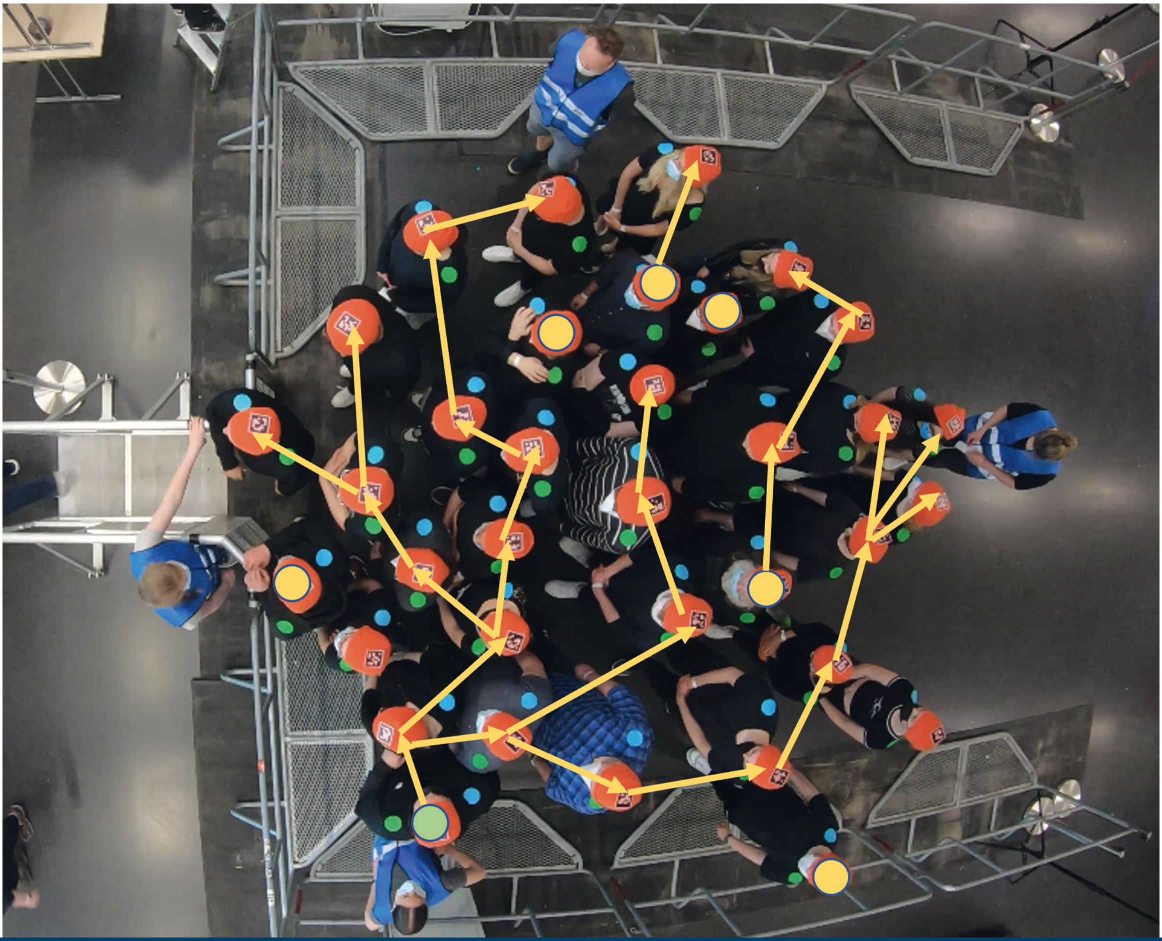

Propagation of Stimuli in Crowds: Empirical insights into mutual influence in human crowds

Helena Lügering

IAS Series
Band / Volume 65
ISBN 978-3-95806-775-2

Mitglied der Helmholtz-Gemeinschaft

JÜLICH
Forschungszentrum

Schriften des Forschungszentrums Jülich
IAS Series Band / Volume 65

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation (IAS)
Zivile Sicherheitsforschung (IAS-7)

Propagation of Stimuli in Crowds: Empirical insights into mutual influence in human crowds

Helena Lügering

Schriften des Forschungszentrums Jülich
IAS Series Band / Volume 65

ISSN 1868-8489 ISBN 978-3-95806-775-2

Bibliografische Information der Deutschen Nationalbibliothek.
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte Bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Herausgeber und Vertrieb:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Tel.: +49 2461 61-5368 Fax: +49 2461 61-6103 zb-publikation@fz-juelich.de www.fz-juelich.de/zb
Umschlaggestaltung:	Grafische Medien, Forschungszentrum Jülich GmbH
Druck:	Grafische Medien, Forschungszentrum Jülich GmbH

Schriften des Forschungszentrums Jülich
IAS Series, Band / Volume 65

D 468 (Diss. Wuppertal, Univ., 2024)

ISSN 1868-8489
ISBN 978-3-95806-775-2

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER) unter www.fz-juelich.de/zb/openaccess.

Abstract

This thesis deals with the mutual influence and propagation of stimuli in crowds. Considering the current state of research, the study of mutual influence in the social sciences, especially social psychology, is usually limited to small social groups and dyads. In crowd psychology, or more generally in research on crowd dynamics, assumptions about the mutual influence and the propagation of stimuli in crowds have more or less stagnated since the ideas of Gustave Le Bon at the end of the 19th century. However, his ideas are criticized in current literature and considered outdated. Nevertheless, there are several psychological mechanisms that make it conceivable for people in crowds to influence each other, such as heuristics, social norms, competition and communication. A common example from everyday life would be a pedestrian who runs a red light and is followed by others. The processes of mutual influence can be unconscious and/or unintentional as well as conscious and/or intentional and include various stimuli such as emotions, behavior, and information. The propagation of some of these stimuli, e.g. pushing behavior, would be particularly relevant for crowd dynamics, as pushing can lead to dangerous situations. Linked to the general topic of propagation is the image of mass panic in the context of crowd accidents, according to which panic propagates easily in crowds. In recent decades, researchers have repeatedly criticized this image for its various problematic implications, in particular the underlying assumption that panic behavior and panic contagion are decisive factors for an accident. This PhD project addressed the question of mutual influence and propagation on different levels by empirically investigating how pushing behavior and information propagate in a crowd and by examining lay people's everyday understanding of crowd accidents. The analytical tool developed for this purpose and the results of the studies were published in four papers.

In order to properly investigate pushing behavior with experimental data, a method for quantifying this behavior on video footage of crowds first had to be established. Therefore, Publication I presents the development of an observation and rating method for pushing and non-pushing forward motion based on the guidelines of quantitative content analysis. For this purpose, two researchers thoroughly examined video recordings of previous bottleneck experiments and identified six different parameters whose manifestations could be classified into four different categories, two pushing and two non-pushing. To demonstrate the suitability and reliability of this method, it was additionally applied to a video recording of another experimental run, resulting in excellent overlap and inter-rater-reliability of the ratings of two

trained raters.

In Publication II, this rating system was then used to assess the behavior of 776 participants in 14 runs of another earlier bottleneck experiment. The aim of this study was to investigate whether people were more likely to push if they already had a lot of contact with other people who push. In order to investigate this psychological form of pushing propagation, it was also necessary to identify who was neighboring whom at any time during an experimental run. To this end, a Python script was developed, using the information from the head trajectories and the Voronoi technique. When analyzing the combined data set (i.e., pushing behavior and neighborhood relations), it was found that the probability of starting to push increased to up to 30% if there were a large number of pushers in the close vicinity shortly beforehand.

The aim of Publication III was to investigate person-to-person communication in a crowd. A newly designed bottleneck experiment, in which the flow through the bottleneck was interrupted shortly after the start of each run and then information was passed to a person in this waiting crowd, who had to pass it on, was used to investigate the systematics of information propagation. It turned out that the efficiency of propagation (i.e., the number of people informed and the time taken) was influenced by the participants' knowledge and experience of the task as well as the density of the group. In addition, verbal transmission appeared to be more influential than visual transmission, and the direction of information propagation was generally maintained or rarely changed. These results were confirmed on the one hand in an experiment with smaller groups but several runs, and on the other hand in an experiment with two larger groups, each of which performed one run.

Finally, to investigate not only people's behavior in an experimental crowd setting but also to shed more light on lay people's underlying assumptions about crowds, Publication IV presents a mixed-method study to investigate lay people's associations with the term "mass panic" and two alternative terms. Using an online questionnaire and semi-structured interviews, participants were asked about their ideas of crowd accidents, including perceived dangers, appropriate behavior and ways to defuse a critical situation. In the questionnaire, all items were formulated using one of the terms, while people were confronted with all three terms in the interviews. Overall, the results showed that notions of panic, irrationality and selfishness shape lay people's understanding of crowd accidents, regardless of the terms used.

Zusammenfassung

Diese Arbeit beschäftigt sich mit der gegenseitigen Beeinflussung und Ausbreitung von Reizen in Menschenmengen. Betrachtet man den aktuellen Stand der Forschung, so sind Studien zur gegenseitigen Beeinflussung in den Sozialwissenschaften, insbesondere in der Sozialpsychologie, in der Regel auf kleine soziale Gruppen und Dyaden beschränkt. Im Gebiet der Massenpsychologie oder allgemein der Forschung zur Dynamik in Menschenmengen sind die Annahmen über die gegenseitige Beeinflussung und die Ausbreitung von Reizen in Menschenmengen seit den Ideen von Gustave Le Bon Ende des 19. Jahrhunderts mehr oder weniger stehen geblieben. Seine Ideen werden jedoch in der aktuellen Literatur kritisiert und als überholt angesehen. Dennoch gibt es eine Reihe psychologischer Mechanismen, die eine gegenseitige Beeinflussung von Menschen in Menschenmengen denkbar machen, wie Heuristiken, soziale Normen, Wettbewerb und Kommunikation. Ein gängiges Beispiel aus dem Alltag wäre ein Fußgänger, der über eine rote Ampel läuft und andere folgen. Die Prozesse der gegenseitigen Beeinflussung können sowohl unbewusst und/oder ungewollt als auch bewusst und/oder gewollt sein und verschiedene Stimuli wie Emotionen, Verhalten und Informationen betreffen. Die Ausbreitung einiger dieser Stimuli, z. B. Drängeln, wäre für die Dynamik in Menschenmengen besonders relevant, da Drängeln zu gefährlichen Situationen führen kann. Mit dem allgemeinen Thema der Ausbreitung verbunden ist das Bild der Massenpanik im Zusammenhang mit Unfällen in Menschenmengen, welchem zufolge sich Panik in Menschenmengen leicht ausbreitet. In den letzten Jahrzehnten haben Forscher:innen dieses Bild immer wieder wegen der verschiedenen problematischen Implikationen kritisiert, insbesondere wegen der zugrundeliegende Annahme, dass Panikverhalten und Panikansteckung entscheidende Faktoren für einen Unfall sind. In dem PhD-Projekt wurde der Frage nach der gegenseitigen Beeinflussung und Ausbreitung auf verschiedenen Ebenen nachgegangen, indem empirisch untersucht wurde, wie sich Drängeln sowie Informationen in einer Menschenmenge ausbreiten, und welches Alltagsverständnis Laien von Unfällen in Menschenmengen haben. Das zu diesem Zweck entwickelte Analyseinstrument und die Ergebnisse der Studien wurden in vier Publikationen veröffentlicht.

Um Drängelverhalten angemessen mit experimentellen Daten untersuchen zu können, musste zunächst eine Methode zur Quantifizierung dieses Verhaltens auf Videomaterial von Menschenmengen entwickelt werden. Daher wird in Publikation I die Entwicklung einer Beobachtungs- und Ratingmethode für drängelnde und nicht-drängelnde Vorwärtsbewegungen auf der Grundlage der Richtlinien der quantita-

tiven Inhaltsanalyse vorgestellt. Zu diesem Zweck untersuchten zwei Forscher:innen sorgfältig Videoaufzeichnungen früherer Engstellenexperimente und identifizierten sechs verschiedene Parameter, deren Ausprägungen in vier verschiedene Kategorien, zwei drängelnde und zwei nicht-drängelnde, eingeteilt werden konnten. Um die Eignung und Reliabilität dieser Methode zu demonstrieren, wurde sie außerdem auf eine Videoaufzeichnung eines anderen Versuchsdurchlaufs angewandt, wobei die Überlappung und Inter-Rater-Reliabilität der Ratings zweier geschulter Rater hervorragend war.

In Publikation II wurde dieses Ratingsystem dann verwendet, um das Verhalten von 776 Teilnehmenden in 14 Durchläufen eines anderen früheren Engstellenexperiments zu bewerten. Ziel dieser Studie war es, zu untersuchen, ob Personen eher drängeln, wenn sie bereits viel Kontakt zu anderen drängelnden Personen hatten. Um diese psychologische Form der Drängelausbreitung zu untersuchen, musste auch identifiziert werden, wer zu einem bestimmten Zeitpunkt während eines Versuchsdurchlaufs mit wem benachbart war. Zu diesem Zweck wurde ein Python-Skript entwickelt, das die Informationen aus den Kopftrajektorien sowie die Voronoi-Technik verwendet. Bei der Analyse des kombinierten Datensatzes (d. h. Drängelverhalten und Nachbarschaftsbeziehungen) zeigte sich, dass die Wahrscheinlichkeit, mit dem Drängeln zu beginnen, auf bis zu 30% anstieg, wenn sich kurz zuvor eine große Anzahl von Drängelern in der näheren Umgebung befand.

Ziel der Publikation III war es, die Kommunikation von Person zu Person in einer Menschenmenge zu untersuchen. Mit einem neu konzipierten Engstellenexperiment, bei dem der Fluss durch die Engstelle kurz nach Beginn eines jeden Durchlaufs unterbrochen und dann eine Information an eine Person in dieser wartenden Menge gegeben wurde, die diese dann weitergeben musste, wurde die Systematik der Informationsweitergabe untersucht. Es zeigte sich, dass die Effizienz der Informationsweitergabe (d. h. die Anzahl der informierten Personen und die benötigte Zeit) vom Wissen der Teilnehmenden über und der Erfahrung mit der Aufgabe sowie von der Dichte der Gruppe beeinflusst wurde. Darüber hinaus schien die verbale Weitergabe einen größeren Einfluss zu haben als die visuelle, und die Richtung der Informationsweitergabe wurde im Allgemeinen beibehalten bzw. nur selten geändert. Diese Ergebnisse wurden zum einen in einem Experiment mit kleineren Gruppen, aber mehreren Durchläufen, und zum anderen in einem Experiment mit zwei größeren Gruppen, die jeweils einen Durchlauf absolvierten, bestätigt.

Um schließlich nicht nur das Verhalten der Menschen im Experiment zu untersuchen, sondern auch mehr Licht in die zugrundeliegenden Annahmen von Laien über Menschenmengen zu bringen, wird in Publikation IV eine Mixed Method Studie vorgestellt, die die Assoziationen von Laien mit dem Begriff "Massenpanik" und zwei alternativen Begriffen untersucht. Mithilfe eines Online-Fragebogens und halbstrukturierter Interviews wurden die Teilnehmenden zu ihren Vorstellungen von Unfällen in Menschenmengen befragt, einschließlich wahrgenommener Gefahren, angemessenem Verhalten und Möglichkeiten zur Entschärfung einer kritischen Situation. Im Fragebogen wurden alle Items mit jeweils einem der Begriffe formuliert, während die Personen in den Interviews mit allen drei Begriffen konfrontiert wurden. Insgesamt zeigten die Ergebnisse, dass Konzepte wie Panik, Irrationalität und Egoismus das Verständnis der Laien von Unfällen in Menschenmengen prägen, und zwar unabhängig vom verwendeten Begriff.

Table of Contents

Acknowledgements

My PhD project would not have been possible without the incredible support of my supervisors, colleagues, friends, family and my boyfriend. When I first started in summer of 2020, we were in the first months of the COVID-19 pandemic. Back then, no one could really predict what would happen, and it seems a little ironic that I started in a research project on crowds at large-scale events in such a situation. When the first lockdown phase was announced a few weeks after I started, I never thought that I would built up such good contact with my colleagues and that, despite everything, I can look back on so many wonderful memories today.

However, that is how it is: Whether in the virtual meetings and coffee breaks or in person, I had great conversations - both thematic and private - and was able to let off steam when something went wrong. Even though I gradually spent more and more time in Berlin, which was the second reason why I spent less and less time at the institute, I always felt welcome. Now, at the end of my PhD, I would like to thank everyone who has accompanied me on this journey over the last 3.5 years, both inside and outside of work.

First of all, I would like to thank my day-to-day supervisor Prof. Anna Sieben. The last years have brought changes for her, too, so she often had to manage numerous tasks at the same time. Nevertheless, she always had time for me when I needed her and I could always rely on her innovative solutions. I thank her for her support and advice for my cumulative thesis, but also for putting things into perspective when I felt a little overwhelmed. I would also like to thank my supervisor, Prof. Armin Seyfried, who, despite his different scientific background, has a huge interest in and understanding of psychological research and has always been open to discussion. Further, I would like to thank Prof. Heather Kaths and Prof. Anna Baumert for their willingness to form the examination committee.

Next, I would like to thank the entire "Pedestrian Dynamics – Social Psychology" division. Although we were physically separated most of the time, we had a great team spirit. My special thanks go to Ezel Üsten, with whom I worked closely on some projects and who was definitely the calm anchor, Mira Beermann, who supported me in all matters and always had an open ear, and Kriszta Konya, with whom I had fascinating conversations about concepts and different research traditions. I would also like thank all the others who have been part of our small division over the years and with whom I have collaborated in some cases: Dilek Tepeli, Zeynep

Kapısız, Lena Blind and Marc Pelzer. Thank you for your support and the wonderful moments and conversations at division meetings, excursions and conferences.

A huge thank you also goes to the rest of the CrowdDNA team. Sina Feldmann started the PhD with me and we have reached (almost) all the milestones together. I have always enjoyed working with her and the effort we put into the CrowdDNA presentation has definitely paid off. Juliane Adrian patiently supported me with all my physical and programming challenges. And of course, all the other CrowdDNA partners: We had great discussions, lots of fun at the all-hands meetings and our punching bag experiments will probably go down in history.

Not to mention the many other colleagues: I would like to thank Mira Küpper for co-leading the project seminar and her help with all questions regarding the PhD process, Ann Katrin Boomers for the geocaching afternoons, the time as my roommate and the great conversations, and Alica Kandler for providing the experimental videos and her help in correcting the Petrack trajectories. Ahmed Alia who restructured my Excel spreadsheets more than once, Tobias Schrödter who supported me when Petrack was not working, and Maik Boltes who welcomed me in his division for a short time.

I am also grateful to Claudio Feliciani, Lena, Xiaolu and all the other wonderful people I met during my research stay in Japan. They helped me find my way around, introduced me to the culture and I have probably never discussed so much about panic terms in different languages before. My time in Japan was definitely a highlight of my PhD. I would also like to thank the Japan Society for the Promotion of Science (JSPS) for making this stay possible.

Finally, I would like to thank the people who have accompanied me through my PhD outside of work. I would particularly like to mention Judith and Veronika - my PhD fellow sufferers - who were probably best able to understand my ups and downs and who were always there to advise me. Also, my boyfriend David, who was always by my side during the 3.5 years, went through all the emotional outbursts (positive and negative), always believed in me and was an incredibly support, especially at the end of my PhD. And of course, I thank all my other friends and family who were happy with me when things went well, but who also listened to me when I complained about the same problem for the thousandth time.

Unfortunately, there is not enough space to thank everyone by name. Every single person from the Forschungszentrum Jülich and the University of Wuppertal has made my PhD a wonderful time and without the support of each and every one of them, such large projects as our experiments would not have been possible. So, many thanks to all of you!

List of Publications

Publication I
Üsten, E., Lügering, H., & Sieben, A. (2022). 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. https://doi.org/10.17815/CD.2022.138

Publication II
Lügering, H., Alia, A., & Sieben, A. (2023). Psychological pushing propagation in crowds—Does the observation of pushing behavior promote further intentional pushing? Frontiers in Social Psychology, 1, 1263953. https://doi.org/10.3389/frsps.2023.1263953

Publication III
Lügering, H., & Sieben, A. (2024). A Rumor has Spread like Wildfire? - Empirical Investigation of Information Propagation in Waiting Crowds. Collective Dynamics, 9, 1-28. https://doi.org/10.17815/CD.2024.146

Publication IV
Lügering, H., Tepeli, D., & Sieben, A. (2023). It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms. Safety Science, 163, 106123. https://doi.org/10.1016/j.ssci.2023.106123

CHAPTER 1

Introduction

1.1 Everything in a crowd is highly contagious? Introduction to a common misconception and its possibly true core

Large social gatherings are an essential part of human society. In addition, due to the growth of cities, more and more people are experiencing large crowds on a daily basis, for example in city centers or train stations. However, being in a crowd is not always a source of joy, spirituality, togetherness or at least an unavoidable part of the everyday life; being in a crowd can also be dangerous, as reflected by the increasing number of crowd accidents [1]. To understand the dynamics of crowds and thus provide guidance for crowd management to increase safety, major research efforts have been undertaken in recent years. Most research published in the interdisciplinary field of pedestrian dynamics comes from the fields of physics, engineering, mathematics or computer science and includes both empirical studies and crowd models. Nevertheless, some collective phenomena cannot be understood without the perspective of social sciences [2], which is why findings from these research areas have become increasingly important.

One crucial aspect that has often been neglected in crowd research is that a crowd is a social situation in which people (intentionally or unintentionally) interact with each other. In contrast, in disciplines traditionally dealing with people influencing each other like social sciences and especially in social psychology, research is usually limited to dyads or small social groups (e.g., families, peer groups, workplace). However, the need for empirical research on mutual influence is becoming more and more evident, as there are already various models of crowd dynamics that integrate this issue [3–8]. Due to the lack of psychological insights, the assumptions of these models remain rather superficial and are largely based on the authors' considerations.

In fact, the idea of people in crowds influencing each other is not new but resonates back to the beginnings of the research field of crowd psychology. At the end of the 19th century, Gustave Le Bon published his book "The crowd: A study of the popular mind", in which he wrote (p. 122): "Ideas, sentiments, emotions, and beliefs possess in crowds a contagious power as intense as that of microbes." [9]. He further argued that this contagion leads to a so-called collective mind in which personal interests and individuality are sacrificed – an assumption that was generally embedded in a very negative and destructive image of crowds. According to Le Bon, crowds are savage and barbaric, and the people who are part of them behave in an uncivilized and uncultured way, like creatures acting only on instinct.

In the following years, his ideas of everything being highly contagious is crowds has been taken up by some researchers [10, 11]. Nowadays, however, Le Bon's views are considered to have been strongly influenced by the political and social circumstances of his time and are therefore regarded as outdated [12, 13]. Nevertheless, the narrative of contagion is still common when it comes to describing and explaining negative and violent crowd behavior, such as riots (for a critical overview see [14, 15]). Additionally, the public image of "highly contagious crowds" is particularly apparent when an accident in a crowd has occurred. Newspaper and news programs, for example, often report the following: "Witnesses said panic spread over rumours of suicide bombers." [16] or "According to police officials present at the place of occurrence, a rumour about electric supply line falling on the crowd near the western exit gate of Gandhi Maidan spread like wildfire, triggering panic among people who ran for cover." [17]. Not only do these passages imply that information (especially rumors) is highly contagious and spreads like wildfire in crowds, but they also encourage the idea of panic that often arises in connection with crowd accidents.

According to this idea, emotions, especially panic, spread through a crowd and lead to mass panic through a self-reinforcing process. People then behave irrationally, selfishly, and competitively which ultimately causes injury and death. Thus, this image of mass panic as an explanation for crowd accidents is an assumption about crowds that can be mainly ascribed to Le Bon [18, 19]. Although it was already questioned in the mid-20th century and has since been repeatedly criticized and even completely rejected [20–28], it still influences reporting on crowd accidents, as terms such as "panic", "mass panic", "stampede" or their translations into other languages can usually be found right in the headlines.

To summarize, from a psychological point of view, a crowd is a social system in which it is reasonable that mutual influence plays a role. However, previous scientific approaches have been rather radical, assuming that everything in crowds is highly contagious, and have led to an understanding of crowds – especially regarding critical situations – that is often criticized by current research. Therefore, almost 130 years after Le Bon, this PhD project takes a further step towards answering the question of how contagious crowds actually are. However, due to their negative connotation and association with disease, the terms "contagion" and "contagious" should be avoided; in this thesis, the term "propagation" is predominantly used. While the term "propagation" is not entirely neutral either, as it is used in biology for reproduction, for example, it focuses more on the concrete underlying mechanisms. With "contagion", by contrast, most people have a clear idea of pathogens spreading from one person to another. Although this image is certainly oversimplified from an epidemiological perspective, as some pathogens are more contagious than others and the mechanism is not always the same, "contagion" is already associated with too many simplistic ideas to be suitable for the study of complex social processes in crowds (for a detailed critique of the ”contagion” metaphor, see [15]). Yet, other alternative terms also evoke different associations. Transmission or transfer, for example, focus mainly on the process between two individuals rather than many, communication has a strong verbal dimension and dissemination also refers more to someone passing on information or ideas (verbally) to others. However, this project was not just about investigating how verbal messages can propagate in a crowd. Therefore, the title “propagation of stimuli in crowds” was chosen, with the term “stimulus” serving as a placeholder for different research topics (e.g., information, behavior).

