiments, are displayed. All average calculations were conducted for each second, and the results are presented in three-second intervals. The X-axis of the plots represents the overall experiment time period in seconds, while the Y-axis represents the total ordinal pushing category data, ranging from 1 to 4. Distinct colors correspond to various experimental runs. The labels offer insights into the experimental setups, providing information on corridor width and pedestrian count for each run.

4. Discussion

The study at hand aims to investigate the different categories of pushing behavior in relation to the spatial and temporal properties of crowds. To obtain ratings for the pushing behavior categories, two trained raters independently annotated 14 videos of laboratory pedestrian experiments, ensuring a sufficient level of inter-rater reliability. Using this rating dataset, analyses were conducted to visualize the spatial dynamics, as well as certain aspects of the temporal dynamics of pushing behavior. These analyses facilitated the observation of how pushing behavior changes, forms clusters, increases, and decreases across different locations and time periods within the crowd.

There are several noteworthy findings from this research. Firstly, in almost every video, a clear pattern becomes apparent where the proportion of pushing behavior increases as individuals move closer to the bottleneck. Conversely, individuals further away from the bottleneck exhibit a lower frequency of pushing behavior and a higher proportion of non-pushing categories. This spatial division highlights the different behavioral dynamics among participants based on their position relative to the bottleneck. In a metaphorical manner, we refer to this observation as the "carrot effect," drawing from the English idiom "carrot and stick." This idiom, although more complex in its full context, signifies that a visible reward can increase an individual's motivation. In our case, pedestrians seem to be more engaged when they perceive the goal (the bottleneck) to be within closer reach. In the psychology and motivation literature, this phenomenon can be explained by the concept of "goal proximity", which suggests that as individuals get closer to their goals, the value and attractiveness of the task increase, subsequently enhancing their motivation [19–21]. Additionally, these findings are consistent with the expectancy and value concept of major motivation theories [22–24], as the expectancy of reaching the bottleneck continuously increases as individuals approach it, consequently amplifying their motivation. In light of these contextual factors, we interpret the observed increase in pushing behavior as a reflection of increased expectancy, goal valuation, and motivation.

Secondly, the category charts reveal that all categories show distinct peaks at specific distances, reflecting an orderly progression as the intensity of the category increases. This observation holds true for all categories in the high motivation instruction, and a similar observation also applies to non-pushing categories in low motivation instruction, as there is a gradual increase in those categories. This pattern is consistent with the previous finding, which showed an increased frequency of pushing categories as pedestrians approach the bottleneck entrance. Additionally, it indicates that behavior change of pedestrians shows a clear and organized sequence, increasing into more intensive categories as they approach the bottleneck.

Another pattern observed in the category charts is the presence of a starting threshold for mild pushing and strong pushing behaviors in terms of proximity to the bottleneck. It appears that for the high motivation instruction, mild pushing behavior starts to increase at approximately three meters before the bottleneck entrance. This finding is not surprising, considering that most crowd formations in the analyzed videos occurred within a three-meter reach of the bottleneck. However, there is also a threshold for strong pushing behavior, which is roughly two meters from the bottleneck. This suggests that intense pushing behavior begins either after the crowd has formed or when pedestrians are in close proximity to the bottleneck. While the increase in mild pushing at three meters could be attributed to a higher number of data points in that range, the two-meter threshold for the strong pushing category is a noteworthy finding.

Furthermore, the findings suggest that the notion that pushing leads to faster access is circumstantial and may not have a significant effect in terms of individual success. The width of the corridor emerges as a crucial factor in determining this particular flow dynamics. Narrower corridors tend to exhibit smoother flow, and among pedestrians rated with pushing categories, none of them demonstrated a visible advantage in narrow corridors. In fact, all pedestrians in narrow corridors, whether rated as pushers or non-pushers, reached the bottleneck in similar time and distance properties. However, in wider corridors where congestion and crowd formations occurred, there were instances where pushing pedestrians reached the bottleneck faster than their non-pushing counterparts. It is worth noting that this pattern was not explicitly consistent across all wide corridor runs; it was clearly evident in only one set of experiment data, while being hinted at in others.

Finally, our temporal analyses have revealed significant insights. The primary objective of these analyses was to investigate whether the observed pushing category results could be attributed to non-experimental factors, such as participants' prepositioning on the platform before the experiments began, particularly in front of the bottleneck due to the provided motivation instructions. We refer to this potential phenomenon as the "sorting effect," which suggests that motivation instructions might have strongly influenced the positioning of pedestrians even before the experiments started, with motivated participants placing themselves in front of the bottleneck and quickly exiting, rather than being evenly distributed across the platform. However, the category increase and mean category value analyses suggested that this effect was not a highly dominant factor. While there was a general tendency for category increases during the initial quarter of the experiments, we also observed multiple peak points for increases, most of which occurred after the first few seconds and in the middle of the experiment periods (or proportionally in the last quarter of the experiments). Similarly, the mean category values were typically highest during the middle stages of the experiments. Furthermore, we have strong confidence that our spatial findings were not significantly influenced by these potential issues, as all spatial results were derived from the experiment periods, excluding the initial five to seven seconds. However, this does not necessarily imply the absence of the "sorting effect," as the initial seconds still contained a substantial increase. Simultaneously, the presence of other effects, such as category increases and mean pushing category values peaking in the middle of the experiment periods, suggests the existence of other dynamics requiring further investigation, while also indicating a semi-homogeneous distribution of motivated pedestrians before the experiments began.