1.2 Psychological mechanisms of propagation: State of research

As mentioned in the previous section, there is various research on mutual influence in social psychology, but it mainly focuses on influence between two people or within a small social group. Based on these findings, however, various psychological approaches can be found that suggest that mutual influence could also occur in crowds and that stimuli may propagate. It is important to note that this influence can manifest itself in multiple ways: For one, the entire process can occur unconscious or at least unintentional, as is probably often the case with the transmission of emotions and perhaps behavior. The person being influenced may observe the emotion or behavior in themselves, but is unaware that it was caused by another person’s emotion or behavior. For another, it may be conscious and intentional imitation of a behavior or following of others. Alternatively, people take in information from others and behave accordingly when necessary. The following sections discuss some of the psychological approaches suggesting that different forms of influence can occur in crowds.

Emotional Contagion

A common example for mutual influence is the study of emotional contagion. Emotional contagion was originally defined as “the tendency to automatically mimic and synchronize expressions, vocalization, postures and movement with those of another person’s and, consequently, to converge emotionally” ([29], p. 96). In recent decades, this topic has motivated numerous studies investigating emotional contagion in different circumstances as well as the underlying mechanisms. In most experiments, participants sit in front of a computer and are confronted with pictures or videos of people showing different emotions. The participants’ emotions are recorded either during the trial, e.g. using facial electromyography (fEMG), or afterwards using a questionnaire. A different research approach was taken by Hatfield et al. [30] who, inspired by Le Bon’s assumptions, reviewed the application of emotional contagion theory to various, mostly historical, examples of collective emotions (e.g., dancing manias and mass hysteria) and cases where mental or physical diseases were reportedly contagious. However, as this was only a post-hoc analysis, the circumstances could not be assessed and possible alternative explanations could not be ruled out. The authors' uncritical use of emotional contagion as an explanation for these examples should therefore be treated with caution.

In general, the original definition of emotional contagion has been increasingly questioned in recent decades. On the one hand, the idea of an automatic, unintentional and irrepressible process that resembles the contagion of a disease (which is already evident in the naming) and is in line with the ideas of Le Bon has been criticized [31]. On the other hand, recent work has shown that, for example, mimicking facial expression is not necessarily a precursor to emotional contagion [32, 33]. Moreover, how strong an emotional response is and whether someone shows a concordant or discordant emotional response depends on the interaction between the specific emotion and the perceived similarity, liking and interpersonal relationship [34–37]. Nevertheless, the general effect of transferring emotions from one person to another is well-established.

Beyond dyads, Barsade [38] showed that emotional contagion also occurs in small groups (3-4 people). Yet, for emotional contagion to work in larger groups, such as at concerts, the emotions would have to be transmitted across several people and even with a time delay. Until now, however, experiments have usually been designed in such a way that all responding people can see the "source" of the emotion directly, which obviously does not correspond to reality in crowds. Initial evidence that emotional contagion also occurs beyond dyads was provided by the experiment of Dezecache et al. [39]. They demonstrated that feelings of joy and fear expressed by person A were transferred to person C, even though person C never saw person A, but only the face of person B, who actually observed person A. To the best of my knowledge, however, no further research has been conducted on this topic.

Imitation and heuristics

Apart from the issue of emotional contagion, there are other empirical findings showing that people in crowds indeed influence each other. For example, when the situation during an evacuation is highly uncertain and people do not have sufficient information, they are more likely to follow others [40]. Alternatively, an example of a non-emergency context: Up to 40% of passerby followed the gaze of a group of people standing on the street looking up to a certain point [41]. One possible explanation for both examples is that the people followed the behavior of others because they believed that the others had more information (i.e., they knew a better way out or they saw something special up there). Psychologically, this is also plausible, as it fits the well-known heuristic "Imitate-the-majority" stating that acting like others is a good idea in many cases [42, 43], and heuristics are often used when there is a lack of time, information, or cognitive capacity to thoroughly compare and weigh options. Another heuristic that is closely related to the first is "Imitate-the-successful" [42, 43]. An example from the crowd context in which this heuristic may have played a fatal role is provided by the work of Sieben & Seyfried [44]. In their analysis of the witness statements of the Love parade disaster in 2010, they found that people who saw others climbing the stairs or the poles wanted to do the same, as they saw this behavior as an opportunity to escape the dangerously dense crowd. Since the climbing was also clearly visible due to the elevated position of the stairs and poles, it led to many people moving there (some intentionally and others through the movement of the crowd). However, this had a dangerous secondary effect as it caused a further densification of these areas.

Social norms

A further mechanism that makes mutual influence in crowds conceivable are social norms. Our societies depend on various rules and laws. Some of them, however, are normative, but not written down and legally binding – at least not necessarily. These so-called social norms provide directives for behavior in social situations such as being quiet in a church or library. Although many social norms have already been established, new ones can emerge, for example as a result of a crisis (see Emergent Norm Theory [45]) or because there is no norm for a specific situation yet. In the context of crowds, research has shown that there are no uniform social norms that apply to all possible crowd situations, but rather that these are context-dependent. Queues, for example, are relatively well-defined social situations in which different norms apply. For instance, queue-jumping is not permitted or only permitted under certain circumstances (e.g., if the train will otherwise be missed), and pushing is generally not accepted [46, 47].

If, however, instead of a queue, a semicircle forms in front of an entrance, it is less clear which behavior is considered socially acceptable. Evidence for this is provided by two empirical studies investigating the behavior in different entrance scenarios [2, 48]. In cases where the structural conditions resulted in the crowd resembling more of a queue (i.e., a smaller corridor), less unfair behavior (particularly pushing and shoving) was observed. In addition, the prevailing social norms were more geared toward orderly, lined-up behavior in which pushing is prohibited. Then again, in the setups that led more to a semicircle situation, more unfair behavior was observed and the participants indicated that either no rules applied or the right of the stronger or faster person prevailed. Entrance scenarios thus create a normatively ambiguous situation, in which the question of whether pushing is allowed or not cannot be answered clearly. Therefore, observing pushing behavior could resolve the normative uncertainty and establish a group norm of pushing being allowed.

Social identity theory

Another prominent approach within the limited psychological research on crowd dynamics which goes in a similar direction is mainly pursued by the group around John Drury and Stephen Reicher. Using the Social identity theory [49], they argue that people in crowds – especially in emergency situations – behave similarly because they share a common social identity (for a review see [50]). Originally, the Social identity theory proposes that people identify with different social groups based on characteristics such as gender, nationality, religion or preferences. It is possible to identify with more than one group, and depending on the context, one identity is more salient than others (e.g., football fans are more likely to identify with the group of other fans of their team when they are in the stadium than during a parentsteacher conference at their children's school). In addition to other influences on a person's self-concept and self-esteem, identification with a group also has an impact on behavior. The reason for this is that people define appropriate behavior by referring to the norms of the groups they belong to and then adapt their behavior accordingly, depending on the salient identity.

Returning to behavior in crowds, Drury et al. [51] investigated cooperative vs. competitive evacuation behavior in an underground environment with a virtual reality approach. Cooperative behavior was operationalized as helping behavior while competitive behavior was attributed to pushing others aside. The results showed that people with high group identification helped more and pushed less than people with low group identification. A questionnaire study conducted after a natural disaster corroborated this finding and provided additionally evidence that people are influenced by the helping or supportive behavior of others [52]. More specifically, the data revealed that observing supportive behavior led people to engage in more supportive behavior themselves, and that this relationship was stronger among high-identifiers. In summary, the social identity approach suggests that the question of what behavior a person exhibits and how strongly a person is influenced by the behavior of others depends on how strong the social identity is.

Competition and Revenge

In addition to the two concepts of social norms and social identity, which tend to emphasize the psychological unity of a crowd, there are also psychological concepts that focus more on individual motives in relation to others, such as competition and revenge. These mechanisms may be particularly relevant for the propagation of harsher behaviors such as pushing. In their experiment, Drury et. al [51] even conceptualized competitive behavior as pushing behavior, and conversely, a sense of competition could indeed be one reason why people start pushing when they see other pushers. However, since pushing and competition can also be fun (think of cultural phenomena as mosh pits and circle pits [53, 54]), this does not always have to be as negative as it initially sounds. Yet, in situations where people have a common goal, competition might be accompanied by a perceived unfairness, as a pusher seems to reach the goal faster. Similarly, when people are directly pushed by others, they may simply push back, either to retaliate or to defend their own position. At this point, it is important to mention that in this thesis I am only referring to intentional pushing in terms of a voluntary action [24]. The fact that people might bump into others when they are pushed from behind is not addressed. Hence, the propagation effect I am referring to is also intentional or psychological. A physical analysis of the propagation of a push can be found in [55].

Communication

As already mentioned, the mutual influence and propagation of stimuli in crowds does not have to be limited to the (un)conscious or (un)intentional imitation of other people. It can also involve information that is passed on from person to person, which in turn can trigger an action. Although there are several general psychological communication models [56–58] illustrating the communication between two people, crowd communication has been rather underrepresented in research. As indicated in the introduction, public perception of information propagation in crowds tends to reflect the image of mass panic, assuming that a rumor spreads like wildfire. This is evident not only by the newspaper articles cited above but also indicated by previous research [23]. The study of Drury et al. [23] revealed additionally, that this idea is widespread even among professionals working in public safety and emergency response. However, the idea is contradicted by the fact that people at large-scale events usually do not know each other, since the large crowd consists of many smaller social groups and dyads, and people tend to be more hesitant to talk to strangers than to people they know. In addition, studies on actual crowd accidents have shown that information transfers rather poorly in very dense crowds [44, 59]. Reasons for this could be an impairment of the human senses, the inability to turn the upper body or the fact that people are more focused on themselves rather than on their surroundings in these critical situations. In cases where information is accompanied by behavior (e.g., moving in a certain direction), however, not only verbal but also visual propagation is possible. Of course, visual perception is also limited in dense crowds, but as the analysis of the Love parade witness statements has shown, there can be behaviors that are visible to a larger amount of people and which then have a high attraction and incentive to imitate [44].

Besides rumors or information that could cause unwanted behavior, crowd managers could also use person-to-person information sharing as a communication strategy. Although there is much research on how to address people in an emergency, such as through announcements or guides [60–63], this strategy has been little discussed. However, in order to use it effectively empirical research is needed, since it is not only important to know the systematics of information propagation (e.g., whether there is a preferred direction of communication), but also how reliably information propagates. Because apart from the fact that some people simply do not pass on the information, it can of course also change during transmission (think of the children's game "whisper down the lane").

Research on the image of mass panic

Although little research has been conducted into the mutual influence and propagation of stimuli in crowds to date, there is a related topic that has been intensively researched in recent decades: the concept of "mass panic". As mentioned in the introduction, this concept is based in the assumption that emotions and especially panic propagate in a crowd and lead to hazardous dynamics. Referring to analyses of real-life scenarios, however, current literature mainly challenges the idea that people panic in critical situations. In a nutshell, the researcher argue that competitive, irrational and egoistical behavior can occur in critical crowd situations, but that most of the time people behave calmly, appropriately and rationally and, above all, help each other [18, 19, 26, 28, 44, 59, 64, 65]. Furthermore, death is rarely caused by panicked behavior of others, but rather by extreme overcrowding, such as when people fall and are accidentally trampled or asphyxiate [24, 59, 66, 67]. In addition, the movements of people in extremely dense crowds are often determined by the movement of the crowd itself and not by their own intentions [59]. Therefore, the term “mass panic” is considered misleading, as it at least implicitly assumes that panic is the decisive factor for the accident and that nothing would have happened if the visitors had not panicked [68].

Although the points of criticism mainly refer to the “panic” part of the term, the “mass” part is also repeatedly questioned. While there are examples of people imitating and following others in crowds, as discussed above and also seen in [50], Haghani et al. [69] conclude in their comprehensive review that the idea of everyone blindly following the masses in an emergency (so-called herd behavior) is empirically untenable. Similarly, it is assumed that some people may panic in critical crowd situations but that this panic does not affect the entire crowd [27, 65]. Most of the cited works relate to the English-speaking world, but there are also critical voices against panic-related terms in other languages like German [70] or Japanese [71, 72]. As a result of this fierce criticism, there are repeated calls to replace these terms [22, 27]. Suggestions for alterative terms include “crowd crush” for the English-speaking world [73] or Massenunglück (“mass accident”) [74] or Massendesaster (“mass disaster”) [75] for the German-speaking area.

To summarize the state of research: Not much is yet known about the mutual influence and propagation of stimuli in crowds. Earlier ideas in the tradition of Le Bon, according to which everything in a crowd is highly contagious, were criticized and considered outdated. Nevertheless, they still shape the ideas about crowds today, e.g. in relation to crowd accidents, which contrasts with the strong scientific criticism of the image of mass panic and the associated ideas of panic contagion and herd behavior. From a psychological point of view, however, it is entirely plausible that people in crowds do influence each other. In a nutshell, this is the area of tension between the different perspectives in which the research work of my PhD project takes place.

1.3 Objective and Approach

To address such an innovative and interdisciplinary research topic, it is important to think outside the box and combine methods from different research traditions. My first objective was to experimentally investigate propagation in crowds using the examples of pushing behavior and information transfer. To do this, however, it was also necessary to overcome prevailing (implicit) ideas about crowds. Previous research on crowds often assumed that crowds are homogenous, e.g. in behavior, either between the people or at least within people themselves, meaning that people always behave the same and that behavior does not change over time. A good example is pushing behavior. Most studies have treated crowds as either medium- or high-motivated (and therefore everyone in the crowd pushes (slightly)) or as low-motivated (and no one pushes) [76–79], neglecting the fact that there can also be pushing people in low motivated crowds and vice versa. In addition, behavior can change over time, and people may start pushing, then stop, and then start again. One explanation for this could be the psychological propagation of pushing behavior. However, since the heterogeneity and temporal dynamics in pushing behavior have hardly been discussed so far, there is also no methodological approach to capture them and thus explore possible underlying mechanisms.

In order to further investigate pushing behavior, a method for observing and quantifying it was therefore first required. By thoroughly examining video recordings of previous bottleneck experiments and based on the guidelines of quantitative content analysis [80, 81], different categories of pushing and non-pushing behavior were defined. Accordingly, a rating system not only for pushing behavior, but also for forward motion in general was developed. To this end, various parameters (e.g., position of heads, usage of hands and interaction with others) were identified and evaluated in terms of their relationship to pushing. It was important that these parameters could be observed from above, as video recordings from above were used to minimize occlusions and to use the technical possibilities of the Software Petrack to annotate the videos [82]. Since at least two people were required to develop such a system using the quantitative content analysis, two researchers jointly collaborated on it. In addition to establishing the rating system, its suitability and quality had to be ensured, which is why the two researchers also applied the rating system to a video of a previous bottleneck experiment [83] and rated the behavior of each participant throughout the course of the experiment. These independent ratings were then checked for statistical overlap and inter-rater-reliability (Publication I).

Since Petrack not only offers the possibility to make annotations for different people at different time points on an overhead recording of a crowd, but also extract their head trajectories, it is possible to link the information about pushing behavior to the spatiotemporal information of the individual people. This feature was used to examine the question of psychological pushing propagation. Despite the multitude of conceivable explanations as to why this intentional form of pushing propagation might occur, it has not yet even been investigated whether this phenomenon can actually be found empirically. Therefore, the rating system of forward motion was applied on 14 experimental video recordings of another previous bottleneck experiment [84]. Additionally, the trajectory information was used to identify who is neighboring whom. By combing these two datasets, it was then statistically analyzed whether the probability of starting to push increased when the participants had more contact with people already pushing (Publication II).

In contrast to the investigation of pushing behavior, for which a secondary analysis of previous bottleneck experiments was conducted using a newly developed method, a novel bottleneck experiment was designed for the question of information propagation. In order to simplify the spatial dynamics, the flow through the bottleneck was interrupted shortly after the start of each run, resulting in waiting crowds where people are oriented in the same direction (such as in front of a concert stage, for example). Then, a piece of information was given to one participant in the waiting group, which should be passed on according to the children's game "whisper down the lane". Since information can also be accompanied by a behavior, not only purely verbal messages were used, but also commands in which the triggered behavior might additionally propagate visually. Based on the overhead video recordings as well as questionnaire completed by the participants, characteristics of information propagation such as transmission direction and speed were analyzed (Publication III).

Since crowd research inherently involves people and people's ideas and conceptions can influence their actions, it is conceivable that the crowd dynamics are also affected by people's assumptions about crowds. Therefore, the second objective of my PhD project was not only to look at the issue of propagation from an artificial, experimental perspective, but also to investigate lay people's understanding of it. Since the investigation of critical situations in an experiment is out of the question for ethical and moral reasons and field studies are only possible in retrospect, this perspective could be particularly relevant for research on crowd accidents. Terms like "mass panic" as description for accidents in crowds have been repeatedly criticized by researchers due to their implications of panic contagion and herd behavior and there are repeated calls to replace these terms. Therefore, a mixed-method study consisting of an online questionnaire and semi-structured interviews was conducted to investigate what lay people associate with the term and how these associations change when two alterative terms are used. To this end, participants were asked about their underlying ideas and assumptions on crowd accidents, e.g. why they occur or what participants consider to be dangerous in these situations. The answers were then compared between the different terms and, in addition, a concept of participants' understanding was developed (Publication IV).

Due to the COVID-19 pandemic restriction in 2020 and 2021, the last research question (Publication IV) was actually one that I started with, because the data collection did not involve studying physically present large crowds. However, as it nicely completes the results of the experimental studies on large crowds and provides some further practical implications, I have decided to present this study last.

CHAPTER 2

Results

The dissertation consists of four studies, which were published as four scientific papers. The following sections briefly present the research questions, methods and key findings of the individual publications. Detailed information on the individual publications can be found at the end of this thesis.

2.1 Publication I

Publication I ("Pushing and Non-pushing Forward Motion in Crowds: A Systematic Psychological Observation Method for Rating Individual Behavior in Pedestrian Dynamics") introduces and evaluates a rating system of forward motion. This system was developed and elaborated to assess the forward motion of people in crowds, both pushing and non-pushing, as well as its temporal dynamics. Using the guidelines of quantitative content analysis, six parameters of people's body posture and behavior that are detectable on video recordings from above were identified, namely the position of their arms and hands, their shoulders and head, the space they have around, their interaction with others, their speed and acceleration and the focus of their attention. On this basis, four different categories, each characterized by a specific set of manifestations of the various parameters, were defined: (1) falling behind, (2) just walking, (3) mild pushing, and (4) strong pushing – representing two pushing and two non-pushing categories. Their characteristics can be briefly summarized as follows:

Individuals assigned to category (1) usually have more distance to the front than to the back, meaning they move more slowly than the group, and their attention is often distracted (e.g., by their smartphone, the environment or because they are talking to others). They do not actively use their arms and hands but rather cross them or simply letting them hang down. The arm and hand position of individuals in category (2) can be similar, but they may also use them to protect their upper body or to stabilize themselves at barriers. Likewise, they may talk to others. However, their attention is more focused on the target (e.g., the exit) and they move with the flow at an appropriate speed. Their shoulders and head perform a penguin-like waddle motion and they maintain a proportionate distance from others (meaning that they may have body contact in very dense crowds). The main characteristic of individuals in category (3) is that they actively increase the density in a group, for example by being disproportionately close to the person in front (and at the same time having more space to the rear), overtaking by using gaps or changing lines. Overall, they are moving fast (fast penguin-like waddle motion) and have more body contact with others. The arms can again be used for stabilization, but are also increasingly used to exert force on others. According to the labels of the categories, category (4) is more extreme than the previous one. Individuals in this category move very quickly through the crowd and are strongly focused on their goal or the best way to get there, respectively. They not only use existing gaps but create them by pushing others aside or even pulling them back or using their shoulder as a plow. Thus, they have the most body contact of all the categories.

It is important to note that a person does not have to meet all the defined characteristics to be assigned to one category, it is rather about the overall dominant behavior. Additionally, the classification is context-dependent, in other words, it depends on the motivation and the behavior of the entire group. So, in a low-motivation group, where people generally keep distance to each other, someone who is disproportionately close to the person in front might get classified as a mild pusher, whereas the same behavior in a high motivation group, in which it is very dense anyway, is still categorized as just walking. To assign a category to each person at each point in time, two human raters independently watch the video recordings person by person and use the annotation function of the software Petrack [82]. However, only behaviors that last at least two to three seconds are considered in order to avoid rating momentary behavioral changes that are likely to occur by chance rather than intentionally. Afterwards, the median is used to condense the rating to a minimum unit of one second, since the temporal resolution of Petrack is 25 frames per seconds, which is too fast for human behavior.

Besides the detailed description of the categories and the rating process, Publication I also showed that the rating method has proven itself in practice and produces reliable results. In a test data set [83], two trained raters achieved an overlap of 90.5% (KALPHA = .79) between their ratings. This was an extremely satisfactory result, because due to the complexity of the process (e.g., the context dependency and the requirement of assessing small behaviors that are difficult to detect from above), two raters are unlikely to obtain completely identical results. Nevertheless, as long as the initial agreement is sufficiently high, a unified data set can be created by correcting small time slippages (up to two seconds) and then forming a compromise between the two ratings.

2.2 Publication II

After the methodological development, the rating system was applied to existing video footage from an earlier bottleneck experiment [84] in Publication II ("Psychological Pushing Propagation in Crowds – Does the Observation of Pushing Behavior Promote Further Intentional Pushing?") to address the question of whether pushing behavior propagates in crowds. In total, the behavior of $N = 776$ participants in 14 different experimental runs (7 low and 7 high motivation ones) was again manually assessed by two trained raters and a unified data set was created. Since only the difference between pushing and non-pushing behavior was of interest in this study, categories (1) and (2) as well as categories (3) and (4) of the original rating system were merged. Furthermore, a Python script was developed to automatically identify who is next to whom in the crowd by using the information of head trajectories or, more precisely, the Voronoi technique [85]. Neighborhood was defined as two individuals sharing a border of their Voronoi cell and being no further than 80 cm apart. These two data sets were then combined resulting in a table that contained, among other things, each participant's own behavior (i.e., pushing or non-pushing), their neighbors, and their neighbors' behavior (i.e., pushing or non-pushing) at each point in time.

The analysis of this data set revealed that there was indeed a small but significant influence of the participants' behavior in the direct neighborhood on a person's behavior. More precisely, the likelihood of a person starting to push increased the more pushers were nearby – this was true for both low motivation and high motivation runs. In order to better delineate this effect, four different measurements of the number of pushing neighbors were calculated, each taking a different aspect into account. The most basic measurement only included the absolute number of pushers who were in the vicinity of a person during the last three seconds. The second measurement considered the total number of neighbors during these three seconds as well and represented the relative proportion of pushers. Likewise, the two remaining measurements captured the absolute and relative number of pushers, but this time not only added up over the last three seconds but over the entire course of the respective run so far. The results showed that the most basic measurement yielded the best prediction of whether or not a person started pushing in the next second, indicating a short-term influence of neighbors' behavior. However, the effect was rather small, as the probability of starting to pushing in the next second was only around 30%, even after strong pushing exposure.

2.3 Publication III

Publication III ("A Rumor has Spread like Wildfire? - Empirical Investigation of Information Propagation in Waiting Crowds") presents two bottleneck studies that examine the systematics of information propagation in a waiting crowd. In the main experiment, five different groups of participants ($n = 33$ to 41) each underwent seven runs with the goal of investigating how information propagation changes when participants know or do not know about their task, when the information is or is not relevant to them, when the information is input at different sides of the group, and when the information is additionally linked to a behavior (i.e., a command to do something). The procedure was the same for every run: At a signal, the group started to walk through the bottleneck and was interrupted after a few seconds. Then one of the four experimenters standing on each side of the group gave an information (i.e., a mere verbal message or a command) to one person who was to pass it on. After a waiting period of 35-120 seconds, the bottleneck was opened again and the participants filled in a questionnaire before regathering in the experimental area. Results showed that propagation was way more effective when participants were properly instructed about their task. In cases they did not know and the experimenter just asked the first participant to pass on the information, the propagation was slow and the information subsided after just a few participants – independent of whether it was relevant or not. However, when explicitly communicated from the beginning that sharing the information was the task in the experiment, many participants followed this instruction and a high percentage were informed across all runs. Overall, a high level of commitment was observed, as some participants were informed multiple times and some also shared the information more than once. In addition, participants looked around to see who may not have received the information. Despite all this, some participants just did not pass on the information, so that it subsided in several runs without everyone being informed. Generally, propagation was faster in the later runs (i.e., when participants were more practiced) and there was a tendency for higher density to lead to faster propagation – at least at the low to moderate densities in this experiment. Further, visual propagation played a rather subordinate role, as the command was often only carried out once it had been received verbally. Although not significant, the message tended to propagate faster than the command, which is reasonable given that the sentence to be transmitted was longer in the command conditions than in the message conditions.