4.1. Limitations

It is important to note that the interpretations and discussions were based on artificial settings, which may limit their generalizability to real-world scenarios. It is crucial to apply the established category system in real-life situations and different experimental environments, such as evacuation scenarios, to ensure the relevance and replicability of the findings. Additionally, although we examined different corridor widths and motivations in bottleneck scenarios, other potential factors, such as bottleneck width or having multiple bottlenecks, were not thoroughly considered due to the limitation of having only one type and size in our study. It is necessary to verify the applicability of the findings in various settings, including wider or narrower bottleneck widths, as wider widths may lead to smoother crowd flow, potentially resulting in less utilization of pushing behavior.

Furthermore, the interpretations we made based on our findings, such as the idea that an increase in motivation leads to pushing as pedestrians get closer to the bottleneck, are somewhat speculative since we did not collect any subjective data from pedestrians regarding their underlying psychological mechanisms. While our interpretations align with existing motivation literature, they focus on one specific explanation, and it is important to acknowledge that there are other plausible explanations within different theories that may also hold true. For instance, competition theories or the propagation concept (see [25]) could also contribute to these results, aside from motivation. The lack of subjective data is a major limitation of this study, and despite our confidence in interpreting the events with motivation literature, we cannot disregard the fact that future studies should investigate this issue by focusing on the subjective psychological mechanisms to see the full picture.

4.2. Practical Implications

The present study has the potential to contribute valuable knowledge to crowd safety by examining the forward motion behaviors displayed in various circumstances, such as different corridor widths and instructed motivations. These behaviors have been observed to follow specific patterns, with excessive behaviors tending to occur in certain areas. Understanding where potentially dangerous behaviors may emerge can inform crowd managers and help them implement practical measures for prevention or mitigation. Additionally, researchers in the fields of pedestrian dynamics and crowd sciences can utilize the key findings of this study or build upon them to further investigate the link between pushing behavior, motivation, and general crowd behavior in terms of spatial and temporal dynamics.

Acknowledgements The authors acknowledge Prof. Dr. Armin Seyfried and Dr. Maik Boltes for providing the PeTrack software. The authors thank Helena Lügering for her role as a co-rater. Additionally, the authors utilized ChatGPT to proofread the manuscript.

Ethics Statement The ethical statements and protocols regarding the experiments and recordings utilized in this study are available in the previous studies from which we sourced the materials (see [5, 13]).

Author Contributions Ezel Üsten, first author: Rating process, conceptualization, analysis, writing, original draft preparation / Jette Schumann: Conceptualization, visualization, analysis / Anna Sieben: General supervision, revision, editing

References

[1] Lügering, H., Üsten, E., Sieben, A.: Pushing and Non-pushing Forward Motion in Crowds: A Systematic Psychological Observation Method for Rating Individual Behavior in Pedestrian Dynamics. Collective Dynamics **7**, 1–16 (2022). doi:10.17815/CD.2022.138

[2] Mann, L.: Queue culture: The waiting line as a social system. American Journal of Sociology **75**(3), 340–354 (1969)

[3] Kneidl, A.: How do people queue? a study of different queuing models. Traffic and Granular Flow '15 (2016). doi:10.1007/978-3-319-33482-0_26

[4] Haghani, M., Sarvi, M., Shahhoseini, Z.: When 'push' does not come to 'shove': Revisiting 'faster is slower' in collective egress of human crowds. Transportation Research Part A: Policy and Practice **122**, 51–69 (2019). doi:10.1016/j.tra.2019.02.007

[5] Adrian, J., Seyfried, A., Sieben, A.: Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology. Journal of The Royal Society Interface **17**(165), 20190871 (2020). doi:10.1098/rsif.2019.0871

[6] Johnson, N.R.: Panic at "The Who Concert Stampede": An Empirical Assessment*. Social Problems **34**(4), 362–373 (1987). doi:10.2307/800813

[7] Helbing, D., Farkas, I.J., Molnar, P., Vicsek, T.: Simulation of Pedestrian Crowds in Normal and Evacuation Situations. Pedestrian and evacuation dynamics **21**(1), 21–58 (2002). doi:20.500.11850/123980

[8] Garcimartín, A., Zuriguel, I., Pastor, J.M., Martín-Gómez, C., Parisi, D.R.: Experimental Evidence of the "Faster Is Slower" Effect. Transportation Research Procedia **2**, 760–767 (2014). doi:10.1016/j.trpro.2014.09.085

[9] Feldmann, S., Adrian, J.: Forward propagation of a push through a row of people. Safety Science **164**, 106173 (2023). doi:10.1016/j.ssci.2023.106173

[10] Drury, J., Cocking, C., Reicher, S., Burton, A., Schofield, D., Hardwick, A., Graham, D., Langston, P.: Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model. Behavior Research Methods **41**(3), 957–970 (2009). doi:10.3758/BRM.41.3.957

[11] Reicher, S.D.: The battle of westminster: Developing the social identity model of crowd behaviour in order to explain the initiation and development of collective conflict. European Journal of Social Psychology **26**(1), 115–134 (1996)

[12] Institute for Advanced Simulation 7 Civil Safety Research, F.J.: Data archive of experimental data from studies about pedestrian dynamics. doi:10.34735/ped.da

[13] Institute for Advanced Simulation 7 Civil Safety Research, F.J.: Crowds in front of bottlenecks from the perspective of physics and social psychology [data set] (2018). doi:10.34735/ped.2018.1

[14] Boltes, M., Seyfried, A., Steffen, B., Schadschneider, A.: Automatic Extraction of Pedestrian Trajectories from Video Recordings. In: Klingsch, W.W.F., Rogsch, C., Schadschneider, A., Schreckenberg, M. (eds.) Pedestrian and Evacuation Dynamics 2008, pp. 43–54. Springer, Berlin, Heidelberg (2010). doi:10.1007/978-3-642-04504-2_3

[15] Krippendorff, K.: Computing krippendorff's alpha-reliability (2011). Retrieved from https://repository.upenn.edu/asc_papers/43.

[16] Hayes, A.F., Krippendorff, K.: Answering the Call for a Standard Reliability Measure for Coding Data. Communication Methods and Measures **1**(1), 77–89 (2007). doi:10.1080/19312450709336664

[17] De Swert, K.: Calculating inter-coder reliability in media content analysis using krippendorff's alpha p. 15 (2012)

[18] Sieben, A., Schumann, J., Seyfried, A.: Collective phenomena in crowds—Where pedestrian dynamics need social psychology. PLOS ONE **12**(6), e0177328 (2017). doi:10.1371/journal.pone.0177328

[19] Jhang, J.H., Lynch, J.G.: Pardon the interruption: Goal proximity, perceived spare time, and impatience. Journal of Consumer Research **41**(5), 1267–1283 (2015). doi:10.1086/679308

[20] Steel, P., König, C.J.: Integrating theories of motivation. Academy of Management Review **31**(4), 889–913 (2006). doi:10.5465/amr.2006.22527462

[21] Latham, G., Seijts, G.: The effects of proximal and distal goals on performance on a moderately complex task. Journal of Organizational Behavior **20**(4), 421–429 (1999)

[22] Atkinson, J.: An introduction to motivation. An introduction to motivation. Van Nostrand, Oxford, England (1964)

[23] Vroom, V.: Work and motivation. Work and motivation. Wiley, Oxford, England (1964)

[24] Porter, L.W., Lawler, E.E.: Managerial attitudes and performance. Irwin-Dorsey series in behavioral science in business. R.D. Irwin, Homewood, Ill. (1968). OCLC: 243614

[25] Lügering, H., Alia, A., Sieben, A.: Psychological pushing propagation in crowds – does the observation of pushing behavior promote further intentional pushing? Journal of Applied Social Psychology (submitted) (2023)