The analysis of the direction of propagation further revealed that there was no preferred direction of communication, but that the information was usually passed on in the same direction from which it came. In some runs, the information also changed, most frequently with the command "tap yourself on your own shoulder". This often became "tap others on the shoulder," which seems to be a natural misinterpretation of the original sentence. Overall, however, participants were eager to share the correct message and to execute the correct command, as evidenced by people looking around to make sure they had understood correctly, as well as people executing an incorrect behavior correcting themselves or being corrected by others. The main findings of this experiment were also confirmed by two runs in the second experiments, in which the procedure was repeated with two larger groups of participants ($n = 91$ and $n = 101$).

2.4 Publication IV

For Publication IV ("It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms"), a mixed-method study with German samples was conducted consisting of an online questionnaire and semi-structured interviews. The aim was to examine lay peoples' everyday understanding of the German term "Massenpanik" and two alternative terms, namely "Massenunglück" ("mass accident"), or "Massendesaster" ("mass disaster").

In the questionnaire, the participants were asked about their general ideas about crowd accidents, the perceived level of danger, the sources of danger, the options for action to defuse the situation, the causes of the occurrence, the parties responsible, their associations and familiarity with the respective term, and their source of knowledge about crowd accidents. Initially, the total sample of $N = 282$ participants was divided among the three terms, meaning that each participant saw only one of them. In other words, the questions were the same for each participant, but framed with a different term each time. However, it turned out that the three terms did not differ in their underlying ideas and assumptions. When asked how familiar participants were with the respective term, mass panic was rated as slightly more familiar than mass disaster, but not than mass accident, and of the many other items, there was only one that showed a significant difference – again between mass panic and mass disaster. Yet, these differences were very small (less than 0.5 scale points). The three conditions were therefore combined for all further analyses.

Overall, participants strongly agreed with items reflecting the image of mass panic, whereas they tended to disagree with the idea of orderly behavior. Nevertheless, they generally believed that people also help each other in critical situations. Further, they stated that not all people behave in the same way, but exhibit heterogenous behavior in critical situations. Another main finding was that all provided causes for and dangers of crowd accidents were rated as equally plausible, suggesting that participants perceive crowd accidents as something very dangerous that can happen for a variety of reasons, but do not have more differentiated assumptions. For instance, the idea of "contagious panic" as a cause was deemed just as plausible as "overcrowding". In contrast, there were clear beliefs on how to behave in a critical situation, with many of the options considered appropriate based on the advice "Don't panic and stay calm". When it comes to responsibility for crowd accidents, participants rated organizers of the event as the most responsible, followed by visitors and security services, and the police the least responsible. The knowledge about crowd accidents is commonly gained from media and social media.

The semi-structured interviews were conducted to explore more openly what lay people associate with the term "mass panic" and the two alternative terms. The main difference between the methodology of the questionnaire and the interview study was that the interviewees were asked about their associations with all three terms (rather than just one). The transcripts of the 17 interviews were analyzed using the qualitative content analysis [86] to determine lay peoples' everyday understanding of mass panic as well as associative fields. The key findings were that, similar to the questionnaires, interviewees reproduce the image of mass panic by describing, for example, a prevalence of decreased rationality and egocentrism in critical crowd situations. Some interviewees stated that the only thing people would care about was their own children. Furthermore, there was a strong link to biological concepts such as instincts (of self-preservation) and swarm behavior. Surprisingly, the notion of contagion and the associative field of spreading a disease were barely used. While many said that panic in a crowd intensifies and people infect each other, no one described this process in detail. Overall, however, most interviewees were dissatisfied with their own "mass panic explanation" as it did not correspond to their image of humans. Nevertheless, they also had hardly any associations with the two alternative terms "mass accident" and "mass disaster" and, in contrast to the questionnaire results, perceived these terms as unfamiliar and rather unsuitable for the crowd context.

CHAPTER 3

Discussion and Outlook

Taking all the results from the four publications together, this PhD project makes a significant contribution to research on crowd dynamics, but also to research in social psychology. Until now, research on mutual influence in social psychology has often been limited to smaller social groups and dyads, and research on crowd dynamics has often completely neglected these psychological phenomena. Models on crowd dynamics, in turn, have already incorporated various forms of people influencing each other, but they often lack an empirical basis. By combining theoretical knowledge, previous research findings and methods of both scientific fields, light has now been shed on the mutual influence and propagation of stimuli in crowds.

The first major contribution of this PhD project is Publications II and III, showing that there are examples of people in crowds influencing each other. In these two studies, this was shown using the examples of pushing behavior and information propagation. However, the mutual influence was less strong and dramatic than originally assumed by Le Bon [9]. Although Publication II revealed an effect of psychological pushing propagation, even a strong pushing exposure only increased the probability of starting to push to up to 30%. Similarly, Publication III has shown that information can be disseminated in groups, but that this does not just happen. These findings are also in line with the previous literature [41] and confirm the conclusion of Haghani et al. [69] that the idea of so-called herd behavior in crowds has to be rejected – also beyond the evacuation context.

Secondly, Publications III and IV contribute to research on the image of mass panic. While Publication III supports the scientific criticism by challenging the notion that information (e.g., a rumor) spreads like wildfire in a crowd, Publication IV brings the perspective of lay people into play. Overall, it was shown that the image of mass panic in general is still quite widespread among lay people and cannot be easily changed by using an alternative term. However, the concrete nature of the underlying assumptions differed somewhat between the questionnaire and interview study. In the questionnaire, participants mainly agreed with the items related to the concept of (panic) contagion (in an epidemiological sense), but also stated that not everyone behaves in the same way and that people may help each other in critical crowd situations. In the interviews, in turn, the answers were shaped by a strong notion of egocentrism and swarm behavior. The image of contagion was mentioned, too, but not explained in detail. Nevertheless, it became clear that lay people believe that panic is a decisive factor in the occurrence of a crowd accident.

The third contribution concerns the methodological and conceptual progress made in Publication I. The combination of the psychological method of questionnaires with physical measurements is already established in crowd research [48, 87–89]. However, to the best of my knowledge, there are no other observation and rating methods to quantify crowd behavior on an individual level. Although this method is very time-consuming and not feasible for a large number of videos or for detecting pushing in real time, its conceptual elaboration provides an important basis for further research on pushing. On the one hand, it can be used for the analysis of other research questions, such as the relationship between forward motion and pushing behavior with people's spatiotemporal characteristics [90]. On the other hand, the manually generated data could be used to train a system for the automatic detection of pushing [91, 92], which in turn could one day be used for real-time detection and thus support crowd management. Further, Publication I makes it clear how heterogenous (pushing) behavior in crowds is: It can vary not only from person to person, but also over time. Of course, this does not only apply to pushing, but can be extended to a variety of other behaviors and psychological states such as motivation and helping behavior. Therefore, this work has broader conceptual implications in terms of heterogeneity and temporal dynamics in crowds that may have a lasting effect on the understanding of crowd dynamics.

However, further research into the psychological mechanisms of mutual influence and propagation in crowds is still needed. In section 1.2, I presented a selection of concepts that make it conceivable that people in crowds influence each other, and the experiments then provided evidence that this influence does indeed take place. Nevertheless, it cannot be deduced from my results whether, for example, the psychological pushing propagation results from the formation of a social norm or is due to competition or revenge. Likewise, the theoretical considerations need to be further specified. For example, it is not yet clear in which cases it is justified to speak of a newly established social norm and in which cases the propagation was only due to imitation resulting from other motives. Moreover, it is important to keep in mind that different factors can interact. This not only concerns the interaction of different social factors (e.g., the norm of helping has usually been established, but some people also perceive unfairness), but also the influence of individual factors (e.g., different motivation, personality traits, gender) and situational factors (e.g., the circumstances at a concert, a conference or a department store sale). The interplay of all these factors can therefore influence whether a person performs a certain behavior, receives information about a certain circumstance, or is influenced by others.

When thinking about future research questions on propagation in crowds, emotional contagion would be an important aspect to consider. Although the idea that emotions spread automatically and inevitably, much like a disease, is empirically untenable, emotional contagion is the concept that underlies many other assumptions about crowds, including the image of mass panic. Since there is empirical evidence that emotional contagion works, at least in principle, between two people, it is compelling to examine whether it also occurs outside of dyads. However, research into emotional contagion in crowds is very challenging. On the one hand, the spatiotemporal characteristics are more complex, as the neighborhood relations change and the question arises to who actually perceives whom. On the other hand, it is also more difficult to quantify emotions. In a crowd, people are usually standing and moving around, so it is not possible to connect them to devices that are classically used in emotional contagion experiments, such as fEMG. Although facial recognition software exists today, in my experience it still needs to be improved before it can be used for actual research purposes. A major problem that emerged during my initial tests was that the most frequently recognized emotion was joy. With the system I used, it was difficult to recognize other emotions such as anger or sadness, as the system only responded to very prototypical expressions of these emotions, which of course are not present in every person. A second problem was that the system also recognized other facial expressions in which the corner of the mouth was pulled slightly upwards as joy, although it was more an expression of an unpleasant emotion. Other challenges included the movement of the people themselves, e.g. turning their faces away or increasing the distance to the camera, as well as overlaps and occlusions. Similar to the rating system of forward motion, manual raters could be used; however, manual raters also require video recordings in which the people and especially their faces are clearly visible, which could be difficult to realize in crowd situations.

Besides the contribution of this thesis to theoretical knowledge about crowds, the results also have some practical implications for crowd management. Firstly, Publication IV suggests that people are convinced that the advice “Don’t panic (and stay calm)” is useful in critical crowd situation. However, this also means that they would not draw attention to the danger or themselves if they noticed something or felt uncomfortable. Even though crowd managers should not rely solely on the feedback from the crowd anyway, they should not automatically assume that everything is fine just because no one in the crowd expresses any discomfort. Secondly, Publication III provides advice on how person-to-person information sharing can be used to complement other communication channels. For example, to increase the likelihood of sharing, it is important to engage visitors in this strategy upfront. Asking them to share information only at the moment it becomes relevant is unlikely to work well. However, if people feel genuinely committed to taking on this task, they will be more willing to comply. To ensure that information is not only shared but also shared correctly, messages and commands should be designed as short and clear as possible. If the triggered behavior is then also visible for a longer period of time, there is an additional likelihood that people who have performed the action incorrectly will correct themselves or will be corrected by others.

In summary, this thesis clearly shows that not only physical influences between agents (e.g., contact forces, collision avoidance, friction) must be considered in the research and modeling of crowd dynamics, but also psychological influences. Although there are already models that have attempted to integrate pushing propagation [4, 6] or information sharing [7, 8], this has mostly been done without a psychological or empirical basis so far. However, the inclusion of a new parameter or even just changing the settings of a parameter can have an enormous impact on the predictive power of a model. In addition to the methodological contribution, the contribution to the expansion of knowledge and hopefully the inspiration for many other empirical studies, the results of this PhD project also have the potential to make crowd models more realistic and improve their predictive power, thus making crowds safer in the future.

Bibliography

[1] Claudio Feliciani, Kenichiro Shimura, and Katsuhiro Nishinari. *Introduction to Crowd Management: Managing Crowds in the Digital Era: Theory and Practice*. Cham: Springer International Publishing, 2021. DOI: 10.1007/978-3-030-90012-0.

[2] Anna Sieben, Jette Schumann, and Armin Seyfried. "Collective phenomena in crowds—Where pedestrian dynamics need social psychology". In: *Plos One* 12 (2017), e0177328. DOI: 10.1371/journal.pone.0177328.

[3] Funda Durupinar, Ugur Gudukbay, Aytek Aman, and Norman I. Badler. "Psychological Parameters for Crowd Simulation: From Audiences to Mobs". In: *IEEE Transactions on Visualization and Computer Graphics* 22 (2016), pp. 2145–2159. DOI: 10.1109/TVCG.2015.2501801.

[4] Colin M. Henein and Tony White. "Agent-Based Modelling of Forces in Crowds". In: *Multi-Agent and Multi-Agent-Based Simulation. MABS 2004. Lecture notes in Computer Science, vol 3415*. Ed. by P. Davidsson, B. Logan, and K. Takadama. Berlin, Heidelberg: Springer, 2005, pp. 173–184. DOI: 10.1007/978-3-540-32243-6_14.

[5] Chaochao Li, Pei Lv, Dinesh Manocha, Hua Wang, Yafei Li, Bing Zhou, and Mingliang Xu. "ACSEE: Antagonistic Crowd Simulation Model with Emotional Contagion and Evolutionary Game Theory". In: *IEEE Transactions on Affective Computing* (2019), pp. 1–18. DOI: 10.1109/TAFFC.2019.2954394. arXiv: 1902.00380.

[6] Nuria Pelechano, Jan M. Allbeck, and Norman I. Badler. "Controlling Individual Agents in High-Density Crowd Simulation". In: *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. 2007, pp. 99–108. URL: https://repository.upenn.edu/hms/210.

[7] Chen Yehua and Bai Jing. "The dynamics of pedestrians' evacuation during emergency situations". In: *Adaptive Behavior* 24 (2016), pp. 399–410. DOI: 10.1177/1059712316672517.

[8] Qiling Zou and Suren Chen. "Simulation of Crowd Evacuation under Toxic Gas Incident Considering Emotion Contagion and Information Transmission". In: *American Society of Civil Engineers* 34 (2020), pp. 1–18. DOI: 10.1061/(ASCE)CP.1943-5487.0000889.

[9] Gustave Le Bon. *The crowd: A study of the popular mind.* New York: Macmillan Company, 1896.

[10] Herbert Blumer. "Collective behaviour". In: *Principles of sociology.* Ed. by Robert E. Park. New York, NY: Barnes & Noble, 1939, pp. 219–288.

[11] Robert Ezra Park. *Masse und Publikum: eine methodologische und soziologische Untersuchung.* Buchdruckerei Lack & Grunau, 1904.

[12] John Drury and Clifford Stott. "Contextualising the crowd in contemporary social science". In: *Contemporary Social Science* 6 (2011), pp. 275–288. DOI: 10.1080/21582041.2011.625626.

[13] Clark McPhail. *The myth of the madding crowd.* London: Routledge, 1991.

[14] John Drury and Stephen Reicher. "Crowds and Collective Behavior". In: *Oxford Research Encyclopedia of Psychology.* 2020. DOI: 10.1093/acrefore/9780190236557.013.304.

[15] Zachary J Warren and Séamus A Power. "It's contagious: Rethinking a metaphor dialogically". In: *Culture & Psychology* 21 (2015), pp. 359–379. DOI: 10.1177/1354067X15601190.

[16] BBC News. *Iraq stampede deaths near 1,000.* 2005-08-31. URL: http://news.bbc.co.uk/2/hi/middle_east/4199618.stm (visited on 09/15/2023).

[17] DNA Web Team. *Patna stampede: Rumours may have a hand in tragedy; Bihar CM Jiten Ram Manjhi promises action, trashes 'Dummy CM' tag.* 2014-10-05. URL: https://www.dnaindia.com/india/report-patna-stampede-rumours-may-have-a-hand-in-tragedy-bihar-cm-jiten-ram-manjhi-promises-action-trashes-dummy-cm-tag-2023339 (visited on 09/15/2023).

[18] Chris Cocking, John Drury, and Steve Reicher. "The psychology of crowd behaviour in emergency evacuations: Results from two interview studies and implications for the Fire and Rescue Services". In: *Irish Journal of Psychology* 30 (2009), pp. 59–73. DOI: 10.1080/03033910.2009.10446298.

[19] Guillaume Dezecache. "Human collective reactions to threat". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 6 (2015), pp. 209–219. DOI: 10.1002/wcs.1344.

[20] Erik Auf der Heide. "Common Misconceptions about Disasters: Panic, the "Disaster Syndrome," and Looting". In: *The First 72 Hours: A Community Approach to Disaster Preparedness.* Ed. by M. O'Leary. Lincoln: iUniverse Publishing, 2004, pp. 340–380.

[21] Lee Clarke. "Panic: Myth or Reality?" In: *Contexts* 1 (2002), pp. 21–26. DOI: 10.1525/CTX.2002.1.3.21.

[22] Chris Cocking and John Drury. "Talking about Hillsborough: 'Panic' as Discourse in Survivors' Accounts of the 1989 Football Stadium Disaster". In: *Journal of Community & Applied Social Psychology* 24 (2014), pp. 86–99. DOI: 10.1002/casp.2153.

[23] John Drury, David Novelli, and Clifford Stott. "Psychological disaster myths in the perception and management of mass emergencies". In: *Journal of Applied Social Psychology* 43 (2013), pp. 2259–2270. DOI: 10.1111/jasp.12176.

[24] Dirk Helbing and Pratik Mukerji. "Crowd disasters as systemic failures: Analysis of the love parade disaster". In: *EPJ Data Science* 1 (2012), pp. 1–40. DOI: 10.1140/epjds7.

[25] Alexander Mintz. "Non-Adaptive Group Behavior". In: *The journal of abnormal and social psychology* 46 (1951), pp. 150–159.

[26] Enrico L Quarantelli. "Images of Withdrawal Behavior in Disasters: Some Basic Misconceptions". In: *Social Problems* 8 (1960), pp. 68–79. URL: https://www.jstor.org/stable/798631.

[27] Enrico L. Quarantelli. "Panic, Sociology of". In: *International encyclopedia of the social and behavioral sciences*. Ed. by N. J. Smelser and P. B. Baltes. New York: Pergamon Press, 2001, pp. 11020–11023.

[28] Jonathan D Sime. "The concept of panic". In: *Fires and human behaviour*. Ed. by D. Canter. London: David Fulton, 1990, pp. 63–81.

[29] Elaine Hatfield, John T. Cacioppo, and Richard L. Rapson. "Emotional contagion". In: *Current Directions in Psychological Science* 2 (1993), pp. 96–99.

[30] Elaine Hatfield, Megan Carpenter, and Richard L. Rapson. "Emotional contagion as a precursor to collective emotions". In: *Collective emotions: Perspectives from psychology, philosophy, and sociology*. Ed. by C. von Scheve and M. Salmela. New York, NY: Oxford University Press, 2014, pp. 108–122. DOI: 10.1093/acprof:oso/9780199659180.003.0008.

[31] Guillaume Dezecache, Terry Eskenazi, and Julie Grèzes. "Emotional Convergence: A Case of Contagion?" In: *Shared Representations: Sensorimotor Foundations of Social Life*. Ed. by Sukhvinder D. Obhi and Emily S. Cross. Cambridge University Press, 2016, pp. 417–436. DOI: 10.1017/cbo9781107279353.021.

[32] Huan Deng and Ping Hu. "Matching your face or appraising the situation: Two paths to emotional contagion". In: *Frontiers in Psychology* 8 (2018). DOI: 10.3389/fpsyg.2017.02278.

[33] Ursula Hess and Sylvie Blairy. "Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy". In: *International Journal of Psychophysiology* 40 (2001), pp. 129–141. DOI: 10.1016/S0167-8760(00)00161-6.

[34] Masanori Kimura, Ikuo Daibo, and Masao Yogo. "The study of emotional contagion from the perspective of interpersonal relationships". In: *Social Behavior and Personality: an international journal* 36 (2008), pp. 27–42.

[35] Job van der Schalk, Agneta Fischer, Bertjan Doosje, Daniël Wigboldus, Skyler Hawk, Mark Rotteveel, and Ursula Hess. "Convergent and divergent responses to emotional displays of ingroup and outgroup". In: *Emotion* 11 (2011), pp. 286–298. DOI: 10.1037/a0022582.

[36] Monika Wróbel, Klara Królewiak, and Anna Z. Czarna. "Do I Mirror Your Mood if Were Peas in a Pod? Similarity and Liking in the Social Induction of Affect". In: *Journal of Social Psychology* 155 (2015), pp. 636–649. DOI: 10.1080/00224545.2015.1047437.

[37] Monika Wróbel and Klara Krolewiak. "Is your mood more contagious if you are likeable? the role of liking in the social induction of affect". In: *Polish Psychological Bulletin* 46 (2015), pp. 413–420. DOI: 10.1515/ppb-2015-0048.

[38] Sigal G. Barsade. "The ripple effect: Emotional contagion and its influence on group behavior". In: *Administrative Science Quarterly* 47 (2002), pp. 644–675. DOI: 10.2307/3094912.

[39] Guillaume Dezecache, Laurence Conty, Michele Chadwick, Leonor Philip, Robert Soussignan, Dan Sperber, and Julie Grèzes. "Evidence for Unintentional Emotional Contagion Beyond Dyads". In: *PLoS ONE* 8 (2013), e67371. DOI: 10.1371/journal.pone.0067371.

[40] Milad Haghani and Majid Sarvi. "Following the crowd or avoiding it? Empirical investigation of imitative behaviour in emergency escape of human crowds". In: *Animal Behaviour* 124 (2017), pp. 47–56. DOI: 10.1016/J.ANBEHAV.2016.11.024.

[41] Andrew C. Gallup, Joseph J. Hale, David J.T. Sumpter, Simon Garnier, Alex Kacelnik, John R. Krebs, and Iain D. Couzin. "Visual attention and the acquisition of information in human crowds". In: *Proceedings of the National Academy of Sciences of the United States of America* 109 (2012), pp. 7245–7250. DOI: 10.1073/pnas.1116141109.

[42] Robert Boyd and Peter J. Richerson. *The origin and evolution of cultures.* Oxford, New York: Oxford University Press, 2005.

[43] Gerd Gigerenzer. "Why Heuristics Work". In: *Perspectives on Psychological Science* 3 (2008), pp. 20–29. DOI: 10.1111/j.1745-6916.2008.00058.x.

[44] Anna Sieben and Armin Seyfried. "Inside a life-threatening crowd: Analysis of the Love Parade disaster from the perspective of eyewitnesses". In: *Safety Science* 166 (2023), p. 106229. DOI: doi.org/10.1016/j.ssci.2023.106229.

[45] Ralph H. Turner and Lewis M. Killian. *Collective behavior*. 3rd ed. Englewood Cliffs, N.J: Prentice-Hall, 1987.

[46] Leon Mann. "Queue Culture: The Waiting Line as a Social System". In: *American Journal of Sociology* 75 (1969), pp. 340–354. DOI: 10.1086/224787.

[47] Bernd H Schmitt, Laurette Dube, and France Leclerc. "Intrusions Into Waiting Lines: Does the Queue Constitute a Social System?" In: *Journal of Personality and Social Psychology* 63 (1992), pp. 806–815. DOI: 10.1037/0022-3514.63.5.806.

[48] Juliane Adrian, Armin Seyfried, and Anna Sieben. "Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology". In: *Journal of the Royal Society Interface* 17 (2020), p. 20190871. DOI: 10.1098/rsif.2019.0871.

[49] Henri Tajfel and John C. Turner. “An integrative theory of intergroup conflict”. In: *The social psychology of intergroup relations*. Ed. by William G. Austin and Stephen Worchel. Monterey, CA: Brooks/Cole, 1979, pp. 33–37.

[50] John Drury. “The role of social identity processes in mass emergency behaviour: An integrative review”. In: *European Review of Social Psychology* 29 (2018), pp. 38–81. DOI: 10.1080/10463283.2018.1471948.

[51] John Drury, Chris Cocking, Steve Reicher, Andy Burton, Damian Schofield, Andrew Hardwick, Danielle Graham, and Paul Langston. “Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model”. In: *Behavior Research Methods* 41 (2009), pp. 957–970. DOI: 10.3758/BRM.41.3.957.

[52] John Drury, Rupert Brown, Roberto González, and Daniel Miranda. “Emergent social identity and observing social support predict social support provided by survivors in a disaster: Solidarity in the 2010 Chile earthquake”. In: *European Journal of Social Psychology* 46 (2016), pp. 209–223. DOI: 10.1002/ejsp.2146.

[53] Gabrielle Riches. “Embracing the Chaos: Mosh Pits, Extreme Metal Music and Liminality”. In: *Journal for Cultural Research* 15 (2011), pp. 315–332. DOI: 10.1080/14797585.2011.594588.

[54] Jesse L. Silverberg, Matthew Bierbaum, James P. Sethna, and Itai Cohen. “Collective Motion of Humans in Mosh and Circle Pits at Heavy Metal Concerts”. In: *Physical Review Letters* 110 (2013), p. 228701. DOI: 10.1103/PhysRevLett.110.228701.

[55] Sina Feldmann and Juliane Adrian. “Forward propagation of a push through a row of people”. In: *Safety Science* 164 (2023), p. 106173. DOI: 10.1016/j.ssci.2023.106173.

[56] Dean C. Barnlund. “A Transactional Model of Communication”. In: *Foundations of Communication Theory*. Ed. by Kenneth K. Sereno and C. David Mortensen. New York, NY: Harper & Row, 1970.

[57] Wilbur Ed Schramm. *The process and effects of mass communication*. Urbana: University of Illinois Press, 1954.

[58] Claude Elwood Shannon and Warren Weaver. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.

[59] Norris R. Johnson. “Panic at “The Who Concert Stampede”: An Empirical Assessment”. In: *Social Problems* 34 (1987), pp. 362–373. DOI: 10.2307/800813.

[60] Guylene Proulx and Jonathan D Sime. “To Prevent ’Panic’ in an Underground Emergency: Why Not Tell People the Truth?” In: *Fire Safety Science: Third International Symposium*. Ed. by Geoffrey Cox and Brian Langford. London: Elsevier Applied Science, 1991, pp. 843–852.

[61] David Purser. "Comparisons of Evacuation Efficiency and Pre-travel Activity Times in Response to a Sounder and Two Different Voice Alarm Messages". In: *Pedestrian and Evacuation Dynamics 2008*. Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 121–134. DOI: 10.1007/978-3-642-04504-2_9.

[62] Anne Templeton, Anezka Beamish Leskovcova, Gareth R. Clegg, Adam Lloyd, Elvira Fraser, Sentia Keyulong, and Sayaka Hinata. *Effective strategies in emergencies: First responders' views on communicating and coordinating with the public*. 2023. URL: https://www.research.ed.ac.uk/en/publications/effective-strategies-in-emergencies-first-responders-views-on-com (visited on 10/08/2023).

[63] C. Natalie van der Wal, Mark A. Robinson, Wändi Bruine de Bruin, and Steven Gwynne. "Evacuation behaviors and emergency communications: An analysis of real-world incident videos". In: *Safety Science* 136 (2021). DOI: 10.1016/j.ssci.2020.105121.

[64] John Drury, Chris Cocking, and Steve Reicher. "The Nature of Collective Resilience: Survivor Reactions to the 2005 London Bombings". In: *International Journal of Mass Emergencies and Disasters* 27 (2009), pp. 66–95.

[65] John Drury, Chris Cocking, and Steve Reicher. "Everyone for themselves? A comparative study of crowd solidarity among emergency survivors". In: *British Journal of Social Psychology* 48 (2009), pp. 487–506. DOI: 10.1348/014466608X357893.

[66] James R. Gill and Kristen Landi. "Traumatic asphyxial deaths due to an uncontrolled crowd". In: *American Journal of Forensic Medicine and Pathology* 25 (2004), pp. 358–361. DOI: 10.1097/01.PAF.0000147316.62883.8B.

[67] Maria Moitinho de Almeida and Johan von Schreeb. "Human Stampedes: An Updated Review of Current Literature". In: *Prehospital and Disaster Medicine* 34 (2019), pp. 82–88. DOI: 10.1017/S1049023X18001073.

[68] Luis-Manuel Garcia. "Pathological Crowds: Affect and Danger in Responses to the Love Parade Disaster at Duisburg". In: *Dancecult: Journal of Electronic Dance Music Culture* 2 (2011), pp. 6–11. DOI: 10.12801/1947-5403.2011.02.01.15.

[69] Milad Haghani, Emiliano Cristiani, Nikolai W.F. Bode, Maik Boltes, and Alessandro Corbetta. "Panic, Irrationality, and Herding: Three Ambiguous Terms in Crowd Dynamics Research". In: *Journal of Advanced Transportation* (Hindawi Limited 2019). DOI: 10.1155/2019/9267643.

[70] Laura Künzer, Robert Zinke, and Gesine Hofinger. "Mythen der Entfluchtung". In: *Tagungsband Jahresfachtagung*. Ed. by Vereinigung zur Förderung des Deutschen Brandschutzes eV. 2012, pp. 725–735.

[71] Tatsuya Nogami. "Disaster myths among disaster response professionals and the source of such misconceptions". In: *Journal of Contingencies and Crisis Management* 26 (2018), pp. 491–498. DOI: 10.1111/1468-5973.12218.

[72] Tatsuya Nogami and Fujio Yoshida. “Disaster myths after the Great East Japan Disaster and the effects of information sources on belief in such myths”. In: *Disasters* 38 (2014), pp. 190–205. DOI: 10.1111/disa.12073.

[73] Ed Galea. “It’s a crowd crush NOT A STAMPEDE - the difference is important!” In: *LinkedIn* (2021-04-30). URL: https://www.linkedin.com/pulse/its-crowd-crush-stampede-difference-important-ed-galea/ (visited on 12/07/2021).

[74] Gianna-Carina Grün. “Was sind crowd dynamics? - Forschungsfrage”. In: *jetzt* (2011-07-22). URL: https://www.jetzt.de/forschungsfrage/was-sind-crowd-dynamics-528031 (visited on 12/07/2021).

[75] Micha Hörnle. “Letzte ”Nachtschicht” in Heidelberg: ”Massenpanik” ist ein Mythos”. In: *Rhein-Neckar-Zeitung* (2018-01-02). URL: https://www.rnz.de/nachrichten/heidelberg_artikel,-letzte-nachtschicht-in-heidelberg-massenpanik-ist-ein-mythos-plus-fotogalerie-_arid,327609.html (visited on 12/07/2021).

[76] Claudio Feliciani, Iker Zuriguel, Angel Garcimartín, Diego Maza, and Katsuhiro Nishinari. “Systematic experimental investigation of the obstacle effect during non-competitive and extremely competitive evacuations”. In: *Scientific Reports* 10 (2020), p. 15947. DOI: 10.1038/s41598-020-72733-w.

[77] A Garcimartín, D R Parisi, J M Pastor, C Martín-Gómez, and I Zuriguel. “Flow of pedestrians through narrow doors with different competitiveness”. In: *Journal of Statistical Mechanics: Theory and Experiment* (2016), p. 043402. DOI: 10.1088/1742-5468/2016/04/043402.

[78] Milad Haghani, Majid Sarvi, and Zahra Shahhoseini. “When ‘push’ does not come to ‘shove’: Revisiting ‘faster is slower’ in collective egress of human crowds”. In: *Transportation Research Part A: Policy and Practice* 122 (2019), pp. 51–69. DOI: 10.1016/j.tra.2019.02.007.

[79] José M. Pastor, Angel Garcimartín, Paula A. Gago, Juan P. Peralta, César Martín-Gómez, Luis M. Ferrer, Diego Maza, Daniel R. Parisi, Luis A. Pugnaloni, and Iker Zuriguel. “Experimental proof of faster-is-slower in systems of frictional particles flowing through constrictions”. In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 92 (2015), p. 062817. DOI: 10.1103/PHYSREVE.92.062817/FIGURES/3/MEDIUM.

[80] Nicola Döring and Jürgen Bortz. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5th ed. Springer-Lehrbuch. Berlin, Heidelberg: Springer, 2016. DOI: 10.1007/978-3-642-41089-5.

[81] Kimberly A. Neuendorf. *The Content Analysis Guidebook*. 2nd ed. California: SAGE Publications, Inc, 2017. DOI: 10.4135/9781071802878.

[82] Maik Boltes, Armin Seyfried, Bernhard Steffen, and Andreas Schadschneider. “Automatic Extraction of Pedestrian Trajectories from Video Recordings”. In: *Pedestrian and Evacuation Dynamics 2008*. Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 43–54. DOI: 10.1007/978-3-642-04504-2_3.

[83] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Entrance 2, entry with guiding barriers (corridor setup) [Data set]*. 2013. DOI: 10.34735/ped.2013.1.

[84] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Crowds in front of bottlenecks from the perspective of physics and social psychology [Data set]*. 2018. DOI: 10.34735/ped.2018.1.

[85] Georges Voronoi. "Nouvelles applications des paramètres continus à la théorie des formes quadratiques." In: *Journal für die reine und angewandte Mathematik* 133 (1908), pp. 97–178.

[86] Philipp Mayring. "Qualitative Content Analysis [28 paragraphs]. Forum Qual. Sozialforsch". In: *Forum Qual. Soc. Res. 1, Art. 20* (2000). DOI: 10.17169/fqs-1.2.1089.

[87] Xiaolu Jia, Claudio Feliciani, Hisashi Murakami, Akihito Nagahama, Daichi Yanagisawa, and Katsuhiro Nishinari. "Revisiting the level-of-service framework for pedestrian comfortability: Velocity depicts more accurate perceived congestion than local density". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 87 (2022), pp. 403–425. DOI: 10.1016/j.trf.2022.04.007.

[88] Krisztina Konya and Anna Sieben. "Waiting and walking with strangers: a socio-psychological pedestrian experiment on joint action in anonymous situations". In: *Royal Society Open Science* 10 (2023), p. 221601. DOI: 10.1098/rsos.221601.

[89] Mira Küpper and Armin Seyfried. "Waiting in crowded places: influence of number of pedestrians, waiting time and obstacles". In: *Journal of The Royal Society Interface* 20 (2023), p. 20230193. DOI: 10.1098/rsif.2023.0193.

[90] Ezel Üsten, Jette Schumann, and Anna Sieben. "Exploring the Dynamic Relationship between Pushing Behavior and Crowd Dynamics". In: *Collective Dynamics* 8 (2023), pp. 1–29. DOI: 10.17815/CD.2023.143.

[91] Ahmed Alia, Mohammed Maree, Mohcine Chraibi, Anas Toma, and Armin Seyfried. "A Cloud-Based Deep Learning Framework for Early Detection of Pushing at Crowded Event Entrances". In: *IEEE Access* 11 (2023), pp. 45936–45949. DOI: 10.1109/ACCESS.2023.3273770.

[92] Ahmed Alia, Mohammed Maree, and Mohcine Chraibi. "A Hybrid Deep Learning and Visualization Framework for Pushing Behavior Detection in Pedestrian Dynamics". In: *Sensors* 22 (2022), p. 4040. DOI: 10.3390/s22114040.

PUBLICATION I

Pushing and Non-pushing Forward Motion in Crowds: A Systematic Psychological Observation Method for Rating Individual Behavior in Pedestrian Dynamics

Note: This article was published as Üsten, E., Lügering, H., & Sieben, A. (2022). 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. https://doi.org/10.17815/CD.2022.138

Author Contributions

Conceptualization: Helena Lügering, Ezel Üsten and Anna Sieben
Data curation: Helena Lügering and Ezel Üsten
Formal analysis: Helena Lügering and Ezel Üsten
Investigation: Helena Lügering and Ezel Üsten
Methodology: Helena Lügering and Ezel Üsten
Supervision: Anna Sieben
Validation: Helena Lügering and Ezel Üsten
Visualization: Helena Lügering and Ezel Üsten
Writing–original draft preparation: Helena Lügering and Ezel Üsten
Writing–review and editing: Helena Lügering, Ezel Üsten and Anna Sieben

Pushing and Non-pushing Forward Motion in Crowds: A Systematic Psychological Observation Method for Rating Individual Behavior in Pedestrian Dynamics

Ezel Üsten[1 2], Helena Lügering[1 2] and Anna Sieben[1 3]

[1] Institute for Advanced Simulation 7: Civil Safety Research, Forschungszentrum Jülich, Jülich, Germany
[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany
[3] Chair of Social Theory and Social Psychology, Ruhr University Bochum, Bochum, Germany

Abstract

Pushing behavior impairs people's sense of well-being in a crowd and represents a significant safety risk. There are nevertheless still a lot of unanswered questions about who behaves how in a crowded situation, and when, where, and why pushing behavior occurs. Beginning from the supposition that a crowd is not thoroughly homogenous and that behavior can change over time, we developed a method to observe and rate forward motion. Based on the guidelines of quantitative content analysis, we came up with four categories: (1) falling behind, (2) just walking, (3) mild pushing, and (4) strong pushing. These categories allow for the classification of the behavior of any person at any time in a video, and thereby the method allows for a comprehensive systematization of individuals' actions alongside temporal crowd dynamics. The application of this method involves videos of moving crowds including trajectories. The initial results show a very good inter-coder reliability between two trained raters with a 90.5% overlap (KALPHA = .79) demonstrating the general suitability of the system to describe forward motion in crowds systematically and quantify it for further analysis. In this way, pushing behavior can be better understood and, prospectively, risks better identified. This article offers a comprehensive presentation of this method of observation.

Keywords: pushing behavior; forward motion; crowd psychology; observation method; content analysis; rating system

1 Introduction

Imagine a crowd of excited fans waiting to enter a concert hall: There is no queuing system and everyone wants to be the first in the hall, for there are no seat reservations either. If you had a bird's eye view to observe this crowd from above, you would likely get the impression that it is just one big throng in which everyone is pushing and shoving. Examining each person individually, however, reveals that the crowd is not actually homogeneous and not everyone is behaving the same. This paper introduces an observation method which focuses on individual behaviors in such crowds and allows for an appraisal of who is pushing at which moment in time to draw a more differentiated picture. The assessment and evaluation of individual behavior is performed by trained observers using videos of crowds and the extracted trajectories.

Crowded situations are common and happen—at least before COVID-19—almost every day. Just think about the jostling at the train station. As ordinary as it may be, the consequences can be very serious. Pushing behavior not only impairs satisfaction during the crowd experience [1], it also poses a safety risk. Different studies show that high motivation, which often involves pushing and shoving, increases density [2, 3], and reports from real-life scenarios indicate that pushing from behind can lead to life-threatening density and pressure resulting in injuries and fatalities [4]. Although there is broad evidence of cooperative behavior in emergencies [4–7], pushing may also occur during evacuations, further increasing the danger. [8] Several simulations of pedestrian crowds have therefore tried to integrate this behavior [9, 10] but without providing a systematic psychological basis.

Aside from the safety issues, pushing and shoving were generally evaluated as inappropriate and unfair in recent studies with a bottleneck set-up [2, 11]. It is quite surprising, though, that the same participants mentioned these behaviors as the most promising strategies for faster access. Whether individuals actually move forward faster by jostling depends, however, on their strength and the density of the crowd. With respect to the crowd as a whole, it has not yet been conclusively determined whether increasing the pressure by pushing changes the flow through the bottleneck. Although it has been suggested that pushing actually decreases the flow—the so called "faster-is-slower" effect [12, 13] — Haghani et al. [3] found no conclusive evidence for this general occurrence in a review of current experimental literature. Their own experiment, however, indicated that at least strong and aggressive pushing prolongs the egress time in a bottleneck situation.

However, not everyone in a crowd pushes to the same extent. In Adrian et al. [2], the percentage of participants engaged in this behavior varied from 29.2 to 78.6%. Reasons for non-pushing were, for example, avoidance of danger or a general aversion to pushing. Additionally, identification with the crowd may influence pushing behavior—high-identification participants tended to push less and to give more help in a mass evacuation scenario [14]. Also, social norms (e.g., triggered by the spatial organization of the crowd) influence whether pushing is appropriate behavior or not. Queuing, for example, is a social system where norms prevail that are opposed to pushing [11, 15, 16]. These results show very clearly that pushing is a complex behavior influenced by several factors. Apart from this general decision for or against pushing, it is also natural that any human behavior is not static but dynamic and can therefore change over time. This means, of course, that pushing behavior is also dynamic and sometimes people push only to stop in the next moment. Researchers addressing crowd dynamics have nevertheless tended thus far to address pushing as a constant behavior in a homogeneous crowd. Our proposed rating method takes into account these fluctuating dynamics of pushing and non-pushing.

But before examining these complex dynamics, it is essential to understand which behaviors are included when talking about pushing. According to the Cambridge Dictionary "(to) push" means "to move forcefully, especially in order to cause someone or something that is in your way to move, so that you can go through or past them" [17]. Further, it must be distinguished between intentional and unintentional pushing [18]. Unintentional pushing is the physical reaction to a push from behind that results in one person being pushed forward into another person. In intentional pushing, on the other hand, individuals exert energy themselves to build up forward pressure. In recent studies [2, 11], participants mentioned the use of elbows, arms, or shoulders, as well as pushing to the front and pushing to the side as different forms of pushing. Additionally, filling gaps is mentioned as a strategy for faster access. It is debatable whether filling gaps is a form of pushing behavior, as it is less aggressive, but it clearly leads to increased density and people moving forward faster. Consequently, for the purpose of our method, we include filling gaps as a form of pushing. This enumeration of possible forms of pushing strongly suggests that simply distinguishing between pushing and non-pushing is too simple to be helpful. Therefore, our method examines two different gradations of pushing, namely, mild and strong. Adapted to this, we also distinguish two gradations of non-pushing: a simple forward movement "with the flow" and a forward movement that is slower than the crowd as a whole and thus "falling back."

The general idea for the observation and rating method is based on quantitative content analysis as used in psychology and the social sciences [19, 20]. With the help of a complete coding system, this method captures the characteristics of a document. The coding system is created before the analysis and contains precise definitions of the characteristic expressions and assigns numbers to them. The details of the coding system, as well as useful examples and explanations for the coding process, are recorded in the codebook. Furthermore, the document is divided into precise units of analysis. The rating is performed by at least two trained raters, and reliability measures serve to ensure their concurrence. While content analysis was initially developed for text documents (such as newspaper articles, diaries), in recent years it has also been adapted for the analysis of images and video material.

Important steps of content analysis for both text and video analysis are [20]: (a) determination of the analysis material, and definition of units of analysis, (b) design of the coding system based on the literature and research questions, (c) tentative application and revision of the coding system, (d) discussion of the validity of the coding system, (e) training of raters, (f) reliability analysis (inter-coder reliability), (g) complete data collection, and (h) statistical evaluation. In this paper, we present our content analytic method for capturing pushing behavior in crowd videos in a step-by-step fashion (with the exception of the last two steps (g and h)—for an analysis of the data at this level has yet to be performed).

2 Method

The method described here uses videos taken of crowds from overhead in confined areas such as in front of bottlenecks. Trained observers pick out individual people one by one and categorize their behavior in every second. To do this, they use a four-level category system that includes pushing and non-pushing behavior. The method is introduced here with a thorough step-by-step explanation, to facilitate its future use by other research groups.

2.1 Determination of the analysis material and definition of units of analysis

Although pushing behavior has been regularly observed in former experiments, an in-depth approach for defining and grading the behavior has not been one of the most prominent objectives in pedestrian dynamics so far. As a result, there is a wealth of video material that can be potentially “recycled” for constituting a base to analyze the behavior (see for example: Pedestrian Dynamics Data Archive [21]). Any video that contains pedestrians in forward motion can be used. The category system can be applied to experiments with very different crowd dynamics (i.e., fast or slow) because this method includes the entire spectrum of pushing and non-pushing behavior. Every participant can be categorized as to the degree and intensity of their behavior, whether pushing is observed or not.

Individual trajectories must be available or first extracted for the video to be evaluated. The detection is done via PeTrack software [22]. PeTrack was mainly developed for automatic extraction of pedestrian trajectories from video recordings that are captured from cameras with a top-down view for measuring the physical properties of crowds (e.g., density). The category system uses these trajectories for individual pedestrians to provide accurate timing (via frame numbers: 1 second is equal to 25 frames) of starting categories, category shifts, ending categories, and their spatial visualizations. PeTrack was upgraded specifically for the current category system; an annotation command and a feature allowing the video to be played in real time were added to the software (Version 0.8.15) in order to have an accurate-timing comment (rating: category 1 to 4) for a specific person and a specific frame. The txt file output shows the rating with the respective frame that is bound to the respective pedestrian.

The rating is executed in specific frames that contain a starting point, an ending point, or a behavior change, for every pedestrian. However, a human observer needs at least one second in order to comprehend the complex behavior (and its potential change in the next second) of an individual and therefore it does not make sense to use the frame units defined in PeTrack. For the category system, a unit of analysis is consequently defined as the behavior of an individual in one second. The frame rate of PeTrack is, however, 25 frames per second. Therefore, it was decided that the median of frame ratings within one second of one participant would be calculated and used as the minimum unit of the rating measure. The process of the rating of pushing behavior is as follows: After the experiment video selection, the ptc (PeTrack) files were gathered from the IAS-7 database and every pedestrian in the chosen video was annotated according to their behavior. The starting point was considered the first frame (usually frame 0) in which PeTrack detects the selected pedestrian, and the ending point was set as either in the last frame of the video or when (if) the pedestrian reaches the bottleneck. In the latter case, the ending frame was always annotated as “END.”

2.2 Design of the coding system on the basis of literature and research questions

As outlined above, pushing is defined as a behavior that can involve using arms, shoulders, or elbows; or simply the upper body, in which one person actively applies force to another person (or people) to overtake, while shifting their direction to the side or back, or force them to move forward more quickly. Pushing usually correlates with speed acceleration. Our approach also includes using gaps as a form of pushing because this is a form of overtaking. We distinguish two gradations of pushing behavior: mild and strong. Accordingly, we also distinguish two gradations of non-pushing forward motion. As a result, a category system with four categories has been created: (1) falling behind, (2) just walking, (3) mild pushing, and (4) strong pushing; as two pushing (3 and 4) and two non-pushing (1 and 2) categories.

Six different parameters were used for rating individuals according to these categories: the position of their arms and hands; the position of their shoulders and heads; their personal space; their interaction with others; speed and acceleration; and attention to the exit. These parameters have different behavioral outputs depending on which category they are in, as can be seen below.

Falling behind (1) is the most passive category in terms of forward motion (Figure 1). People in this category use their hands and arms less. Their arms are generally crossed or dropped by their sides, apart from cases in which they were chatting with other people and using their hands to gesture (arms and hands position). They show frequent head movements because their attention is scattered; they can hence focus on non-specific things in their environment (shoulder and head position). They mostly have some distance to the group and minimal physical contact. In most cases, they are at the back of the crowd, but, when they are in the front, they may actively increase the distance to the person in front by slowing down (personal space). They might be actively involved in chatting with other participants (interaction with others). They are slow overall—even stopping in some cases or changing their direction to somewhere different than toward the exit—and obstruct the pedestrian flow (speed and acceleration). They are focused on other people or things in the environment or become distracted via cell phones instead of focusing on the exit (attention to the exit).

The second category, just walking (2), is applied to people who are not pushing but also not as passive as the people in the falling behind (1) category; they are basically just going with the flow (Figure 1). People in this category have similar properties with the former category as they can have crossed and dropped arm positions, but since they are mostly within the crowd, they can use their arms close to their upper body to protect against possible pushing behaviors and they may hold onto fixed objects or barriers to stabilize themselves (arms and hands position). They move slowly and methodically, and they can form a penguin-like waddling motion (shoulder and head position). While they are mostly maintaining their position relative to the crowd and staying in their line, they can be in close body contact with others around them if they are jammed or shoved but under normal circumstances they have sufficient space around them to avoid body contact, as they do not actively increase or decrease the distance to others under a length of half a meter (personal space). They sometimes chat while they are walking (interaction with others). They are also slow and steady, and they may let others go first (speed and acceleration). They can focus on protection or the environment while they are walking toward the exit (attention to the exit).

Mild pushing (3) is a genuine pushing category but, as the name implies, a less active category than the fourth (Figure 1). People in this category actively increase the density of the crowd. They may raise their arms to apply force to the back of other persons or extend their elbows and arms, or even stabilize themselves by holding on to barriers to prevent others from overtaking (arms and hands position). They often move fast and methodically; consequently, they can form a "fast" penguin-like waddling motion (shoulder and head position). They have much more body contact, they tend to close gaps, change their lines, and overtake for faster access, but without applying excessive force. They may be disproportionately close to the next person without trying to overtake as a tailgating movement or as "psychological pushing," or the closeness can even occur out of an affiliation motive such as hugging someone they know (personal space). They mostly do not chat with other people (interaction with others). They are fast, and they actively decrease their distance to others (speed and acceleration). Their attention is focused on the exit or possible gaps providing a better route (attention to the exit).

The last category, strong pushing (4) is created due to the need for an advanced pushing category for extreme cases (Figure 1). People with strong pushing behavior tend to use their elbows and hands more strongly to create gaps, they can use barriers to pull themselves forward, they may collide with other people or even pull other pedestrians backward, as they are actively changing their position (arms and hands position). They can move sideways and use a shoulder as a plow, and in most cases, they lean forward (shoulders and head position). They have the most physical contact, and they may create some space behind them due to their rapid movement (personal space). They might communicate with others to engage in coordinated pushing (interaction with others). They are fast and accelerate rapidly when possible (speed and acceleration). Like the mild pushers in the former category, the strong pushers' attention is focused on the exit or possible gaps that might provide a better route (attention to the exit).

All actions in these categories are fully observable in overhead video analysis. This does not mean, however, that people show every parameter in their respective category as they move forward. A person does not necessarily use their arms close to their upper body as protection in just walking (2) if there is no pushing behavior around. There might be no coordinated pushing for people in the strong pushing (4) category if the strong pusher is alone. Consequently, people can be annotated and put in a category depending on their prominent behavior even if they do not meet all the parameters.