A. Appendices

A.1. Time-Distance Trajectories

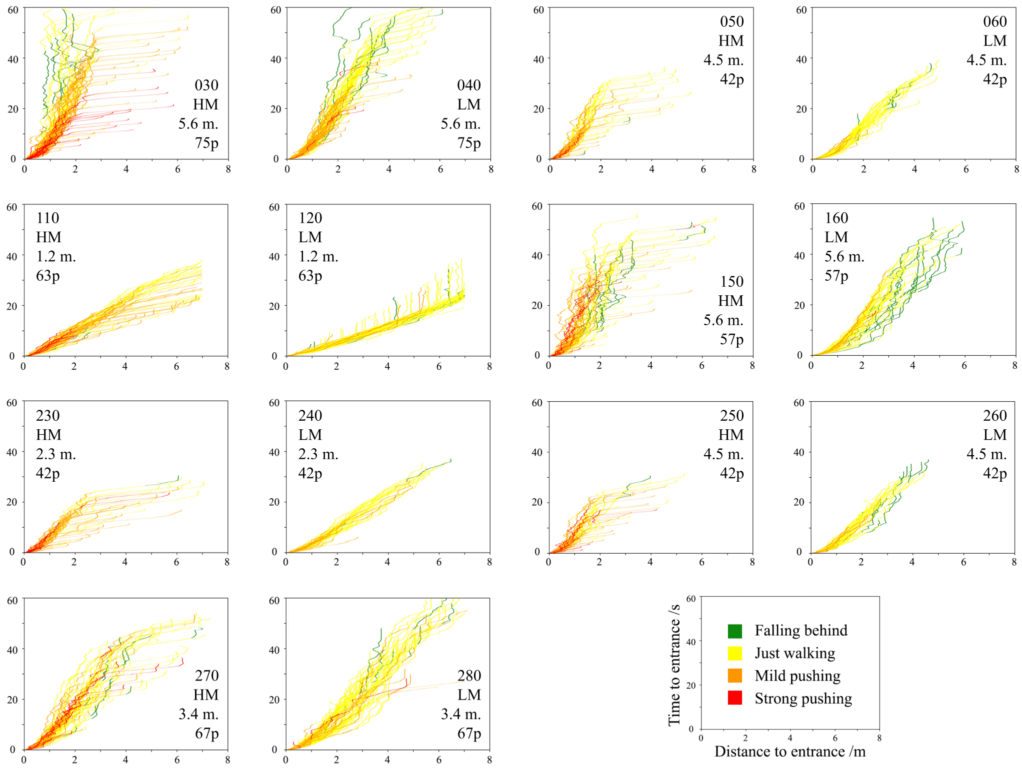


Figure 10 Complete time-distance analysis of the experimental runs is depicted in the figure. The X-axis of each plot represents "distance to entrance" in meters (0 to 8 m.), and the Y-axis represents "time to entrance" in seconds (0 to 60 seconds). Individual time and distance trajectories of pedestrians are color-coded according to category labels. Red indicates "Strong pushing," orange indicates "Mild pushing," yellow indicates "Just walking," and green indicates "Falling behind." Trajectories start from the initial position of each pedestrian and end when the pedestrian reaches the bottleneck. Side notes provide additional information, including the motivation level (HM = high motivation; LM = low motivation), corridor width (5.6 m.; 4.5 m.; 3.4 m.; 2.3 m.; 1.2 m.), and the number of pedestrians in each specific run. The order of the plots reflects the sequence of the conducted experiments.

A.2. Distance Bins (Relative Frequency)