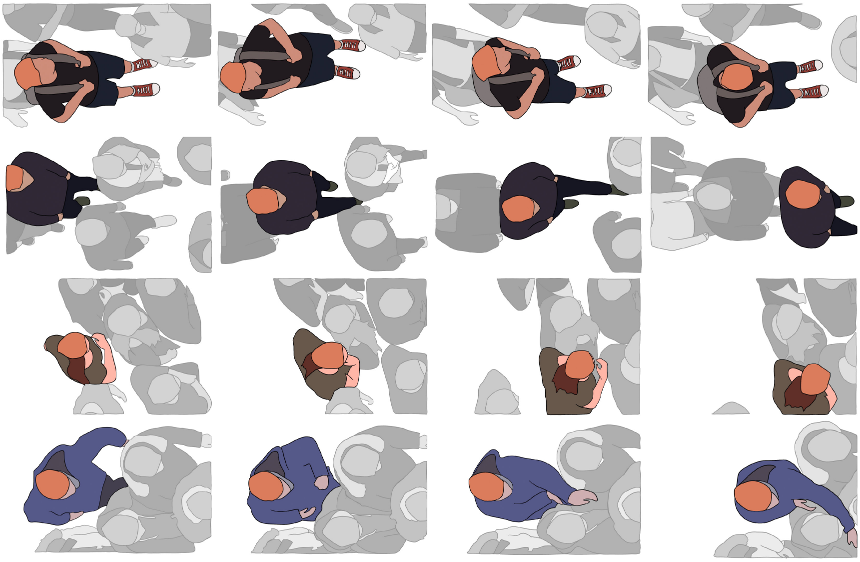

Figure 1: Illustrations of four categories. Each line represents one category. From top to bottom: Category (1) Falling Behind, Category (2) Just Walking, Category (3) Mild Pushing, Category (4) Strong Pushing

Another crucial point is that people are not bound to their initial category; as outlined above, they can change their behavior in real life and the category system adapts accordingly to account for these changes. A person might start out as just walking (2) but some time later switch to mild pushing (3) depending on the environment or a shift in motivation. This allows us to describe not only individual differences between people in the crowd but also to capture temporal dynamics.

2.3 Tentative application and revision of the coding system

Once the base structure and the technical properties of the pushing behavior system had been established, raters participated in a series of trials to develop the system further using existing datasets from the project BaSiGo [11, 23, 24] as well as interdisciplinary experiments performed at the University of Wuppertal [2, 25]. All the former experiment video recordings, along with trajectories of each pedestrian, had already been prepared for earlier research and studies and subsequently stored and published in the pedestrian dynamics data archive. The ethic statements for these experiments and recordings can be found in the corresponding papers; no additional ethical approval was necessary for the current study.

The selected empirical setup for the main trial video was an L-shaped bottleneck scenario, where all participants were instructed to reach the exit with high motivation [11, 23]. People were gathered on a platform, each wearing a unique hat (enabling their individual detection from cameras), and were instructed to pass through the bottleneck and exit the platform. Forty pedestrians were randomly selected (out of 123) for the trial dataset and rated accordingly.

The trial ratings revealed that understanding short-term behavior changes is notably challenging: Behavioral shifts of the pedestrians (i.e., category changes from 2 to 3) require more than a second to be comprehended by their actors since there were many examples of momentary behavioral changes for some pedestrians that appear to have happened only by accident (being pushed increases acceleration momentarily in a passive way) or to have been unconscious decisions on the part of the pedestrian (accidental line changing toward a gap), with the former behavior being resumed after one second. It was thus decided that the time gap for a valid and intentional behavioral change should be at least 2 or 3 seconds depending on its context.

2.4 Discussion of the validity of the coding system

Revision of the coding system after some trials revealed some significant points regarding the pushing behavior system. Raters were concerned that they were focused on the observable motivation (having high or low motivation) rather than actual pushing behavior in some cases. While being highly motivated and using high or mild pushing behavior are potentially highly correlated, the actual behavior can possibly be disregarded while observing the crowd due to the primed motivation of the pedestrian. This vague issue has come up during high-motivation-priming video trials where it was observed that, although most of the pedestrians were highly motivated to reach the bottleneck, not all of them were using pushing behavior. Overall, the main concern was that raters might inadvertently appraise the motivation of the pedestrians instead of their observable pushing behavior.

After careful consideration, raters agreed to conduct the rating process with a context-dependent perspective to avoid this issue. For instance, being fast and accelerated in a calm and slow crowd was agreed to be an indicator for mild (3) or strong pushing (4), but the same behavior can be seen as just walking (2) if the crowd is highly energetic and the average flow speed is similar to the "fast" pedestrian. It is thus helpful to watch the video once before the actual rating to get a feel for the respective context. Raters favored this approach as it is much more accurate for detecting and annotating pushing behavior, as it frames the question to be answered in more concrete terms.

The exactness of timing was also an issue for the consistency between both raters: After several test appraisals, some selected annotations done by two raters were analyzed and found, in fact, to be comparable except for a small time slippage by one or two seconds. It was later decided that the observed behaviors were actually the same but coded differently in time either by mistake or by a time lag caused by the software. Nevertheless, it is only natural for human observers to have minor errors in the timing of their ratings in a highly detailed and complex dataset, and those minor errors should not be problematic especially if the raters are in agreement about what they have seen. Consequently, raters decided to look more closely at the cases with a time slippage of up to two seconds between ratings and select the proper timing together for the main dataset. This process was called “correction” and was done for all the related cases.

2.5 Training of raters

The same L-shaped bottleneck video [11, 23] was selected for use again as the training dataset for two raters to annotate pedestrians. The remaining participants from the main trial dataset (n = 83, out of 123) were annotated by the raters. The rating was done via PeTrack and a txt output was collected afterward.

The output shows only the respective frames for which a rate comment was inserted (i.e., frame 0 = 3, frame 523 = 4, frame 801 = 3, frame 1792 = END; for participant *number*), hence it always needs to be prepared for data analysis. The first preparation was done manually; the total frame numbers were written in Excel and all the ratings were dragged in between the frames (i.e., frame 0,1,2,3. . . .520,521,522 = 3). After every rating for every frame and every pedestrian was prepared, the median of the ratings for units of seconds was calculated and written accordingly. The final procedure was to assemble all the ratings in one column. These proceedings were done separately for the two annotations of two raters. Later, data columns of two raters were merged (as two columns) and collected in one Excel file. The file was stored for later analyses in IBM SPSS.

2.6 Reliability analysis (inter-coder reliability)

It was decided that the inter-coder reliability should be calculated via Krippendorff’s alpha (KALPHA) [26]. Having multiple coders and an ordinal level of measurement (i.e., categories increase from 1 to 4 depending on the behavior), KALPHA was found to be the most effective reliability coefficient for our rating system.

For calculating KALPHA we used a macro by Andrew F. Hayes for IBM SPSS [27]. This macro provides a proper syntax where only the last line must be manually adapted to the respective data set and the required output. This looks as follows: KALPHA judges = judgelist/level = lev/detail = det/boot = z. “Judgelist” contains the names of the raters, “lev” is the measurement level (in our case: ordinal = 2), “det” is a selection of whether there is a need for a more detailed output (0 for only KALPHA value), and “z” is the bootstrapping number (in our case: 10000) [28]. As database for the inter-coder reliability, we used the ratings from the training section. So, N = 83 participants were rated by two independent raters. Please note that N = 43 participants were rated twice by one rater with 4.5 months in between because the first rating was performed before the method had been described in detail for this article. In the process of writing, the categories underwent additional differentiation and clarification, so we decided that both raters should conduct their observations at the same time. As the quality of the first rating thus might remain below what is possible, we repeated it for this paper to demonstrate more accurately the potential of our system. The second rating round was almost a new one since the rating process is very complex and there was a big time gap between the two ratings. The rater could thus not remember the former ratings and was of course not aware of the rating of the second observer during the process. Finally, the dataset for reliability analysis consisted of 143,172 rated frames. After aggregating 25 frames into one second, 5,717 units of analysis remained. We adjusted 60 units due small time slippages as explained in Section 2.4. For this prepared data set, the results show 90.5% overlap between the raters and KALPHA = .79.

Even though De Swert [28] mentioned KALPHA = .80 as an established limit for good reliability, he also stated that lower values (minimum of .67, or even .60 for extreme cases) are acceptable if there are good reasons for it. In our case, there is an extremely large number of analysis units, and our categories further rely on rather minor behaviors which are context dependent and sometimes difficult to detect from above. Additionally, behavioral shifts over time are considered, and the analysis units are somehow dependent from each other (e.g., if one observer sees a shift to mild pushing and therefore changes the rating from 2 to 3 but the other evaluates the behavior differently, the rating does not only differ for one second but immediately for several). Given this complexity of the rating system, a value of .79 is, in our view, more than satisfactory.

Despite this high level of agreement between raters, we nevertheless have partially divergent ratings for some participants. If the data is to be used for further analysis, however, there cannot be two data sets with divergent values. Therefore, the question is how to combine these different values into one value. The calculation of the mean value, for instance, makes no sense for the method (e.g., 2.5 as mean between just walking and mild pushing). Instead, the raters have to reach a later compromise in cases of disagreement. For that purpose, all divergent cases must be observed again and discussed. This leads to a completely consistent data set that can be used to answer the following research questions. It is essential to note that this step may only be performed after the inter-coder reliability has proven to be high enough.

2.7 Preliminary visualization

For visualization of the rating, we took one video from the Pedestrian Dynamics Data Archive [2, 21, 25]. Screenshots are depicted in Figure 2 and the full video can be found in the 'Supplementary files' section. This visualization is only preliminary to illustrate our rating system. More sophisticate forms can be created using special software (e.g., JuPedSim) [29] or including other quantities (e.g., density).

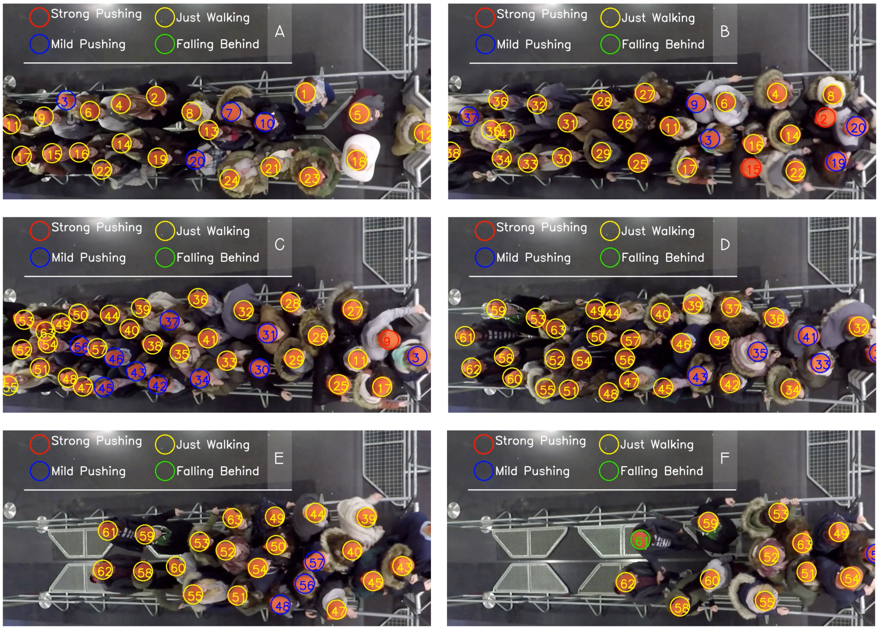

Figure 2: Preliminary visualization of ratings. Screenshots were taken from one exemplary video. Letters (A, B, C, D, E, F) state the order of the crowd flow. Timepoints of the screenshots are: A = 00.00 s, B = 00.08 s, C = 00.16 s, D = 00.24 s, E = 00.32 s, F = 00.40 s.

3 Discussion

Pushing behavior impairs people' sense of well-being in a crowd and also poses a significant safety risk. Nevertheless, to date it has been barely investigated. Following the idea that a crowd is not thoroughly homogenous in behavior and that there can also be changes over time, we developed a rating system of individual behavior in crowds. Prospectively, this can be used to systematize and quantify all kinds of forward motion as we not only capture pushing but also non-pushing behavior. However, since pushing can have various forms, having just a binary distinction would have been too easy. Therefore, we came up with four categories to take this diversity of forward motion into account: (1) falling behind, (2) just walking, (3) mild pushing, and (4) strong pushing. These categories thus enable us to classify the behavior of any person at any time in a video. In this way, we can not only consider the individuality of people but also the temporal dynamics of behavior. Our rating system was built on the scientific basis of content analysis [19, 20] and showed a very good inter-coder reliability between two trained raters.

3.1 Limitations

Although the rating system was found to be reliable, it is also worth mentioning its challenges and limitations in order to have a well-rounded perspective on the system. One major concern was noticed during the training process: The rating procedure was too time consuming. Annotating forward motions of numerous pedestrians involves repeated watching of the videos, focusing on a specific person, and determining the exact time periods of behavioral changes. Overall, annotating one pedestrian required at least five minutes of observation and consideration, as well as inserting the actual rates into the software. Complex cases, however, required as much as ten (or even fifteen) minutes. In order to have a complete annotation of 83 participants, each rater spent at least seven hours preparing the data. Raters spent an additional two hours correcting the data before the statistical analysis could occur (check Section 2.4). In the long run, these durations cannot and (more importantly) should not be decreased since the nature of the system depends on detailed observations. Speeding up the rating process might cause human observers to miss valuable information concerning the pedestrians.

The second observed issue was related to the properties of the selected video. Even though the video was high-resolution, image distortion (flattened fish-eye) sometimes made it hard to perceive and determine actual behavior. The software distorts images in this way to depict an accurate trajectory from the pedestrians from the first standing point through the bottleneck, but this also causes pedestrians to be shaped somewhat bizarrely when they move away from the center. The raters tried to adjust their observation and rate accordingly, although some information might have been lost throughout the process due to this situation. In a broader perspective, using only a bird's-eye view could potentially lead to a loss of information, as well, since the observation becomes slightly limited when seen only from this vantage point. Future studies could incorporate secondary cameras with frontal or side angles where it is thought that these could be beneficial.

Finally, the method was limited by the use of only one video for introducing the pushing behavior system. Even though the selected video contains a crowd scenario with varied behaviors, a different kind of environment (i.e., less crowded, high motivation, low motivation) could potentially be constructive for determining the applicability of the system itself. Raters have conducted some informal trials with different videos that suggest that the system is valid in all the cases mentioned. Additionally, investigating multiple exit scenarios or pedestrians moving in different directions could also be beneficial for showing how feasible the system is, although, we firmly believe that the system would be valid in these cases as well. If a crowd scenario contains forward motion of the pedestrians, then the system can potentially be used since it is based on individual observations regardless of the direction of the pedestrian moving. However, crowd contexts such as watching a sport or a music performance cannot be investigated with the current rating system because these situations do not contain forward motion. Nonetheless, regardless of the selected crowd scenario, it has proven beneficial for raters to confer in advance about the category system for each individual experiment and agree on a set of individual examples of the four categories. This minimizes the context effects.

3.2 Practical implications

While on the subject, possible future applications are described below. The first and probably the most prominent future study could be automating behavior detection by utilizing artificial intelligence (AI) [30, 31]. As it was mentioned in the limitations, the rating process is time consuming and laborious, but an automated AI system could dramatically decrease the rating time by assisting raters in appraising clear cases while flagging the ambiguous ones. All in all, the rating system and the actual annotations might be considered as the beginning of further pushing behavior-related studies since the system opens a door to measure behavior in space and time and potentially be applied to related research questions. If an automated detection system could be created, later research could use it to acquire the annotations of multiple videos in a short time.

Regarding future research in social and crowd psychology, behavioral effects can be easily observed and measured with the rating system. Observing one person or one group within a crowd is quite difficult due to having a massive amount of information from the environment, but reducing this data to four ordinal categories could be useful for observing what is really happening in the crowd. For instance, behavior propagation can be observed if it exists (i.e., strong pushing behavior propagates between pedestrians over time via exposure) or behavioral clustering can be identified in some specific locations (i.e., mild pushing behavior localizes in front of the bottleneck). The authors are currently working on these research questions in regard to the rating system's future application. These examples could potentially yield crucial insights for crowd management and evacuation studies, as well, since the system allows interested parties to understand pushing and pushing-related behaviors. Ultimately, the rating system should make it easy to recognize if behavior categories affect each other in any way, depending on the time and their position.

Although the rating method is far too time consuming to be directly useful in the application field of crowd management, it directs the focus toward observing individual behavior as a key to understanding the strategies people use in crowds. Such knowledge could be very useful for practitioners in the long run since (potentially dangerous) shifts in crowd movement could be better understood. Likewise, using the system can be beneficial in evacuation studies, such as observing the effects of given directions or instructions on the crowd at an individual level. Potentially, researchers can identify unfair or unwanted behaviors and their effects in an evacuation scenario, and then design or model alternate scenarios to avoid dangerous situations. Furthermore, the detailed descriptions of pushing behavior developed for this method could provide a starting point for thinking about automated observation tools for crowds to detect characteristic indicators of problematic behavior.

3.3 Conclusion

Our rating system provides an important and adequate basis for better understanding the complex dynamics of pushing behavior and forward motion in general. In the video we tested, the agreement between two raters was very good, and a consistent and highly reliable dataset can be generated through the subsequent strategy of compromising. In the future, however, the system must prove its suitability for other videos in different contexts (e.g., different motivations, different moving directions or even CCTV footage). An automated solution for speeding up the rating process would be also beneficial. In any case, this idea is worth pursuing since the quantification of pushing behavior is necessary to answer further research questions which will allow researchers to better understand crowds and thus contribute to public safety.

Acknowledgements

The authors acknowledge Prof. Dr. Armin Seyfried and Dr. Maik Boltes for providing the PeTrack software, and Deniz Kilic and Tobias Schrödter for updating the software with the necessary functions for category annotation. The authors thank Ahmed Alia for creating video files with ratings/annotations made by authors embedded on pedestrians. The authors thank Panar Ege Usten for creating the category illustrations.

References

[1] Victoria Filingeri, Ken Eason, Patrick Waterson, and Roger Haslam. “Factors influencing experience in crowds - The participant perspective”. In: *Applied Ergonomics* 59 (2017), pp. 431–441. DOI: 10.1016/j.apergo.2016.09.009.

[2] Juliane Adrian, Armin Seyfried, and Anna Sieben. “Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology”. In: *Journal of the Royal Society Interface* 17 (2020), p. 20190871. DOI: 10.1098/rsif.2019.0871.

[3] Milad Haghani, Majid Sarvi, and Zahra Shahhoseini. “When ‘push’ does not come to ‘shove’: Revisiting ‘faster is slower’ in collective egress of human crowds”. In: *Transportation Research Part A: Policy and Practice* 122 (2019), pp. 51–69. DOI: 10.1016/j.tra.2019.02.007.

[4] Norris R. Johnson. “Panic at “The Who Concert Stampede”: An Empirical Assessment”. In: *Social Problems* 34 (1987), pp. 362–373. DOI: 10.2307/800813.

[5] Chris Cocking, John Drury, and Steve Reicher. “The psychology of crowd behaviour in emergency evacuations: Results from two interview studies and implications for the Fire and Rescue Services”. In: *Irish Journal of Psychology* 30 (2009), pp. 59–73. DOI: 10.1080/03033910.2009.10446298.

[6] John Drury, Chris Cocking, and Steve Reicher. “Everyone for themselves? A comparative study of crowd solidarity among emergency survivors”. In: *British Journal of Social Psychology* 48 (2009), pp. 487–506. DOI: 10.1348/014466608X357893.

[7] Lee Clarke. “Panic: Myth or Reality?” In: *Contexts* 1 (2002), pp. 21–26. DOI: 10.1525/CTX.2002.1.3.21.

[8] Dirk Helbing, Illés J. Farkas, Péter Molnár, and Tamás Vicsek. "Simulation of pedestrian crowds in normal and evacuation situations". In: *Pedestrian and Evacuation Dynamics*. Ed. by M. Schreckenberg and S. D. Sharma. Berlin, Heidelberg: Springer, 2002, pp. 21–58.

[9] Colin M. Henein and Tony White. "Agent-Based Modelling of Forces in Crowds". In: *Multi-Agent and Multi-Agent-Based Simulation. MABS 2004. Lecture notes in Computer Science, vol 3415*. Ed. by P. Davidsson, B. Logan, and K. Takadama. Berlin, Heidelberg: Springer, 2005, pp. 173–184. DOI: 10.1007/978-3-540-32243-6_14.

[10] Sujeong Kim, Stephen J Guy, Karl Hillesland, Basim Zafar, Adnan Abdul-Aziz Gutub, and Dinesh Manocha. "Velocity-based modeling of physical interactions in dense crowds". In: *Vis Comput* 31 (2015), pp. 541–555. DOI: 10.1007/s00371-014-0946-1.

[11] Anna Sieben, Jette Schumann, and Armin Seyfried. "Collective phenomena in crowds—Where pedestrian dynamics need social psychology". In: *Plos One* 12 (2017), e0177328. DOI: 10.1371/journal.pone.0177328.

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

[13] Dirk Helbing, Illés Farkas, and Tamás Vicsek. "Simulating dynamical features of escape panic". In: *Nature* 407 (2000), pp. 487–490. DOI: 10.1038/35035023. arXiv: cond-mat/0009448.

[14] John Drury, Chris Cocking, Steve Reicher, Andy Burton, Damian Schofield, Andrew Hardwick, Danielle Graham, and Paul Langston. "Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model". In: *Behavior Research Methods* 41 (2009), pp. 957–970. DOI: 10.3758/BRM.41.3.957.

[15] Leon Mann. "Queue Culture: The Waiting Line as a Social System". In: *American Journal of Sociology* 75 (1969), pp. 340–354. DOI: 10.1086/224787.

[16] Bernd H Schmitt, Laurette Dube, and France Leclerc. "Intrusions Into Waiting Lines: Does the Queue Constitute a Social System?" In: *Journal of Personality and Social Psychology* 63 (1992), pp. 806–815. DOI: 10.1037/0022-3514.63.5.806.

[17] *Cambridge Dictionary — push*. n.d. URL: https://dictionary.cambridge.org/de/worterbuch/englisch/push (visited on 02/01/2022).

[18] Dirk Helbing and Pratik Mukerji. "Crowd disasters as systemic failures: Analysis of the love parade disaster". In: *EPJ Data Science* 1 (2012), pp. 1–40. DOI: 10.1140/epjds7.

[19] Kimberly A. Neuendorf. *The Content Analysis Guidebook*. 2nd ed. California: SAGE Publications, Inc, 2017. DOI: 10.4135/9781071802878.

[20] Nicola Döring and Jürgen Bortz. *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. 5th ed. Springer-Lehrbuch. Berlin, Heidelberg: Springer, 2016. DOI: 10.1007/978-3-642-41089-5.

[21] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Data archive of experimental data from studies about pedestrian dynamics [Data archive]*. n.d. URL: https://ped.fz-juelich.de/da/doku.php.

[22] Maik Boltes, Armin Seyfried, Bernhard Steffen, and Andreas Schadschneider. "Automatic Extraction of Pedestrian Trajectories from Video Recordings". In: *Pedestrian and Evacuation Dynamics 2008*. Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 43–54. DOI: 10.1007/978-3-642-04504-2_3.

[23] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Entrance 2, entry with guiding barriers (corridor setup) [Data set]*. 2013. DOI: 10.34735/ped.2013.1.

[24] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Entrance 1, entry without guiding barriers (semicircle setup) [Data set]*. 2013. DOI: 10.34735/ped.2013.2.

[25] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Crowds in front of bottlenecks from the perspective of physics and social psychology [Data set]*. 2018. DOI: 10.34735/ped.2018.1.

[26] Klaus Krippendorff. *Computing Krippendorff's Alpha-Reliability*. 2011. URL: https://repository.upenn.edu/asc_papers/43.

[27] Andrew F. Hayes and Klaus Krippendorff. "Answering the call for a standard reliability measure for coding data". In: *Communication Methods and Measures* 1 (2007), pp. 77–89. DOI: 10.1080/19312450709336664.

[28] Knut De Swert. *Calculating inter-coder reliability in media content analysis using Krippendorff's Alpha*. Center for Politics and Communication, 2012, pp. 1–15.

[29] Armel U. Kemloh Wagoum, Mohcine Chraibi, and Gregor Lämmel. "JuPedSim: An open framework for simulating and analyzing the dynamics of pedestrians". In: *3rd Conference of the Transportation Research Group of India*. 2015. URL: https://www.researchgate.net/publication/289377829_JuPedSim_an_open_framework_for_simulating_and_analyzing_the_dynamics_of_pedestrians.

[30] Ahmed Alia, Mohammed Maree, and Mohcine Chraibi. "A Hybrid Deep Learning and Visualization Framework for Pushing Behavior Detection in Pedestrian Dynamics". In: *Sensors* 22 (2022), p. 4040. DOI: 10.3390/s22114040.

[31] Ahmed Alia, Mohammed Maree, David Haensel, Mohcine Chraibi, Helena Lügering, Anna Sieben, and Ezel Üsten. "Two Methods for Detecting Pushing Behavior from Videos: A Psychological Rating System and a Deep Learning-based Crowd Behavior Analysis". In: *Proceedings of the 10th International Conference on Pedestrian and Evacuation Dynamics (PED2021), Paper No. 62*. 2021. URL: https://drive.google.com/file/d/1OdQIxcxaoCuQJUjQE4AC0NJ_qzC4mVHH/view.

PUBLICATION II

Psychological pushing propagation in crowds—Does the observation of pushing behavior promote further intentional pushing?