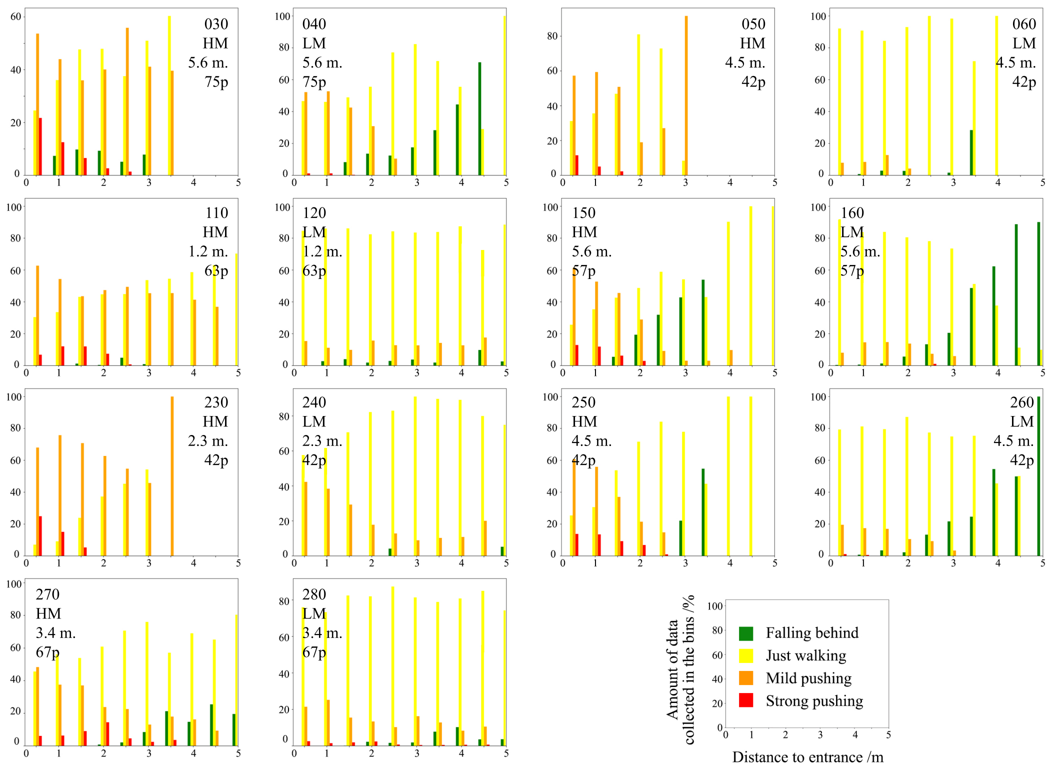


Figure 11 The figure depicts a complete (relative frequency) distance analysis of the experimental runs using bin columns. The X-axis of each plot represents "distance to entrance" in meters (0 to 5 m., with each bin placed at 0.5 m. intervals), and the Y-axis represents the percentage of category data collected within each semi-circle, presented as bins. The collected data in each bin are color-coded according to category labels. Red indicates "Strong pushing," orange indicates "Mild pushing," yellow indicates "Just walking," and green indicates "Falling behind." Side notes provide additional information, including the motivation level (HM = high motivation; LM = low motivation), corridor width (5.6 m.; 4.5 m.; 3.4 m.; 2.3 m.; 1.2 m.), and the number of pedestrians in each specific run. The order of the plots reflects the sequence of the conducted experiments.

A.3. Category Charts

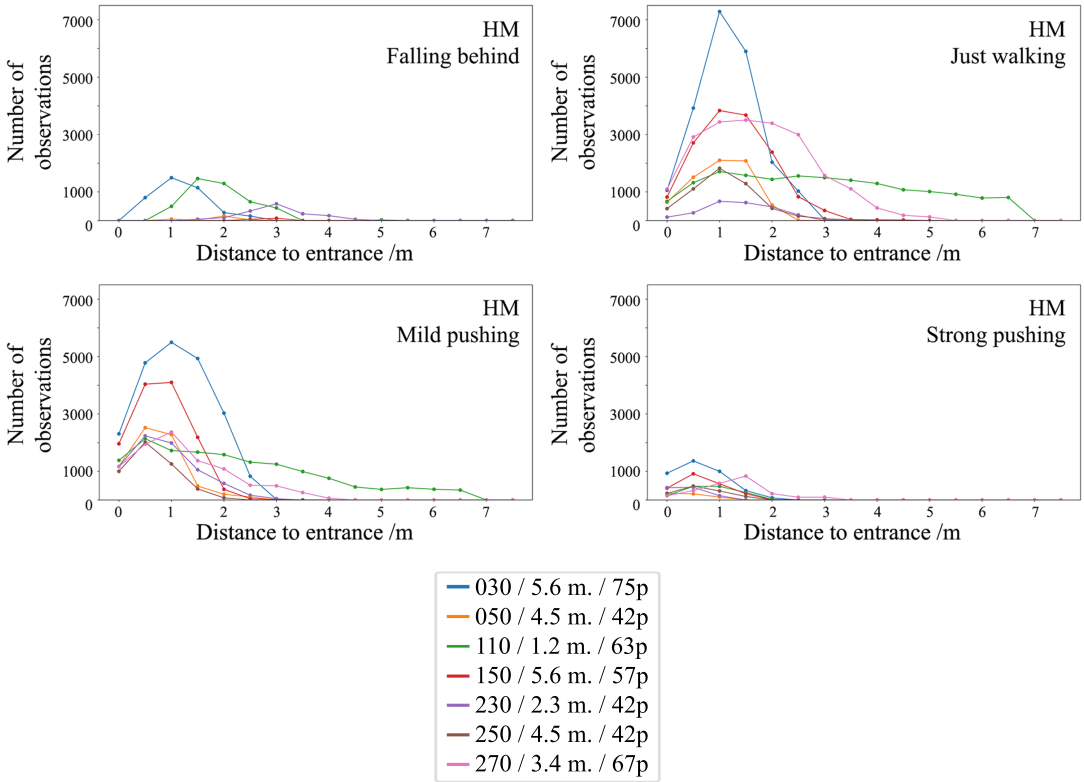


Figure 12 The figure illustrates a complete (absolute frequency) distance analysis of the pushing categories using coordinate points for high motivation runs. The X-axis of each plot represents “distance to entrance” in meters (0 to 7 m., with each point placed at 0.5 m. intervals), and the Y-axis represents the number of category data collected from all frames within each semi-circle. Colors represent the different experimental runs. Side notes provide additional information, including the motivation level (HM = high motivation) and the specific pushing categories being presented (falling behind, just walking, mild pushing, strong pushing).