Note: This article was published as Lügering, H., Alia, A., & Sieben, A. (2023). Psychological pushing propagation in crowds—Does the observation of pushing behavior promote further intentional pushing? Frontiers in Social Psychology, 1, 1263953. https://doi.org/10.3389/frsps.2023.1263953

Author Contributions

Conceptualization: Helena Lügering
Data curation: Helena Lügering and Ahmed Alia
Formal analysis: Helena Lügering and Ahmed Alia
Investigation: Helena Lügering
Methodology: Helena Lügering
Project administration: Helena Lügering
Software: Ahmed Alia
Supervision: Anna Sieben
Validation: Helena Lügering
Visualization: Helena Lügering and Ahmed Alia
Writing–original draft preparation: Helena Lügering
Writing–review and editing: Helena Lügering, Ahmed Alia and Anna Sieben

Psychological pushing propagation in crowds—Does the observation of pushing behavior promote further intentional pushing?

Helena Lügering[1 2], Ahmed Alia[1 2 3] and Anna Sieben[4]

[1] Institute for Advanced Simulation 7: Civil Safety Research, Forschungszentrum Jülich, Jülich, Germany

[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany

[3] Department of Information Technology, An-Najah National University, Nablus, Palestine

[4] School of Humanities and Social Sciences, University of St. Gallen, St. Gallen, Switzerland

Abstract

When large numbers of people come together (e.g., at concerts or religious gatherings), critical situations can arise easily.While physical factors such as crowd density play a role, people's behavior can also affect crowd dynamics. For example, pushing and shoving, which are closely related to density, can quickly contribute to potentially dangerous dynamics. There is little extant research, however, on why people start pushing in the first place. Aside from individual reasons (e.g., motivation), social reasons might also play a role: an initial instance of pushing might be imitated or spark a competition if the pusher seems to reach the goal faster or the behavior of individuals defines a group norm whether pushing is allowed or not. Practically speaking, these social factors should lead people to push because they perceive other pushers, or, in other words, a psychological pushing propagation occurs. To address this question, the behavior (pushing or non-pushing) of people in 14 different experimental runs of crowds walking through a bottleneck ($N = 776$) was assessed by two independent raters with the help of a rating system of forward motion. This assessment was then linked to the spatiotemporal positions of the participants to combine it with the neighborhood relations. Based on that, it was analyzed whether individuals who started to push were more likely to be in the direct neighborhood of individuals who were already pushing. Results showed a small but significant effect suggesting that there is an influence, but that pushing is not overly "contagious".

Keywords: pedestrian dynamics; pushing propagation; pushing behavior; social psychology; crowds; social norms

1 Introduction

There are many different occasions for which people come together in small and large groups. Whether at concerts, religious gatherings, or train stations, crowded situations are everywhere. In such situations, however, people do not just behave in an orderly manner—by, for example, forming a queue. Sometimes they push and shove, and such behavior can have serious consequences: It has been shown, for instance, that pushing worsens the satisfaction of the people in the crowd [1]. Furthermore, various bottleneck studies have indicated that high motivation of the group, which is often accompanied by pushing behavior, significantly increases the density [2–4] Reports from real-life scenarios have shown that pushing also plays a role in the context of crowd accidents. However, people in critical situations usually do not push out of a sense of panic or selfishness to save themselves (for a critical discussion of the term “panic” in crowd accidents, see [5]). Instead, people from behind continue to push because they want, for example, to get faster access and do not realize how serious the situation has become in the front [6–8].

When talking about pushing, it is important to know that this behavior can be understood in different ways and to distinguish between intentional and unintentional acts [9]. In unintentional pushing, an external impact on the person’s body causes them to push against other individuals. In intentional pushing, the person applies energy themselves and thus builds up pressure on others. Unless stated otherwise, we refer to intentional pushing when talking about pushing in this paper. However, in addition to actively building up pressure from behind, we also consider other behaviors to be intentional pushing. In principle, we include any behavior that increases density in a crowd and can cause turbulence from behind, such as using gaps and being disproportionately close to the person in front. Thus, pushing behavior itself is very diverse. Moreover, crowds are usually not homogeneous with all people either pushing or not, but heterogeneous and dynamic, meaning that some people push, others do not, and the behavior changes over time [10]. These complex dynamics lead inevitably to the question of why people push in the first place.

Due to its impact on realistic crowd dynamics, various crowd models have attempted to integrate pushing behavior [11–15]. In these models, pushing is usually considered as an action which is performed with a certain probability when an individual wants to move in a specific direction (e.g., toward a goal) and others obstruct the way. Henein and White [12] added a second circumstance, namely that individuals push when someone else wants to move in their direction, i.e., they push to defend their position in the crowd. Sometimes people who push more are also characterized as being impatient, rude, or in a hurry [13, 14]. When thinking of (dense) crowds with a common goal, such as in bottleneck scenarios, however, there is in fact always someone in the way when moving forward. In addition, as long as no one gets hurt, pushing can also be fun. This has been shown anecdotally in our own experiments, in which the participants repeatedly expressed the joy of competing with each other. Furthermore, there are cultural phenomena at concerts which people initiate in order to push or to collide with each other, such as mosh pits, circle pits or the wall of death [16, 17]. Although these situations actually have a high risk of being injured [18, 19], those involved enjoy them and explicitly want to be pushed [16]. Yet the models largely disregard such aspects and generally lack a nuanced understanding about when and why some people push and others do not. But in fact, to the best of our knowledge, there are hardly any empirical studies that could provide these insights thus far.

One of the few psychological studies to yield initial evidence about pushing behavior is Drury et al. [20]. With a virtual reality approach, these researchers studied cooperative vs. competitive evacuation behavior in an underground environment. Competition was operationalized as the frequency with which others were pushed aside. They found out that there was more pushing behavior among low group identifiers (compared to high group identifiers) in mass emergencies. Additionally, the concern for others seemed to reduce the amount of pushing. In Adrian et al. [3], participants evaluated their perceptions and behavior after taking part in a bottleneck experiment. Among other things, the results showed that pushing forward and filling gaps were perceived as strategies for faster access. This view was also already indicated by a previous study [21]. Further, many said that they had pushed because it was their task to reach the goal first or more quickly. Other reasons for pushing were that the participants were pushed from behind, that they saw others pushing, or that they wanted to escape.

Starting from these answers and elaborating and systematizing them further, a number of reasons for pushing are conceivable, which can be roughly classified as "individual" and "social." According to this logic, individual reasons include everything that involves only the person itself, such as the demographics, the motivation to reach a goal or a personal preference for or against pushing. For instance, in Adrian et al. [3], significantly more men reported to have been engaged in pushing. However, since such individual factors are rather stable, but pushing is a highly heterogenous behavior that can potentially change several times within a short period of time, it is improbable that individuality alone can explain this phenomenon. Therefore, it is essential to consider social reasons, too, as these reasons take the interaction between several people into account. Several years ago, Zeitz et al. [22] reviewed the literature on crowd behavior at mass gatherings and concluded that undesired crowd behavior emerges when there is first a "seed," i.e., individuals or a small group who show this behavior initially, which then incites the crowd to join in. This idea could also be applied to pushing behavior, whereby different mechanisms are conceivable as to why someone engages in pushing behavior when others do so. For instance, one person might simply imitate the behavior of others or competition might occur because a pusher seems to reach the goal faster and this is perceived as unfair. In cases where members of a crowd not only see the pushing behavior but are also pushed directly, they might also push back, whether in retaliation or to defend their own position [similar to the assumption in the model of Henein and White [12]].

Moreover, social norms could be relevant in determining whether pushing is perceived as appropriate or allowed in a certain situation. Zeitz et al. [22] stated that engaging in abnormal crowd behavior requires the modification of existing norms—an idea that can be mainly traced back to the Emergent Norm Theory [23]. Originally, this theory hypothesized that new, non-traditional collective behavior emerges as a result of a crisis. However, even excluding exceptional precipitating events, there may not be clearly defined social norms for every situation (e.g., for different entrance scenarios). A queue, for instance, is a relatively well-defined social system in which cutting into the line or pushing is not acceptable. Instead, people get in line and wait unless they are allowed to overtake for a specific reason (e.g., queue-jumping in front of a ticket machine when one might otherwise miss the train) [24, 25]. If, on the other hand, the entrance is designed in a way that leads the crowd to form a semicircle and not a queue (i.e., an open space and not a corridor), the behavior perceived as appropriate changes and the situation is normatively more unclear. Evidence for this phenomenon is provided by the study of Adrian et al. [3] in which the corridor width was varied during a bottleneck experiment. The questionnaire data showed that participants engaged in pushing slightly less frequently and observed less unfair behavior (i.e., mainly pushing forward) in the narrowest corridor that most resembled a queue. In another study, participants watched pictures and videos (taken from above) of two entrance scenarios, a semicircle and a corridor setup [21]. Afterward, they had to answer different questions including the inappropriate behavior observed and the social norms that apply in each scenario. Concerning forms of inappropriate behavior, participants mentioned pushing, shoving, and jostling for both setups. However, less inappropriate behavior was observed in the corridor compared to the semicircle setup. So, even if pushing is generally considered inappropriate, participants demonstrated this behavior—especially in experimental setups that lead to a semicircle formation. This pattern was also reflected in the norms mentioned. For the corridor, reminiscent of a queue, the three most frequent answers were "norm of queuing / lining up," "orderly behavior," and "pushing and shoving are forbidden." Whereas for the semicircle it was "the strongest wins / right of the stronger," "no rules," and "first come, first served" ([21], p. 14). Thus, the prevailing norms in the latter setup do not clearly indicate whether pushing is permitted or prohibited.

So, all in all, the social norms that apply in an entrance situation, especially concerning pushing behavior, are not generally clear but seem to be partly influenced by the spatial structure. Thus, participants' observations of pushing behavior could resolve this ambiguous social situation and define a group norm that pushing is actually allowed. Regardless of the exact mechanism, though, all these social reasons cause individuals to start pushing because they are in contact with people who are already pushing, or in other words, a psychological propagation of pushing behavior occurs. We call this form of propagation "psychological," because we are only dealing with the intentional onset of pushing. The purely physical reaction of the body when being pushed and therefore perhaps bumping into another person is not part of this research. Instead, we are interested in voluntary behavior.

The hypothesis of possible pushing propagation is already supported by both the results of Adrian et al. [3]—pushing because one is pushed from behind or sees others pushing— and our own informal observations. During these unsystematic observations, we noticed that people who had contact with strong pushers intensified their behavior afterward. This does not, of course, apply to everyone. In Adrian et al. [3] there were people who did not push at all. Frequently mentioned reasons for this in the questionnaire included a general rejection of pushing behavior, consideration for others, avoidance of danger, and the belief that pushing is inefficient. In such cases, propagation would probably not occur. Depending on the underlying mechanism, pushing propagation might also turn out differently under different circumstances. If, for example, the formation of a social norm is decisive, it is possible that the effect only occurs when several people in the vicinity are pushing, or that a certain proportion of pushers relative to the non-pushers must be reached. Furthermore, it may not only be a short-term propagation effect, but also related to crowd members' previous experiences in this situation. In other words, people who have had many contacts with pushers in the past may themselves start pushing at some point, even if the last contact was some time ago. Here again, the ratio between pushers and non-pushers could be decisive. Given the many potential complicating factors, this issue is very complex and requires systematic investigation. We therefore conducted a study using existing empirical material to examine whether and under what circumstances the behavior of neighbors influences whether a person starts pushing or not. Our main hypothesis was that there is a general influence. However, as this research approach is completely novel, most of the analysis was exploratory.

2 Method

2.1 Video material and rating of pushing behavior

To answer our research question, we used data from a previous bottleneck experiment conducted in Wuppertal, Germany, in 2018. The videos and trajectories including a short description of the experiment can be accessed via the Pedestrian Dynamics Data Archive [26]. For a more detailed description of the method and relevant ethical issues see Adrian et al. [3]. From all participants who took part in the experiment written and informed consent was obtained. In the experiment, there was no explicit instruction on how the participants should pass the bottleneck (e.g., being fast, being highly motivated, moving normally) but a more complex story was used. The participants were told to imagine that they were on their way to a concert by their favorite artist. In the "high motivation" condition the view was restricted from the back and participants sought to access the concert as quickly as possible to stand near the stage, whereas in the "low motivation" condition the view was good everywhere. Due to this implicit manipulation, we expected higher diversity of forward motion within the crowd as well as over time. This assumption was well supported by the results of the subsequent survey indicating that some participants pushed while others did not. This heterogeneity in behavior made the data set perfectly suited for our new research purpose.

For this secondary analysis, we chose all the experimental runs with more than 40 participants since the runs with fewer participants were too short to expect any meaningful social interactions. This selection resulted in 14 videos (out of 22)—seven low motivated and seven high motivated ones. For one pair of a high and a low motivation run, the group of participants was the same and consisted of $N = 42$ to $N = 75$ people. The videos lasted on average 48 sec for the high motivation condition and 53 sec for low motivation condition. The physical quantities of each run were analyzed by Adrian et al. [3]. To quantify the behavior of every participant over time, we used the rating system of forward motion established in Üsten et al.

[10]. With this system, the behavior of each person at each moment in the video was assigned to one of four categories: (1) falling behind (e.g., being slower than the group), (2) just walking (e.g., going with the flow at an appropriate pace), (3) mild pushing (e.g., being disproportionately close or overtaking by using gaps), and (4) strong pushing (e.g., pulling people backwards or using a shoulder as a plow). In order to link the assessment of the behavior with the spatiotemporal position of the participants, the annotation function of the software Petrack [27] was used. In this way, all participants received a rating for each frame—which was later aggregated to seconds using the median. A detailed description on how the method works and exactly which behaviors the individual categories include can be found in Üsten et al. [10].

To ensure high quality of the rating, all the videos were evaluated by two independent raters. Please note that although the raters were aware of the pushing propagation hypothesis, due to its complexity, an intentionally hypothesis-compliant rating would not have been possible.Within the process, each person is looked at individually and independently of their neighbors, and the order of persons is determined by the random assignment of IDs by Petrack. Thus, when rating one person, the raters were not aware of what rating a neighboring person might have already received. In total, the forward motion of $N = 776$ participants was rated. The resulting data set had an overlap of 75.6% and a Krippendorff's alpha of 0.65. Given the complex nature of the chosen experimental videos (e.g., densities up to 9 people per m^2 [3]) which increases the difficulties of the rating system discussed in Üsten et al. [10] (e.g., assessing small changes in behavior that are difficult to detect from above), this inter-coder reliability was considered acceptable [28]. Following the system's guidelines, the raters then revisited any cases of disagreement and compromised on a rating for each person in question. In this way, a unified data set was created. Since the original system consists of four categories, but for this question we are only interested in whether people push or not, the categories (1) falling behind and (2) just walking, as well as (3) mild pushing and (4) strong pushing, were subsequently combined. Thus, we had only two categories in our data set: (1) non-pushing and (2) pushing. Whether people are pushing or not is probably also the most relevant qualitative difference for crowd safety. Of course, it may get even more dangerous when pushing is stronger, however, as a first step, it seems more relevant whether people are pushing at all.

2.2 Identification of neighborhood relations

The videos of the analyzed experimental runs were recorded with 25 frames per second. For each time frame, we identified who is a neighbor with whom by a Python script. The identification was based on Voronoi cells [29], with two pedestrians being neighbors if their two individual cells share one border. However, we defined one restriction: The persons must not be further than 80 cm from each other since, in our view, a greater distance makes a direct influence less plausible (see Figure 1). The exact threshold of 80 cm was chosen because, in a test sample, it produced almost the same results with automatic neighbor detection as with manual detection. At a smaller distance, participants were not identified as neighbors, although they were evaluated as such by a human rater, and a larger distance resulted in the opposite issue. Further, similar to the rating of pushing behavior, once a person reached the bottleneck, they were no longer considered. The moment that there were only two participants or less in the experimental area, the identification process was terminated. That way, we obtained a list of neighboring persons for every single participant (i.e., target person) for every frame. But since one second was chosen as a minimum unit for the rating system [10] and this is, in fact, more appropriate for human behavior than a frame-based analysis, the neighborhood relations were also condensed to one second. For this purpose, the state in the middle frame was chosen as reference for the respective second (e.g., frame 12 for second 1, frame 37 for second 2...).

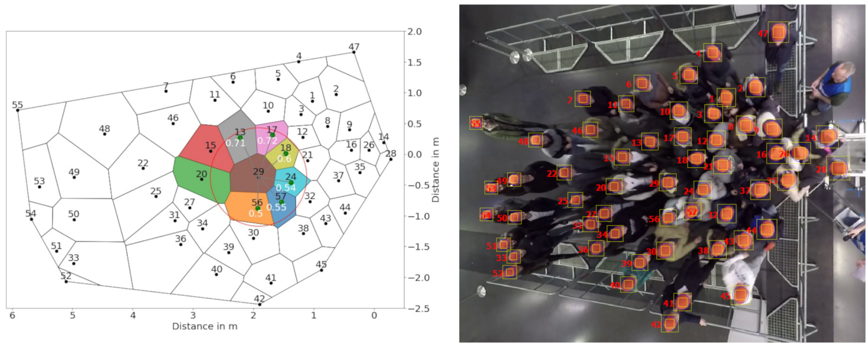

Figure 1: Exemplary neighborhood relations of target person (ID 29) based on the Voronoi technique. Participants with IDs 13, 17, 18, 24, 57, and 56 were considered neighbors of the target (ID 29) in this frame. Although participants with IDs 15 and 20 shared a border of a Voronoi cell with the target, they were not neighbors by definition because their distance was greater than 80 cm (red circle).

2.3 Combination of neighborhood relations and pushing rating

After combining the neighborhood relations and the assessment of the pushing behavior, we obtained a data set including every target person at every second plus the following information: (1) the condition (1 = "low motivation", 2 = "high motivation"), (2) the second when the bottleneck is reached, (3) the behavior (1 = "non-pushing", 2 = "pushing") of the target person in this second, (4) all neighbors of the target person in this second, (5) the behavior of the neighbors (1 = "non-pushing", 2 = "pushing") in this second, (6) the absolute number of neighbors pushing in this second, (7) four different pushing scores (see next section), and (8) whether the target person will start pushing in the next second (0 = "no", 1 = "yes").

Different pushing scores

In order to examine whether the behavior of the neighbors can predict whether a target person will start pushing or not, four different pushing scores indicating the number of pushing neighbors were defined. Each score took different aspects into account and was calculated for every second in which the participants were in the experimental area regardless of their own pushing state.

- *Pushing score over the last three seconds (PS_3sec).* With the help of this score, we wanted to investigate whether there is a short-term pushing propagation. Since it is not realistic that the behavior of the neighbors influences a person from one second to the other, the behavior of all persons who were in the vicinity of the target person in the respective second plus two seconds before was included. Each pusher in the neighborhood scored one point. These points were added up over the three seconds. This also means that a single pusher neighboring the target person in all three seconds scored three points. Please note that for the first two seconds of each person, the pushing score obviously could not cover three seconds. However, in order to have a complete data set, we decided to still calculate a score by including only the first or the first two seconds.

- *Pushing score over the last three seconds relative to the total number of neighbors (PS_3sec_neigh).* It is not necessarily the absolute number of pushers in the neighborhood that is decisive, but the ratio between pushers to non-pushers. Therefore, the PS_3sec was divided by the total number of neighbors in the respective seconds to obtain the proportion of pushers in all neighbors. To make the results easier to interpret, the score is not given as a decimal number but as a percentage.

- *Cumulative pushing score (PS_total).* This score considered the "experiences" that the participants had gathered so far during the experimental run. Similar to the PS_3 sec, each pushing neighbor scored one point every second. This time, however, the points were not only added up over the last few seconds but over the entire period since the start of the run.

- *Cumulative pushing score relative to the total number of neighbors (PS_total_neigh).* Again, it might be the ratio between pushers and non-pushers that is relevant. So, the PS_total was divided by the total number of neighbors the target person had up to that time. Like the PS_3sec_neigh, this score is given as a percent.

2.4 Statistical analysis

While the data preparation was done with Python, all statistical analysis was either performed using IBM SPSS Statistics 29 or RStudio v2023.03.0. To first get an impression of the frequency of pushing in the experimental runs, we calculated the mean proportion of participants who push at least once during a run and the mean duration of individual pushing phases. Since we expected very different amounts of pushing behavior in high and low motivation videos, they were considered separately. This approach was also supported by the fact that the pushing rating is to some extent context-dependent and a pusher in a low motivation video is not necessarily assessed as a pusher in a high motivation video (see [10]). The distinction was maintained for all further analysis.

Next, we looked more closely at the moments in which participants changed from a non-pushing status to a pushing state, or in other words, the moments in which they started pushing. In order to investigate whether the behavior of the neighbors shortly before a change had had an influence, all people who were in the vicinity of a target person in the prior three seconds were identified. Additionally, information on the behavior of the neighboring persons themselves in the respective second (i.e., non-pushing or pushing) was provided. All pushing neighbors were then summed up for each target person and each change (as some participants changed their behavior several times during one run) separately. Please note that in contrast to the pushing scores, every neighbor only counted once, even if they were around for several seconds and also pushed for several seconds. This procedure served to get a first idea on how many different pushers were actually in the vicinity shortly before a change in behavior. However, since we expected a high base rate of pushing behavior, especially in the high motivation videos, individuals probably often had one or more pushing neighbors, regardless of their own behavior. Therefore, a descriptively high percentage of changes involving pushing neighbors would be neither astonishing nor informative. We thus calculated the expected number of changes with pushing involvement using the overall percentage of time that participants had at least one pushing neighbor (regardless of their own behavior). We then used a Chi-square goodness-of-fit test to compare whether the observed number differed significantly from the expected number. Finally, repeated measurement logistic regressions (using the “glmer” function in the “lme4” package) were calculated to determine whether the start of pushing behavior in the next second can be predicted based on the four different pushing scores. The significance level was set at $p < 0.05$ for all statistical tests.

Transition phase

The statistical method of logistic regression assumes implicitly that there is a hard cut-off between non-pushing and pushing behavior, meaning that participants are classified as non-pushing one second, while they are pushing the next. Since it is not realistic for people to suddenly start pushing without mental preparation, this assumption might distort our statistics. To deal with that issue, we already considered not only the neighborly influences from one second to the next but included all neighbors in the three seconds before a change for our analysis. Likewise, it did not make much sense to count a target person two and three seconds before a change as a non-pushing case if the neighbors in these seconds were already perceived as possibly influencing a pushing start. This period of two seconds was accordingly treated as a transition phase and excluded from the analysis.

3 Results

3.1 Descriptive analysis of pushing propagation

In the high motivation videos, 82.2% of the participants pushed at least once during a run and, on average, each person stayed 11.35 sec in this state of pushing behavior. For the low motivation runs, these numbers were expectably lower. Nevertheless, 38.9% of the participants pushed at least once and they stayed in the pushing state for a mean of 8.28 sec. For a visualization of these changing dynamics in pushing behavior see the example in the Supplementary Video 1. Looking more closely at the moments of change, we had a total of 210 (high motivation) and 106 (low motivation) cases in which a participant moved from a non-pushing to a pushing state. When examining the behavior of the target person's neighbors three seconds before each change, it turned out that, in 91.4% (192 cases, high motivation) and 72.6% (77 cases, low motivation), there was at least one pushing neighbor in the vicinity. In many cases, however, there was even more than one—in runs with high motivation, the average number of pushing neighbors was 3.1 (range 1–10), while in runs with low motivation it was 2.1 (range 1–8).

Overall, and regardless of their own behavior, participants had at least one pushing neighbor in 85.7% (high motivation) or 51.7% (low motivation) of the time. Based on these percentages and the total number of actual cases in which a participant moved from a non-pushing to a pushing state, we would expect 180 (high motivation) and 55 (low motivation) changes to happen under pushing involvement. On a descriptive level, these numbers are lower than the observed ones. Calculating a Chi-square goodness-of- fit test, it turned out that the expected frequency also differed significantly from the observed ones for both conditions (high motivation: $\chi^2 = 5.60$, $df = 1$; $p = 0.018$; low motivation: $\chi^2 = 18.29$, $df = 1$; $p < 0.001$).

		N	*M*	*SD*	**Repeated Measurement Logistic Regressions**				
					AIC	*OR*	*CI*	*p*	*Marginal R^2*
PS_3sec	0	4175	4.23	4.12	1581.2	1.14	1.09 – 1.19	<0.001	0.056
	1	210	6.46	4.33					
PS_3sec_neigh	0	4175	33.88	29.53	1592.2	1.02	1.01 – 1.02	<0.001	0.047
	1	210	48.67	27.95					
PS_total	0	4175	26.14	28.47	1616.3	1.01	1.00 – 1.01	0.123	0.004
	1	210	28.01	29.53					
PS_total_neigh	0	4175	34.84	24.80	1605.2	1.02	1.01 – 1.02	<0.001	0.029
	1	210	45.05	24.14					

Table 1: Descriptive statistics and results of repeated measurement logistic regressions for high motivation runs.