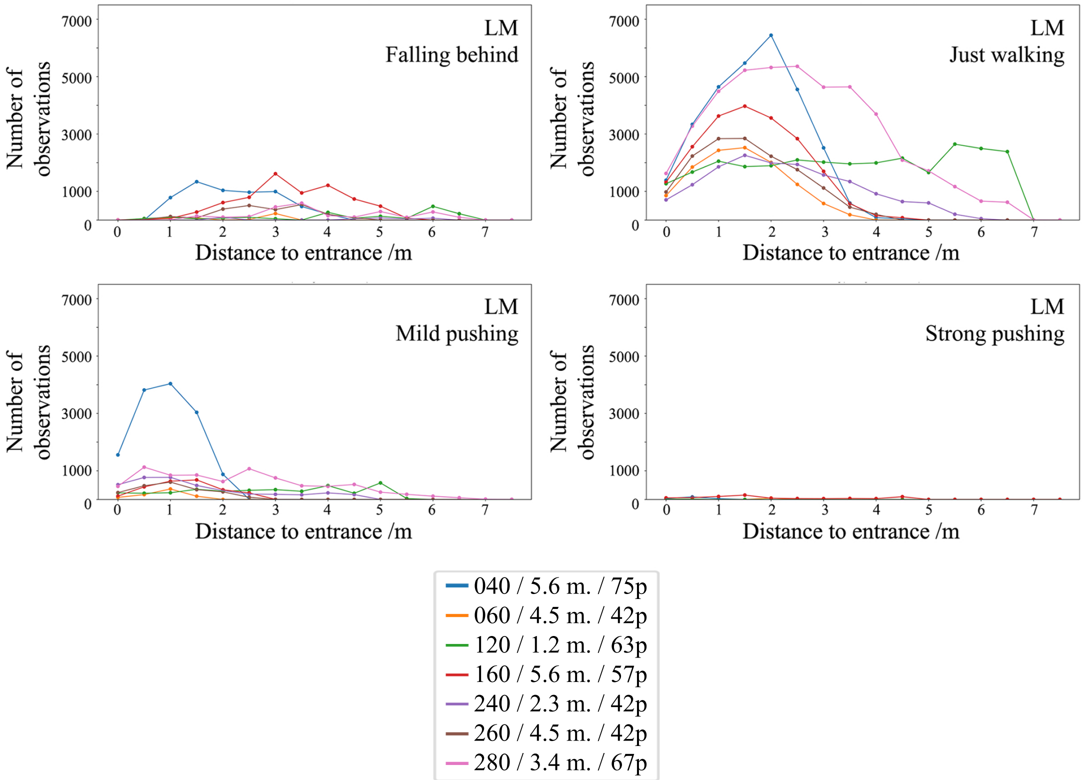


Figure 13 The figure illustrates a complete (absolute frequency) distance analysis of the pushing categories using coordinate points for low motivation runs. The X-axis of each plot represents "distance to entrance" in meters (0 to 7 m., with each point placed at 0.5 m. intervals), and the Y-axis represents the number of category data collected from all frames within each semi-circle. Colors represent the different experimental runs. Side notes provide additional information, including the motivation level (LM = low motivation) and the specific pushing categories being presented (falling behind, just walking, mild pushing, strong pushing).