3.2 Prediction of target persons' behavior based on neighbors' behavior

Next, we used repeated measurement logistic regressions to examine if the different pushing scores significantly predicted the target persons' behavior in the next second. The results of all models can be found in Table 1 for high motivation and in Table 2 for low motivation runs. Dependent on whether the target person starts to push in the next second (= 1) or not (= 0), sample size, mean, and standard deviation are indicated separately for each pushing score. Further, the Akaike Information Criterion (*AIC*), the Odds Ratio (*OR*) with confidence interval (*CI*), the *p*-value as well as the *marginal* R^2 [30] for each model are listed. An OR greater than 1 indicates that the odds of starting to push increases. The OR of 1.14 for the PS_3sec in the high motivation condition means, for example, that for each unit increase in this pushing score the odds of starting to push increases by 14%. The other ORs can be interpreted accordingly. The relation between the pushing score in one second and the predicted probability of a target person moving from a non-pushing to a pushing state in the next second are also depicted in Figure 2 (high motivation) and Figure 3 (low motivation) separately for each score. All in all, the logistic regressions revealed small but significant predictions for all models except for the PS_total in high motivation runs.

		N	*M*	*SD*	**Repeated Measurement Logistic Regressions**				
					AIC	*OR*	*CI*	*p*	*Marginal R^2*
PS_3sec	0	7889	1.66	2.46	1078.1	1.25	1.17 – 1.34	<0.001	0.061
	1	106	3.56	3.77					
PS_3sec_neigh	0	7889	13.11	18.41	1091.1	1.03	1.02 – 1.04	<0.001	0.045
	1	106	24.67	25.09					
PS_total	0	7889	7.98	13.23	1086.1	1.04	1.02 – 1.05	<0.001	0.042
	1	106	16.58	20.80					
PS_total_neigh	0	7889	10.94	15.28	1094.3	1.03	1.02 – 1.04	<0.001	0.044
	1	106	19.87	20.02					

Table 2: Descriptive statistics and results of repeated measurement logistic regressions for low motivation runs.

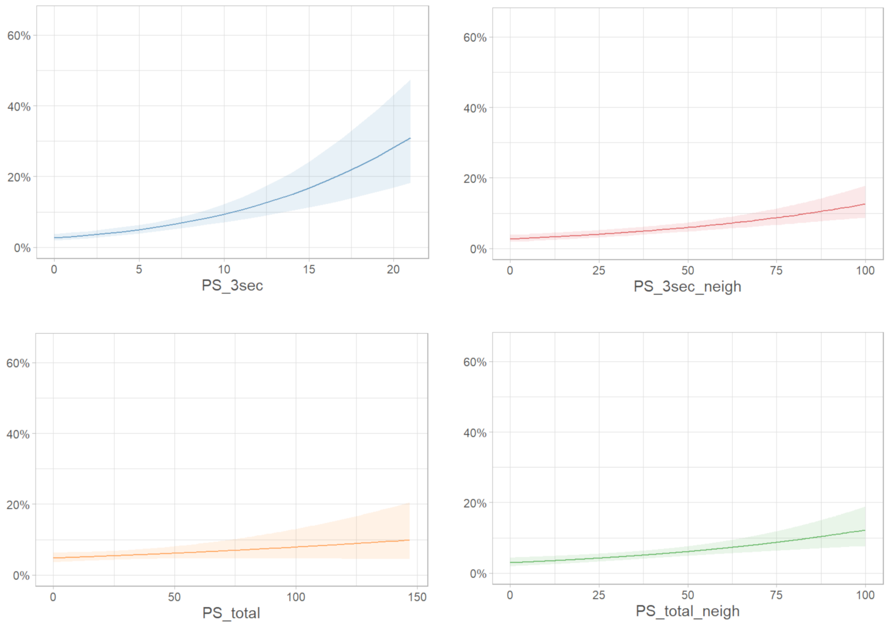

Figure 2: Predicted probabilities for start of pushing behavior in the next second in high motivation runs.

4 Discussion

The aim of this study was to show that the behavior of neighboring persons can influence whether someone starts to push or not. In addition, we wanted to examine whether the number of pushing neighbors at one moment or summed up for the entire situation is more important, and whether it is the absolute or relative number of pushers that is more decisive. All in all, our main hypothesis was confirmed as the analysis yielded small but convincing evidence of a psychological pushing propagation. The observed frequency of changes to a pushing state involving pushing neighbors was significantly higher than the expected one and an increase in a pushing score corresponded to an increase in the odds of starting to push for almost all scores. Only the PS_total does not provide a significant prediction in the high motivation runs, although it does in the low motivation runs. This somewhat surprising exception may be explained by the fact that, in high motivation runs, (strong) pushing is often accompanied by a faster reaching of the bottleneck. Participants who pushed at least once during these videos left the experimental area on average after 24 seconds. For those who did not push at all, the mean was 33 seconds. However, the PS_total "rewards" a longer stay in the experiment, since individuals who generally have more neighbors should also tend to have more contact with pushing neighbors and thus a higher score. As a result, participants who never push might have a higher PS_total than participants who push once in a while just because they have more time to collect points. In our analysis, this distortion is

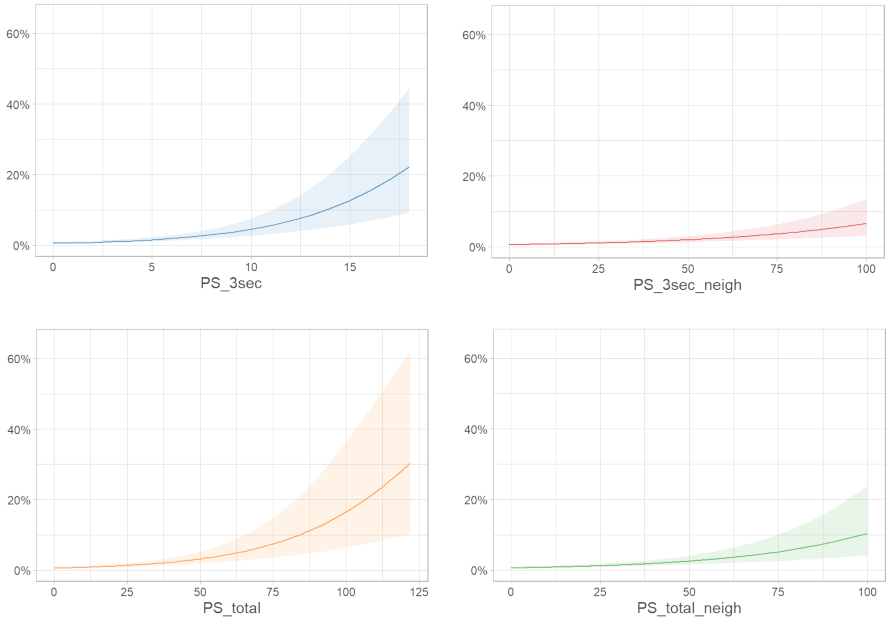

Figure 3: Predicted probabilities for start of pushing behavior in the next second in low motivation runs.

especially crucial when people who start pushing do so at a very early stage. Because then they may have a rather low PS_total, although there has already been comparatively much contact with pushers. This explanation is supported by the fact that the PS_total_neigh, which is relativized at the absolute number of neighbors, again provided a significant prediction. In the low motivation condition, by contrast, this problem was not quite as pronounced, since here pushing behavior is mostly expressed as being close, applying force from behind or, at most, using gaps. Pushing people to the side and thus reaching the bottleneck much faster hardly ever happened. Therefore, participants who never pushed stayed in the video for a similar amount of time ($M = 29$ seconds) as those who did at least once ($M = 26$ seconds). Presumably, this is why the score provided a better prediction for the low motivation runs, however, due to the susceptibility to the temporal component, it does not seem to be a reliable measure for the exposure to pushing behavior anyway.

Looking at the AIC, which is a measure of model fit, of each model in general, we found that actually neither the PS_total nor the PS_total_neigh nor the PS_3sec_neigh were good fits. In both motivation conditions, the model of the PS_3sec has the lowest AIC, indicating the comparatively best fit. Since according to Burnham and Anderson [31] an increase of two units in the AIC already implies a significantly worse model, the performance of all other pushing scores was clearly lower. The PS_3sec is also conceptually the most straightforward value as it only considers pushing neighbors within the last three seconds. The relative proportion of pushers compared to the total number of neighbors thus seemed less relevant to the question of pushing propagation, as does the total pushing exposure throughout the whole experiment. This indicated that neighbors' behavior has a rather short-term effect on a pushing start.

Unfortunately, our study was not able to shed any light on the exact (social) psychological mechanism behind this effect. In particular, we could not conclusively determine whether the unclear social situation regarding the eligibility of pushing behavior was resolved by the emergence of a social norm. We assumed that this emergence might depend not on the absolute but on the relative number of pushers in the vicinity. In our study, however, both pushing scores that took the relative proportion into account (i.e., PS_3sec_neigh and PS_total_neigh) performed worse than the one that only depicts the absolute number (i.e., PS_3sec). Nevertheless, many participants who changed to a pushing state had not only one but several pushing neighbors in that situation. In addition, our results generally showed a "the more, the higher"-effect in terms of the relationship between pushing neighbors and the odds of starting to push. Hence, the social norm could also emerge as a result of several individuals engaging in pushing behavior, no matter how many others do not. In any case, further research is needed to understand the underlying mechanisms.

Despite all the significant results, however, it is also important to keep in mind that the effect is quite small. Increasing the PS_3sec by one point (i.e., one pushing neighbor for one second more), increased the odds for starting pushing by 14% in high motivation runs and 25% in low motivation runs. This means that, even with strong pushing exposure, the probability of starting to push in the next second is only about 30%. Furthermore, participants had high pushing scores even in situations in which they did not start pushing (e.g., for PS_3sec in high motivation runs: $M = 4.23$, $SD = 4.12$). We can thus conclude that pushing neighbors do not necessarily lead to intentional pushing. This is also a plausible and realistic finding because, contrary to Le Bon's assumptions [32], not everything in crowds is contagious. Even though apart from him, there are several other early theories on collective behavior that "accept the unanimity of the crowd as its most salient feature" ([33], p. 21), the idea of crowds being completely homogenous was already critically questioned at that time [33]. A more recent empirical study similarly showed that only a fraction of people engage in violent interactions at protest events or football matches [34]. Systematic observations showed that at no time did more than 10% of individuals belonging to the same group (e.g., home fans) actively engage in violence in the same situation. In our study, 82.2% pushed at least once in the high motivation videos and 38.9% in the low motivation videos. These numbers are of course considerably higher than those found in the study of Adang [34], but violent behavior is also more extreme than pushing behavior. Overall, however, it can be concluded that not everyone in a crowd engages in a certain type of behavior. This also corresponds to the findings of Adrian et al. [3], according to which some participants simply do not want to push. When comparing the results of our secondary analysis with the original study, though, please note that the percentage of people who were rated as pushing at least once in our study was significantly higher than the percentage who self-reported pushing [3]. The self-report data only exists for the high motivation runs but here the proportion varies between 29.2 and 78.6% per run. However, this difference may simply be due to the fact that our rating was stricter and also included single, short pushing events that the participants may not have remembered when answering the questionnaire.

4.1 Further research

To the best of our knowledge, our study is the first to deal with psychological pushing propagation and to attempt to quantify its impact on crowd dynamics. Therefore, there are many open research questions to further pursue on this topic. For one thing, experiments could be conducted in which some individuals are specifically instructed to push, while all others are not given any specific instruction or are simply asked to behave as usual. In this way, it would be easier to explore the behavior of the neighbors in response to the pusher and to investigate pushing propagation more systematically. In such experiments, other influencing factors could also be examined to make the prediction more accurate. For example, the distance between the person pushing and the others or the spatial positioning (i.e., if the pusher is more in front of or behind someone) might be crucial for the strength of propagation. There may even be an interaction between the distance and the spatial positioning insofar that a propagation from behind only happens when the distance is very small, while the distance can be larger when the pusher is in the field of view. Moreover, the interaction with individual factors could be investigated. For example, certain personality traits may increase the susceptibility to propagation (e.g., less agreeable individuals may be more likely to push themselves after being pushed than highly agreeable individuals). The gender constellation of pusher and pushed person could change the dynamics, too, since at least in mosh pits, some men refuse to push women or push them less out of fear of injuring them [16]. Following the study of Drury et al. [20], it is also possible that social identity moderates the effect of pushing propagation. Further, as pushing behavior tends to increase when approaching the bottleneck [35], pushing propagation might become more likely as well. There might be even a temporal influence, since highly motivated individuals might position themselves close to the bottleneck right at the beginning to be among the first, and at the same time these individuals might be particularly prone to pushing propagation. In general, the crowd context could be relevant as well. Our experiment was conducted with a concert framing where dancing, jumping and shoving is rather typical, whereas a conference (i.e., a more orderly setting) or a department store sale (i.e., an even more competitive setting) might evoke quite different dynamics. These examples already point out how intertwined individual, social and other situational factors are and that carefully planned experiments are needed to separate these theoretically distinct concepts from each other and better understand their influence. To additionally explore the exact mechanisms behind the effect of pushing propagation (e.g., imitation or competition), participants could be asked after the experiment to list reasons why they did or did not push when they saw others pushing. While they may not be conscious of some of the mechanisms, a closer inquiry could provide important insights.

An entirely different research idea would be to ask whether observing pushing behavior leads to changes in attitude toward that behavior rather than encouraging pushing per se. As Turner and Killian [33] stated, it is possible for individuals who do not engage in a behavior to nevertheless change their attitude or feelings about it after seeing it in others. Adang [34] reported something similar. In his study, a much larger proportion was indirectly involved in crowd violence with up to 80% supporting the behavior verbally. On the other hand, it is conceivable that witnessing people pushing might not lead to pushing propagation but in fact have an opposite reactance effect as people are reluctant to join in and even slow down as a result. In the original rating system, this behavior would correspond to the category "falling behind." With the data set used in this study, however, the investigation of such an effect was rather impossible. Not only because we merged the categories "falling behind" and "just walking" into one "non-pushing" category, but also because the crowds studied were often so dense that it was hardly possible to fall behind. This behavior was observed more at the edges of the crowd, and, even if there were some participants in the middle who gave up, it would have been difficult to detect this in the videos taken from above. There were in fact a few extreme cases in which participants withdrew from of the crowd. However, these were too rare to systematically investigate a possible reactance effect.

4.2 Limitations

While our study does provide new and potentially important insights into the initial stages of pushing behavior, the methodology was admittedly highly simplified in many respects. First of all, we only examined bottleneck scenarios and no other crowd constellations. These setups are, however, the situations in which pushing and pushing propagation are most likely to occur, as everyone is trying to reach a common goal with limited accessibility. Furthermore, the analysis in this study only included the number of pushing neighbors and, to some extent, the length of the contact, but it neglected other factors that could improve the prediction of propagation, such as the distance between the persons. And even though the length of contact is at least partially accounted for in the two pushing scores that sum over the entire experiment (i.e., PS_total and PS_total_neigh), another limitation becomes apparent here. In the pushing scores, one second of pushing and one new pushing neighbor were weighted equally, meaning that three points could result either from one pusher being around for three seconds or three distinct pushing neighbors being there for one second each. For one thing, it is questionable whether this equation is conceptually appropriate. For another, it is not possible to distinguish whether a high pushing score is due to a target person's having had prolonged contact with a single pushing person or whether there was short-term contact with many pushers. Further, we aggregated the two categories "mild pushing" and "strong pushing" to just one pushing category which is why it was impossible to distinguish these different levels of pushing in our analysis. However, it is conceivable that simply being disproportionately close to the next person has a different impact on propagation than a strong push from behind. Moreover, the rating system of forward motion only allowed us to determine whether participants were pushing in general but not whom they were pushing specifically. As a result, we could not account for whether a target person was directly affected by a pushing neighbor or whether someone else nearby was actually being pushed. However, whether or not someone is being pushed directly could make a significant difference for their own behavior (due to retaliation or defense of the position). Finally, as no personal ID codes were used in this experiment, the questionnaire data as well as the demographics could not be reassigned to the participants in the videos which is why we could not include this information in our analysis. Overall, many of the limitations are due to the fact that this was a first attempt to study a psychological pushing propagation using secondary analysis of previous experiments that were not originally conducted for this purpose. Therefore, building on the results obtained here, new experiments should be designed to examine the effect in a more systematic and complex manner.

5 Conclusion

All in all, our research showed that psychological pushing propagation exists but that the effect is small. This is an important insight when it comes to the question as to why people start to push, and it could be useful for modeling crowd dynamics. Further research is essential to develop a deeper understanding of the underlying mechanisms and other factors bear consideration, for they might be equally important for the occurrence of pushing behavior. Pushing is multicausal and we should abandon the idea that everything in a crowd is highly contagious. People do influence each other, but only to a limited extent.

Data availability statement

The datasets presented in this study can be found in online repositories. The name of the repository and accession number can be found here: Pedestrian Dynamics Data Archive hosted by the Forschungszentrum Jülich, `https://doi.org/10.34735/ped.2022.4`

Ethics statement

Ethical approval and written informed consent for participation were not required for the current study in accordance with the local legislation and institutional requirements. The original study involving humans was approved by the ethics board of the University of Wuppertal, Germany. The studies were conducted in accordance with the local legislation and institutional requirements. In the original study, written informed consent for participation and for the publication of any potentially identifiable images or data was obtained from the participants.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 899739 (CrowdDNA).

Acknowledgments

We thank Mira Beermann for her advice on the statistical analysis and Juliane Adrian for the video visualization of the pushing rating data.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frsps.2023.1263953/full#supplementary-material

References

[1] Victoria Filingeri, Ken Eason, Patrick Waterson, and Roger Haslam. "Factors influencing experience in crowds - The participant perspective". In: *Applied Ergonomics* 59 (2017), pp. 431–441. DOI: 10.1016/j.apergo.2016.09.009.

[2] Milad Haghani, Majid Sarvi, and Zahra Shahhoseini. "When 'push' does not come to 'shove': Revisiting 'faster is slower' in collective egress of human crowds". In: *Transportation Research Part A: Policy and Practice* 122 (2019), pp. 51–69. DOI: 10.1016/j.tra.2019.02.007.

[3] Juliane Adrian, Armin Seyfried, and Anna Sieben. "Crowds in front of bottlenecks at entrances from the perspective of physics and social psychology". In: *Journal of the Royal Society Interface* 17 (2020), p. 20190871. DOI: 10.1098/rsif.2019.0871.

[4] Iker Zuriguel, Iñaki Echeverría, Diego Maza, Raúl C. Hidalgo, César Martín-Gómez, and Angel Garcimartín. "Contact forces and dynamics of pedestrians evacuating a room: The column effect". In: *Safety Science* 121 (2020), pp. 394–402. DOI: 10.1016/j.ssci.2019.09.014.

[5] Helena Lügering, Dilek Tepeli, and Anna Sieben. "It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms". In: *Safety Science* 163 (2023), p. 106123. DOI: 10.1016/j.ssci.2023.106123.

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

[7] John Drury, David Novelli, and Clifford Stott. "Managing to avert disaster: Explaining collective resilience at an outdoor music event". In: *European Journal of Social Psychology* 45 (2015), pp. 533–547. DOI: 10.1002/ejsp.2108.

[8] Anna Sieben and Armin Seyfried. "Inside a life-threatening crowd: Analysis of the Love Parade disaster from the perspective of eyewitnesses". In: *Safety Science* 166 (2023), p. 106229. DOI: doi.org/10.1016/j.ssci.2023.106229.

[9] Dirk Helbing and Pratik Mukerji. "Crowd disasters as systemic failures: Analysis of the love parade disaster". In: *EPJ Data Science* 1 (2012), pp. 1–40. DOI: 10.1140/epjds7.

[10] Ezel Üsten, Helena Lügering, and Anna Sieben. “Pushing and Non-pushing Forward Motion in Crowds: A Systematic Psychological Observation Method for Rating Individual Behavior in Pedestrian Dynamics”. In: *Collective Dynamics* 7 (2022), pp. 1–16. DOI: 10.17815/CD.2022.138.

[11] Dirk Helbing, Illés Farkas, and Tamás Vicsek. “Simulating dynamical features of escape panic”. In: *Nature* 407 (2000), pp. 487–490. DOI: 10.1038/35035023. arXiv: cond-mat/0009448.

[12] Colin M. Henein and Tony White. “Agent-Based Modelling of Forces in Crowds”. In: *Multi-Agent and Multi-Agent-Based Simulation. MABS 2004. Lecture notes in Computer Science, vol 3415*. Ed. by P. Davidsson, B. Logan, and K. Takadama. Berlin, Heidelberg: Springer, 2005, pp. 173–184. DOI: 10.1007/978-3-540-32243-6_14.

[13] Nuria Pelechano, Jan M. Allbeck, and Norman I. Badler. “Controlling Individual Agents in High-Density Crowd Simulation”. In: *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. 2007, pp. 99–108. URL: https://repository.upenn.edu/hms/210.

[14] Sujeong Kim, Stephen J Guy, Karl Hillesland, Basim Zafar, Adnan Abdul-Aziz Gutub, and Dinesh Manocha. “Velocity-based modeling of physical interactions in dense crowds”. In: *Vis Comput* 31 (2015), pp. 541–555. DOI: 10.1007/s00371-014-0946-1.

[15] Michael Fischer, Gaspard Jankowiak, and Marie-Therese Wolfram. “Micro- and macroscopic modeling of crowding and pushing in corridors”. In: *Networks and Heterogeneous Media* 15 (2020), pp. 405–426. DOI: 10.3934/nhm.2020025.

[16] Gabrielle Riches. “Embracing the Chaos: Mosh Pits, Extreme Metal Music and Liminality”. In: *Journal for Cultural Research* 15 (2011), pp. 315–332. DOI: 10.1080/14797585.2011.594588.

[17] Jesse L. Silverberg, Matthew Bierbaum, James P. Sethna, and Itai Cohen. “Collective Motion of Humans in Mosh and Circle Pits at Heavy Metal Concerts”. In: *Physical Review Letters* 110 (2013), p. 228701. DOI: 10.1103/PhysRevLett.110.228701.

[18] Timothy Janchar, Chris Samaddar, and David Milzman. “The mosh pit experience: Emergency medical care for concert injuries”. In: *The American Journal of Emergency Medicine* 18 (2000), pp. 62–63. DOI: 10.1016/S0735-6757(00)90051-2.

[19] Andrew M. Milsten, Joseph Tennyson, and Stacy Weisberg. “Retrospective Analysis of Mosh-Pit-Related Injuries”. In: *Prehospital and Disaster Medicine* 32 (2017), pp. 636–641. DOI: 10.1017/S1049023X17006689.

[20] John Drury, Chris Cocking, Steve Reicher, Andy Burton, Damian Schofield, Andrew Hardwick, Danielle Graham, and Paul Langston. “Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model”. In: *Behavior Research Methods* 41 (2009), pp. 957–970. DOI: 10.3758/BRM.41.3.957.

[21] Anna Sieben, Jette Schumann, and Armin Seyfried. "Collective phenomena in crowds—Where pedestrian dynamics need social psychology". In: *Plos One* 12 (2017), e0177328. DOI: 10.1371/journal.pone.0177328.

[22] Kathryn M. Zeitz, Heather M. Tan, M Grief, Pc Couns, and Christopher J. Zeitz. "Crowd Behavior at Mass Gatherings: A Literature Review". In: *Prehospital and Disaster Medicine* 24 (2009), pp. 32–38. DOI: 10.1017/S1049023X00006518.

[23] Ralph H. Turner and Lewis M. Killian. *Collective behavior*. 3rd ed. Englewood Cliffs, N.J: Prentice-Hall, 1987.

[24] Leon Mann. "Queue Culture: The Waiting Line as a Social System". In: *American Journal of Sociology* 75 (1969), pp. 340–354. DOI: 10.1086/224787.

[25] Bernd H Schmitt, Laurette Dube, and France Leclerc. "Intrusions Into Waiting Lines: Does the Queue Constitute a Social System?" In: *Journal of Personality and Social Psychology* 63 (1992), pp. 806–815. DOI: 10.1037/0022-3514.63.5.806.

[26] Institute for Advanced Simulation 7 - Civil Safety Research - Forschungszentrum Jülich. *Crowds in front of bottlenecks from the perspective of physics and social psychology [Data set]*. 2018. DOI: 10.34735/ped.2018.1.

[27] Maik Boltes, Armin Seyfried, Bernhard Steffen, and Andreas Schadschneider. "Automatic Extraction of Pedestrian Trajectories from Video Recordings". In: *Pedestrian and Evacuation Dynamics 2008*. Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 43–54. DOI: 10.1007/978-3-642-04504-2_3.

[28] Knut De Swert. *Calculating inter-coder reliability in media content analysis using Krippendorff's Alpha*. Center for Politics and Communication, 2012, pp. 1–15.

[29] Georges Voronoi. "Nouvelles applications des paramètres continus à la théorie des formes quadratiques." In: *Journal für die reine und angewandte Mathematik* 133 (1908), pp. 97–178.

[30] Shinichi Nakagawa and Holger Schielzeth. "A general and simple method for obtaining R^2 from generalized linear mixed-effects models". In: *Methods in Ecology and Evolution* 4 (2013). Ed. by Robert B. O'Hara, pp. 133–142. DOI: 10.1111/j.2041-210x.2012.00261.x.