A.4. Heat Maps

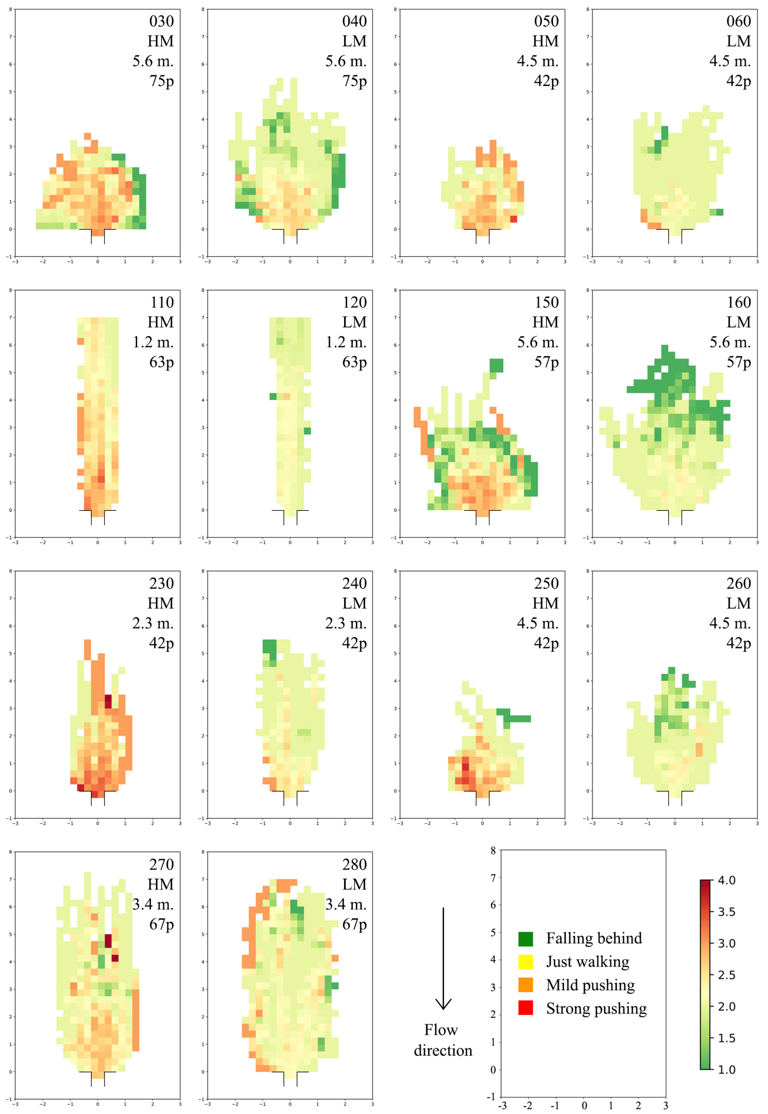


Figure 14 The figure illustrates the mean category data collected in 25cm to 25cm cells across all experimental runs. The X-axis represents the width of the bottleneck platform (-3 m. to 3 m., with a total width of 6 m.), while the Y-axis represents the length of the bottleneck platform (8 m.). The data flow, or pedestrian flow, is depicted from top to bottom. The collected data in each cell are color-coded based on the color scale derived from the category labels. Red represents “Strong pushing,” orange represents “Mild pushing,” yellow represents “Just walking,” and green represents “Falling behind.” Intermediate colors from the color scale are also used. The color scale is presented on the bottom-left of the figure. Side notes provide additional information, including the motivation level (HM = high motivation; LM = low motivation), corridor width (5.6 m., 4.5 m., 3.4 m., 2.3 m., 1.2 m.), and the number of pedestrians in each specific run. The order of the plots reflects the sequence of the conducted experiments.

A.5. Category Increase and Decrease Frequency

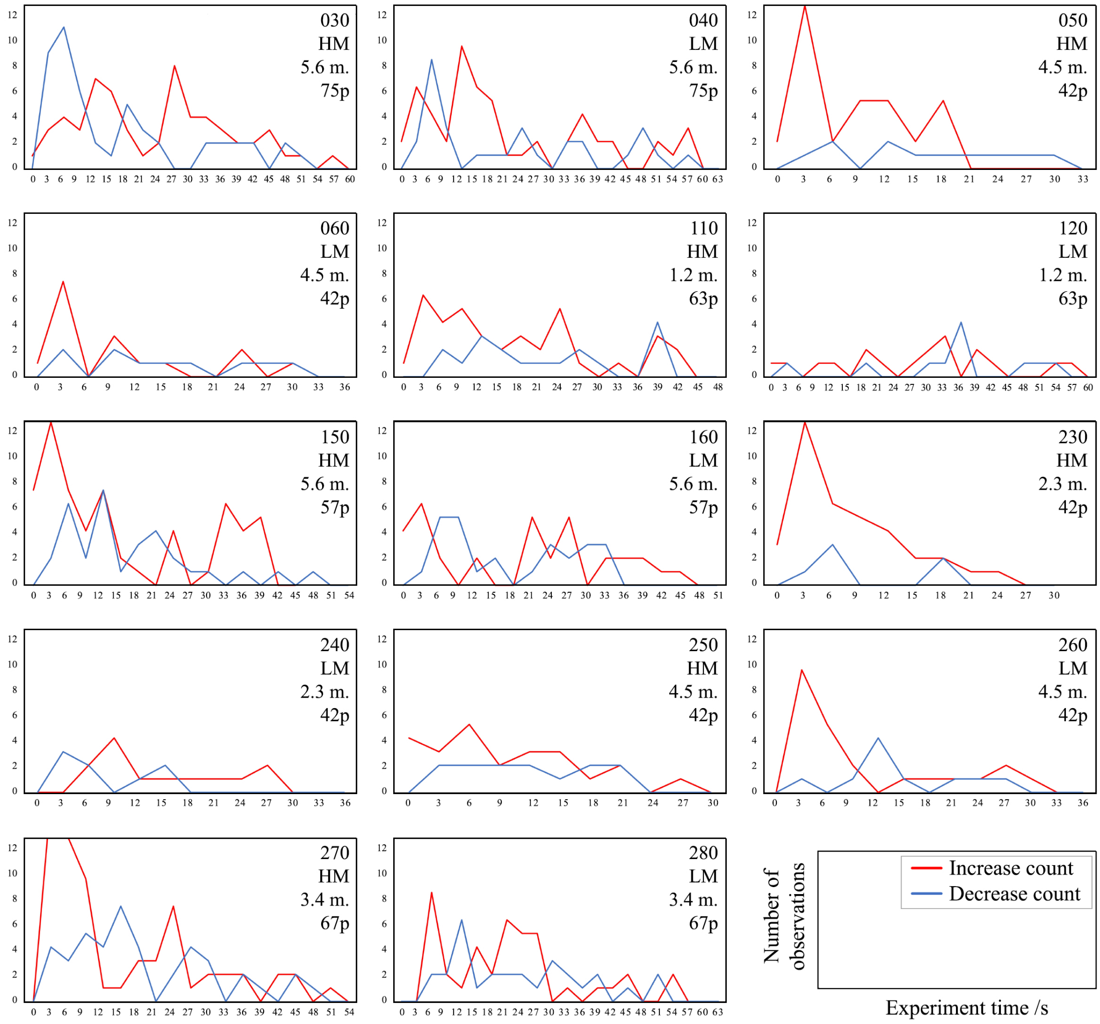


Figure 15 The figure illustrates time/frequency data for pushing category increases and decreases across all experimental runs. The category increases and decreases were collected in each second and aggregated in three-second intervals. The X-axis of the plots represents the overall experiment time period in seconds, while the Y-axis represents the frequency of category increases (in red) and decreases (in blue). Labels provide additional information, including the motivation level (HM = high motivation; LM = low motivation), corridor width (5.6 m., 4.5 m., 3.4 m., 2.3 m., 1.2 m.), and the number of pedestrians in each specific run. The order of the plots reflects the sequence of the conducted experiments.

A.6. Category Increase and Decrease Percentage

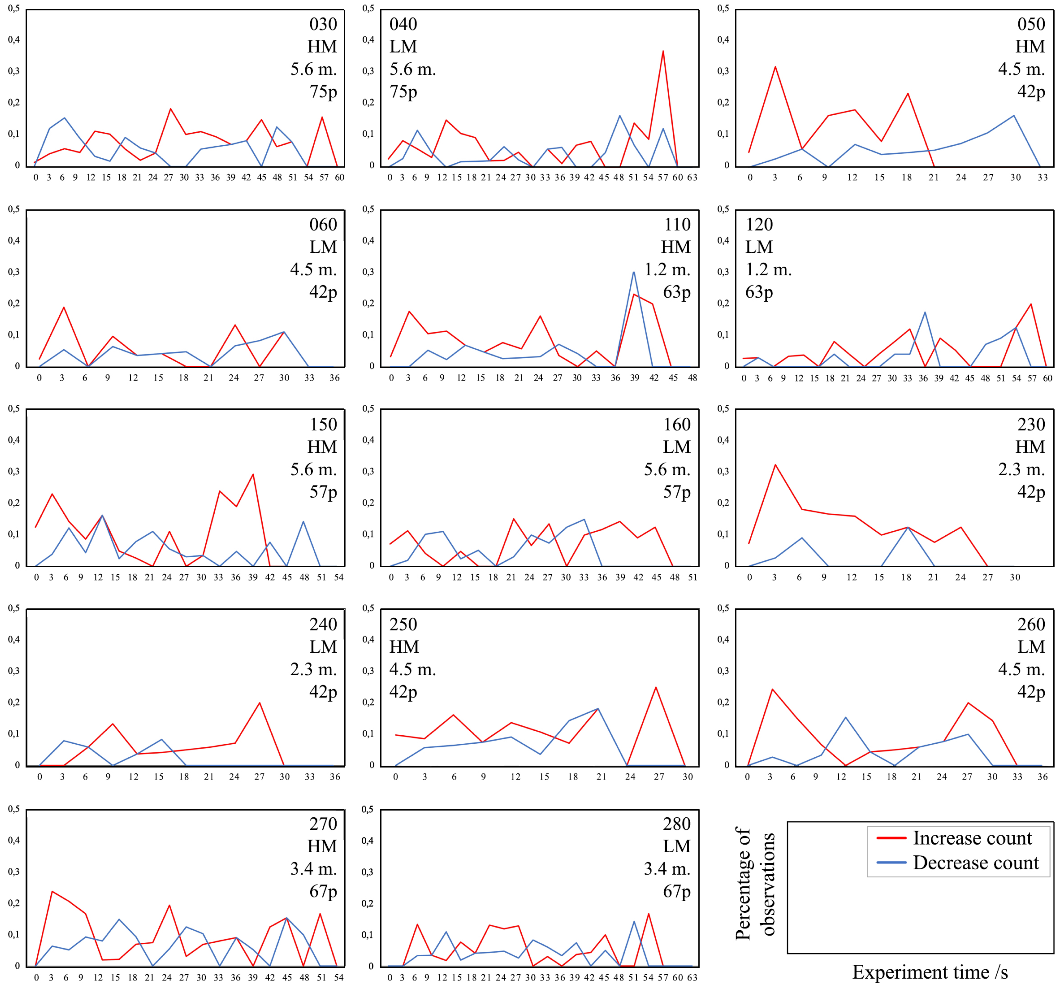


Figure 16 The figure illustrates the proportion of the time/frequency data for pushing category increases and decreases across all experimental runs. The category increases and decreases were collected in each second and aggregated in three-second intervals. The ratio is calculated from increase and decrease frequency divided by the total pedestrian count on the platform for that specific second. The X-axis of the plots represents the overall experiment time period in seconds, while the Y-axis represents the percentage of category increases (in red) and decreases (in blue) with a range of 0 to 50 percent. Titles provide additional information, including the motivation level (HM = high motivation; LM = low motivation), corridor width (5.6 m., 4.5 m., 3.4 m., 2.3 m., 1.2 m.), and the number of pedestrians in each specific run. The order of the plots reflects the sequence of the conducted experiments.