[31] Kenneth P. Burnham and David R. Anderson. "Multimodel Inference: Understanding AIC and BIC in Model Selection". In: *Sociological Methods & Research* 33 (2004), pp. 261–304. DOI: 10.1177/0049124104268644.

[32] Gustave Le Bon. *The crowd: A study of the popular mind*. New York: Macmillan Company, 1896.

[33] Ralph H. Turner and Lewis M. Killian. *Collective behavior*. 2nd ed. Englewood Cliffs, N.J: Prentice-Hall, 1972.

[34] Otto M. J. Adang. "Initiation and Escalation of Collective Violence: An Observational Study". In: *Preventing Crowd Violence*. Ed. by Tamara D. Madensen and Johannes Knutsson. Lynne Rienner Publishers, 2011, pp. 47–68. DOI: 10.1515/9781588269416-006.

[35] Ezel Üsten, Jette Schumann, and Anna Sieben. "Exploring the Dynamic Relationship between Pushing Behavior and Crowd Dynamics". In: *Collective Dynamics* 8 (2023), pp. 1–29. DOI: 10.17815/CD.2023.143.

PUBLICATION III

A Rumor has Spread like Wildfire? - Empirical Investigation of Information Propagation in Waiting Crowds

Note: This article was published as Lügering, H., & Sieben, A. (2024). A Rumor has Spread like Wildfire? - Empirical Investigation of Information Propagation in Waiting Crowds. Collective Dynamics, 9, 1-28. https://doi.org/10.17815/CD.2024.146

Author Contributions

CONCEPTUALIZATION: Helena Lügering
DATA CURATION: Helena Lügering
FORMAL ANALYSIS: Helena Lügering
INVESTIGATION: Helena Lügering and Anna Sieben
METHODOLOGY: Helena Lügering
SUPERVISION: Anna Sieben
VALIDATION: Helena Lügering
VISUALIZATION: Helena Lügering
WRITING–ORIGINAL DRAFT PREPARATION: Helena Lügering
WRITING–REVIEW AND EDITING: Helena Lügering and Anna Sieben

A Rumor has Spread like Wildfire? - Empirical Investigation of Information Propagation in Waiting Crowds

Helena Lügering[1 2] and Anna Sieben[3]

[1] Institute for Advanced Simulation 7: Civil Safety Research, Forschungszentrum Jülich, Jülich, Germany
[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany
[3] School of Humanities and Social Sciences, University of St. Gallen, St. Gallen, Switzerland

Abstract

Information propagation in crowds tends to have a negative image. A common narrative is that information about a danger spreads like wildfire and leads to panic. In contrast, using person-to-person information sharing in crowd management as a complement to other communication channels has been discussed less. Even though previous research indicated that information does not propagate easily in crowds, more detailed research is lacking. In this study, two different experiments are presented to provide initial insights. In the main experiment, five groups of 33-41 participants took part in a total of 35 runs. In each run, a person in a waiting group was given a message or command that had to be passed on, whereby the knowledge about the task, the relevance of the message and the input side were varied. In the second experiment, this procedure was repeated with two larger groups of participants ($n = 91$ and $n = 101$). Overall, results showed that information propagated better when people were properly briefed on their task and have performed it several times. There was also a tendency for a higher density to foster faster propagation and for participants to rely on the spoken word rather than seeing a behavior performed. Yet, some participants did not receive the information at all or did not pass it on. In general, the direction of communication (e.g., back to front or left to right) was not always the same but information was usually passed along in a similar direction from where it came.

Keywords: information propagation ; crowd dynamics; crowd management; communication

1 Introduction

Communication is an essential part of human interaction. We talk with friends, family or at work, in pairs or in small groups. Generally speaking, people can be multipliers of information. If a person knows something, he or she might pass it on to others. However, this is not always the case; sometimes information is withheld - consciously or unconsciously, intentionally or unintentionally. The impact of people withholding information - often referred to as “stiflers” - and the proportion of people who receive a piece of information under different circumstances have motivated numerous theoretical works in recent decades [1–7]. Sudbury [8] was one of the first to use a simplified mathematical model to address this issue. He supposed that there are three types of villages: (1) those that have not heard of the rumor, (2) those that have heard of it and want to spread it, and (3) those that have heard of it but do not want to spread it further on. The villages then randomly call each other. If a village that does not know the rumor yet is called by one that wants to share it, the rumor is passed on. If a village that wants to share the rumor calls a village that already knows it, the first village loses interest and no longer wants to share it. All other calls have no effect. Given these assumptions, the probability of never hearing the rumor approaches 0.203 if the number of villages approaches infinity. Depending on the model and the parameters chosen, the exact number varies. However, most calculations indicate that information subside even if not all possible recipients have heard about it, and while the models are quite artificial and rarely have an empirical basis, this is definitely realistic. For larger crowds, such as concerts or religious gatherings, this fact could be especially relevant since these crowds typically consists of many smaller social groups and dyads which do not know each other. Even if people in waiting situations sometimes make contact with each other, they are usually more hesitant to talk to strangers than to people they know. Thus, the information transfer may be additionally impeded. Nonetheless, even when information is passed on, it does not always reach its recipient correctly. Simple models of communication, such as the Shannon-Weaver model [9], indicate that the transmission of a message between sender and receiver can be affected by various sources of interference (e.g., noise), and also the children’s game “whisper down the lane” taught us about this from an early age.

1.1 Information Transfer in Crowds

Looking at research on large-scale events, spreading of a rumor is referred to as one cause of crowd accidents [10, 11]. This assumption is often based on media reports of single disasters, in which, for example, is written “According to police officials present at the place of occurrence, a rumour about electric supply line falling on the crowd near the western exit gate of Gandhi Maidan spread like wildfire, triggering panic among people who ran for cover.” [12] or “Witnesses said panic spread over rumours of suicide bombers.” [13]. Empirical research revealed that ideas of false rumors spreading quickly in an emergency [14], or of information about something bad happening triggering crowd accidents [15] prevail among the general public as well. Drury et al. [14] showed that these assumptions are even held by professionals working in the field of public safety and emergency response. Aside from the fact that researchers criticize the emergence of panic [15] and assume that disasters are caused by multiple causes [10, 16], the notion that information spreads through a crowd like wildfire should be critically questioned, too. Not only do the general limitations of information transfer mentioned in the previous section contradict this view, but analyses of actual crowd accidents indicate that information travels poorly in very dense crowds, as people often do not realize how dangerous it has already become elsewhere [17, 18].

On the other hand, however, if there was a way for human-to-human information transfer to actually work well, crowd managers could make use of it. So far, there has been a variety of research on how to address people effectively in emergencies. For example, it is already established that personal approach (e.g., announcement, staff guidance) works better than the mere use of an evacuation alarm [19–21]. Additionally, Templeton et al. [22] recently published preliminary findings on the tone of voice when instructing people in face-to-face scenarios. Emergency personnel surveyed indicated that they find a facilitative tone, which conveys the requests as if the public is helping, and a teamwork tone, which creates a collective atmosphere, to be most effective. However, besides directly addressing people, the transfer of information between visitors could also be used. In fact, researchers have been arguing for some time that crowds should also be seen as a resource in emergencies that can be actively engaged [23–26]. But for this strategy, it is necessary to know whether information propagates in certain, predictable patterns, how reliable and quick information is passed on, and in which way a message has to be designed to ensure reliable and fast propagation.

Besides the propagation of mere verbal messages, it should also be considered that information can be accompanied by or trigger the execution of a behavior. In these cases, it is possible that information is not or not only distributed verbally in a crowd, but that an action propagates visually. For instance, a message might instruct people to move in a certain direction. While the message is likely to be passed on verbally at the beginning, at some point, people might just join the moving group without having received the word actively. From a psychological point of view, such behavior makes sense, since the well-known heuristic "Imitate-the-majority" states that it is generally a reasonable idea to follow the actions of others [27, 28]. Nevertheless, of course, this does not mean that everyone blindly follows the majority. In line with this, Haghani et al. [29] concluded in a comprehensive review that the idea of so-called herd behavior in an evacuation context is empirically untenable. Instead, the behavior of individuals partly depends on the behavior of others, but there are many factors determining the extent, for example, the type of behavior (e.g., movement initiation, exit choice) or the degree of uncertainty or urgency of the situation. People may even behave exactly contrary to the majority and/or follow a minority.

Similarly, studies on behavior in non-emergency situations showed that although people influence each other, they never all behave the same way. For instance, when a group of people on a street gazed upward at a certain point, some of the passersby followed this gaze [30]. The proportion of passersby looking up increased as the size of the stimulus group increased, up to 40% at a group size of 15. How long they looked up additionally depended on their own speed as well as on the density of the crowd around them. Lügering et al. [31] found out that the likelihood of someone starting to push increased by up to 30% if a person had a large number of contacts with other people pushing. All in all, however, these effects are rather small. Furthermore, it is important to keep in mind that behavior in a crowd is often only visible to a limited number of people. Those who stand too far away or with their backs to it may not even notice it and as Wirth et al. [32] recently showed, people are only influenced by the motions of neighbors who are visible to them. In other words, people who are in close proximity but occluded by others do not affect people's behavior. In extremely dense groups, this aspect is particularly relevant, since perception is usually limited to the people in the immediate vicinity [18]. However, when it comes to behavior that is visible to almost everyone, this can be highly appealing – sometimes with serious consequences. In their analysis of the witness statements of the Love parade disaster 2010, Sieben and Seyfried [18] showed that as soon as the first visitors began to climb the stairs or the poles out of the dangerously dense crowd, others tried to reach these places as well. While some moved there deliberately because they saw an opportunity to escape, others were carried along by the movement of the crowd which overall resulted in a densification in these areas.

Looking at these findings, it becomes clear that there is actually little research on how messages and behaviors propagate in crowds. Nevertheless, assumptions on person-to-person communication and sharing of information have already been incorporated into some crowd models. For example, Henein and White [33] implemented an egress scenario, where agents have different knowledge about the world and the exits. This knowledge can be expanded through discovery of spatial information or communication with others. Communication occurs when an agent wants to move to a field that is already blocked by another agent. The moving agent then communicates its worldview to the blocking agent in order to align it with the former agent's goals. Other models simply suppose that agents receive the information about hazards or exits when they are close to agents who are already informed [34, 35]. Such assumptions usually lead to the information propagating rather quickly. Kullu et al. [36] distinguished between meaningful and hollow communication and defined various complex conditions for different scenarios in which two people talk to each other with a certain probability (e.g., when someone just wants to communicate and a target is nearby or when someone wants to ask for directions during an evacuation). Furthermore, Yehua and Jing [37] addressed an idea similar to the newspaper articles cited above, namely that information transmission between people within a certain distance around a hazard leads to panic behavior unless there is beneficial calming information (e.g., proper instructions). Hoogendoorn et al. [38] additionally integrated the interplay of information and emotions into their model and their simulation showed that information propagates more or less easily depending on its relevance, its positivity and the level of fear of the agents.

1.2 The present study

This paper presents two experiments to address the research gap on information transfer in crowds and to provide initial empirical evidence. In order to simplify spatial information, we studied information propagation in a static phase of a moving crowd (in contrast to the models just introduced). In the first experiment, small groups (n = 33-41) took part in several runs, both as repeated runs (i.e., each group performed seven runs) and as between-subject runs (i.e., five different groups took part in the experiment). In the second experiment, we had two larger groups (n = 91 and n = 101), each performing one run. In every run, information was given to a person in the waiting group who then had to pass it on. For the analysis, the propagation of the information was first depicted descriptively and checked for any noticeable features. Then, it was examined whether there was a system in the direction the information traveled in. Finally, the speed of the information was investigated descriptively. This first part of the analysis was exploratory. In addition, one hypothesis was tested: We hypothesized that the information would travel faster if the group was more practiced, the people stood closer together and the information was additionally linked to a behavior, as it could then also be passed on visually. In terms of density, our hypothesis was contrary to the literature, which suggests that information is transmitted worse in extremely dense crowds. However, we assumed that at the low to moderate densities that we created in our experiment, it would be advantageous for the participants to be closer together since communication paths would be shorter and information could be overheard.

2 Method

2.1 General Experimental Set-up and Procedure

The analysis presented in this paper is based on two different data sets. The first one - which is our main data set - was collected during a series of experiments in the foyer of a building of the University of Wuppertal in 2022. The second one was collected during a series of experiments in an event hall in Duesseldorf in 2021 (see [39] for a detailed description of the overall organization of these experiments). The Duesseldorf experiment was designed to test the method and to initially examine the systematics in the propagation of information. In addition, these runs could be used for further interpretation and generalization of our results.

The general procedure was the same for both experiments and, in both cases, a bottleneck set-up was used (Figure 1). At the beginning of each run, the group of participants gathered within the experimental area and, on a signal, started to move through the bottleneck. About five seconds after the start, the flow was interrupted and one of the experimenters gave a piece of information to a participant standing at the edge. The information should then be passed on to the other participants. After an interruption time of maximum two minutes, the group was signaled to move again, passed the bottleneck and reassembled in the experimental area. In both runs of the experiment in Duesseldorf, the input information was "This is an additional experiment; we are playing "whisper down the lane". Tap yourself

Figure 1: Experimental set-up of the Wuppertal (left) and the Duesseldorf (right) experiment.

briefly on the shoulder as a sign that you received the information, and then pass it on." Both times, it started from the right side of the crowd. As the experimental design of the Wuppertal experiment was more complex, it is described in more detail below. Table 1 provides an overview of the experimental set-ups and procedures to compare the two experiments. In both experiments, the runs were filmed through an overhead camera and the participants wore colored hats, which allowed to analyze the resulting videos with the software Petrack (e.g., to extract the trajectories) [40]. Further, in accordance with the Covid-19 rules back then, a face mask had to be worn during both experiments.

	Wuppertal	Duesseldorf
Set-up	Bottleneck	Bottleneck
Number of groups	5	2
Number of participants per group	33-41	91-101
Number of runs per group (relevant for this study)	7	1
Type of runs	Surprise, message, command	Command
Side of information input	Left, right, back, front	Right
Length of interruption	35-120 sec	120 sec

Table 1: Overview of the experimental set-up and procedure in the Wuppertal and Duesseldorf experiments

2.2 Procedure of the Main Experiments

The Wuppertal experiment took place on two mornings of a three-day series of experiments in May 2022. In total, it was repeated five times with different groups of participants, each consisting of 33 to 41 people. However, since a few participants already left the experimental area in the first seconds until the stop signal, the actual number analyzed per run varied between 30 and 38.

Upon arrival, the participants read the conditions of participation and signed the informed consent. They were then given an orange hat with a personal ID code, a wristband with the corresponding number, and a green and a blue sticky dot. The green sticky dot was to be placed on the left shoulder and the blue one on the right shoulder. Furthermore, age, gender and body height were noted using the personal ID. The height was required for the processing of the videos in Petrack. After the preparations were completed for all participants, the group gathered in the experimental area and was greeted by the main experimenter. The group was told that their task would be to walk at a brisk pace, without pushing, through the bottleneck and gather back in the experimental area afterwards. The participants were also informed that there would be a questionnaire (paper-pencil) after some of the runs and that crowd noises would be played throughout the experiment to make it less quiet. After answering all questions, the group positioned itself in front of the bottleneck and the experiment began.

The experiment consisted of a test run without interruption, seven runs with interruption and information input (with the first one being a surprise run, see section below), and three to four further runs, which, however, were conducted with a different purpose and hence will not be addressed in this paper. In total, everything took about 1.5 hours. During this time, the main experimenter stood slightly elevated at the head end outside the bottleneck and four additional experimenters stood at the four sides of the group to input the information without attracting much attention. The experimenter standing in front additionally blocked the bottleneck with her arm whenever participants were not supposed to pass through it. All experimenters wore blue vests to make them recognizable as such. Each run began with a "Go" signal and the opening of the bottleneck. After finishing all runs, the participants were thanked, they returned their hats and codes and received 15 € for their participation.

Runs 1-2: Test Run and Surprise Run

The experiment was advertised as a social interaction experiment. Thus, at the beginning, the participants did not know the actual purpose. In order to maintain this cover story, each group started with a test run in which the participants simply walked through the bottleneck (without interruption). This also served to familiarize them with the procedure. In the second run (i.e., the surprise run), the flow was then interrupted about five seconds after the start for no apparent reason. This was followed by a 60-second waiting period with no interaction with the participants. Shortly before the end of this period, the experimenter, who was standing slightly elevated, made a secret sign to the experimenter standing to the left of the group

Run	Type	Information	Position of Input
3	Command	"Peel the green sticky dot off your shoulder."	right
4	Message	"The color is yellow."	back
5	Message	"The color is red."	left
6	Command	"Peel the blue sticky dot off your shoulder."	front
7	Message	"The color is purple."	front
8	Command	"Tap yourself briefly on your own shoulder."	back

Table 2: Exemplary sequence of runs 3-8

(i.e., touching one elbow with the hand of the other arm). This experimenter then discretely approached one of the participants standing at the edge and whispered a message to her/him, which s/he was supposed to pass on. Depending on the condition, this message was either "We have a technical problem." (relevant condition) or "The color is blue." (irrelevant condition). Three of the five groups received the relevant information and the other two the irrelevant information. After another 60 seconds, the bottleneck was reopened, participants passed through it and assistants instructed them to fill in a questionnaire. After everyone had gathered in the experimental area again, the true purpose of the study and the interruption was revealed. The participants were now informed that further interrupted runs would follow, in which they were to pass on either a message or a command which they were also to execute (according to "whisper down the lane").

Runs 3-8: Interruption Runs

The surprise run was followed by six interruption runs. In order to keep participants' attention directed to the front, a colorful 2D animation video (without sound) was played on a screen. At the beginning of each interruption, either a message or a command was given by one of the four experimenters. Each group received three different messages and three different commands. The type of information and the side of the input was randomized (two times each from the front and from the back, and one time each from the right and from the left side). However, both the order of the messages and the order of the commands as such were the same for each group. For an exemplary sequence, see Table 2. The interruption phase lasted between 35 and 60 seconds ($M = 48$, $SD = 11$) and the participants filled in a short questionnaire after each run.

Questionnaire

The questionnaires were mainly intended to record whether the information was received and passed on. However, they differed somewhat depending on the type of the run. The questionnaire after the runs for passing on a message was the shortest. It only asked whether the participants received the message (yes/no) and, if so, what the message was (free text), how often it was heard (1-2x, 3-4x, more than 4x), and whether it was passed on (yes/no). In the runs where a command was passed on, the questions were a bit more extensive. Likewise, the participants stated if the command was received and, if so, what it was, and whether it was passed on. Furthermore, it was asked whether the command was heard, seen, or both, whether the behavior was executed (yes/no), and what triggered the execution (hearing or seeing it). If the command was (also) heard, it was asked how often (1-2x, 3-4x, more than 4x). If the behavior was (also) seen, participants indicated in how many people they observed the behavior (1-2 persons, 3-4 persons, more than 4 persons), and – in case the execution was triggered by seeing the behavior – whether they imitated the behavior immediately after observing it once or only when they observed it several times. For the surprise runs, the questionnaire was the same as for the message runs but there were additional items on participants' feelings during the run and on other psychological constructs. However, these items were not included here as it would be beyond the scope of this paper. In order to later link the answers to the questionnaires with the people in the videos, participants wrote down their personal ID on each sheet.

2.3 Participants

Both experiments were advertised through various channels, including print and social media, email lists of previous experiments, lectures and other information channels of multiple German universities. In the two runs of the Duesseldorf experiment, we had $n = 91$ and $n = 101$ participants with all but one person participating in only one of the two runs. Of the 191 participants, 46.6% were female, 50.8% were male, and 2.6% did not specify their gender. The 189 participants who reported their age were on average 41.3 years old ($SD = 17.1$). The overall sample ($N = 184$) in the Wuppertal experiment was similarly balanced in terms of gender (51.6% female, 46.7% male, 1.6% non-binary), but considerably younger ($M = 24.5$, $SD = 8.4$). This can be explained by the fact that the experiments in Duesseldorf explicitly aimed at a broad sample, which was fostered by the temporal framework (i.e., participants took part in different experiments during a complete day), the location (i.e., an event hall in the city), and a broad recruiting. In contrast, the experiment in Wuppertal lasted only 1.5 hours for each participant, took place in the foyer of a university building, and was advertised mainly through university channels, leading to a primarily student sample.

2.4 Data Preparation and Statistical Analysis

In order to analyze the videos in addition to the questionnaire data, they were processed as follows: First, the path on which the information traveled through the group was traced by looking closely at the videos, i.e., it was manually identified who communicated with whom. To do this, a screenshot was taken from the video about one second before the information input and from where to where the information was shared was marked with arrows as precisely as possible. The focus here was on verbal transmission, meaning that, in the command runs, cases in which participants performed the behavior without being actively talked to were not included. Then, the orientation of the arrows measured in angular degrees was determined using the

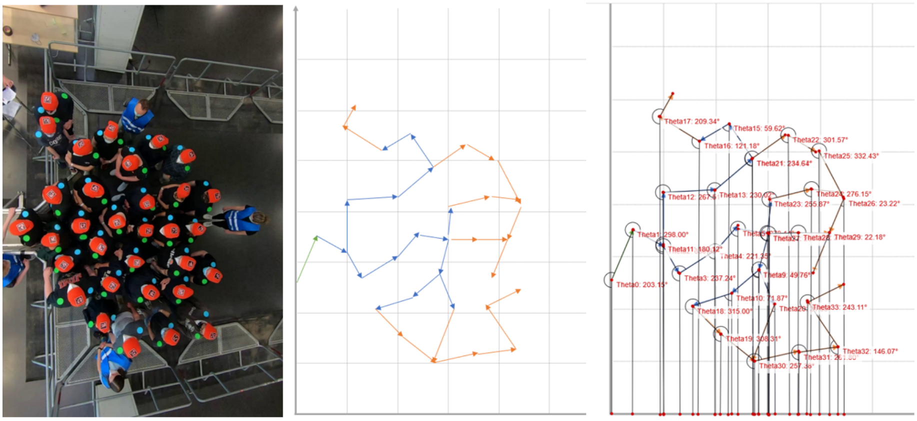

Figure 2: Identification of the direction of information transmission. (a) A screenshot of an exemplary run around one second before the information input. (b) Communication paths indicated by arrows (green = information input, blue = first 15 communication paths indicating the initial propagation, orange = all further communication paths). (c) Determination of the angles using WebPlotDigitizer. The coordinate system is for visual orientation only and has no further meaning.

WebPlotDigitizer [41]. Based on the orientation of the videos, an input from the right side corresponded to an angle of 0°, from the back to 90°, from the left side to 180° and from the front to 270° (Figure 2).

Additional to the directions, we identified the time at which participants received the information and, if applicable, executed the behavior. For this, the software Petrack was used which extracts the trajectories of all participants in the experimental area during an experimental run. This resulted in a data set including the exact position of every participant at each time point. Via an annotation function it was possible to enrich this spatiotemporal information with the data on the information propagation. Similarly to the analysis of the direction of information, this was done manually by carefully watching the videos. In addition, it was noted whether participants received and executed the message/command correctly. This was based on the responses from the questionnaires, particularly for the message runs. Please note that we always marked the first moment in which the participants received the information. If a person was contacted more than once, this was no longer taken into consideration. In the command runs, the participants typically were assigned a time for both the receipt of the message and the execution of the behavior. However, if the command was executed without prior communication, only this time was marked (even if there was an act of communication afterwards). If participants indicated in the questionnaire that they knew about the information but no one communicated directly with them and they did not perform the behavior, the moment was identified in which it was most likely that they became aware of it (either by overhearing or by seeing the command being performed). These two preparatory steps (i.e., identifying communication paths and time points of receipt) were done by a person who was not aware of the research questions and the hypothesis. To ensure

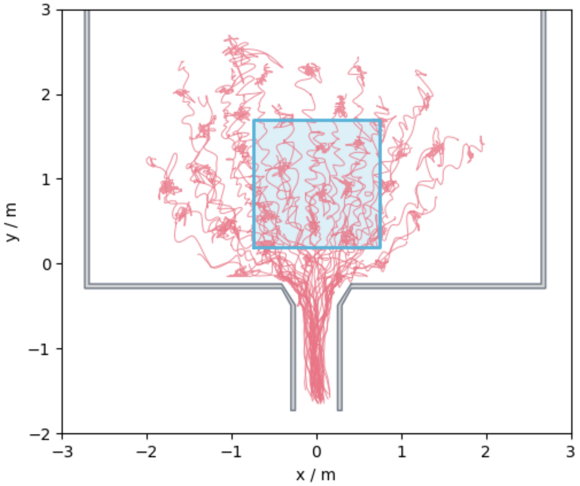

Figure 3: Measurement area for density calculation. The picture shows the experimental area (in meters) with individual trajectories (in red) of one exemplary run. "Knots" indicate the places where participants stood still. The blue square indicates the measurement area for the density calculation.

data quality, several runs were additionally checked by a second person revealing no systematic error and a satisfactory overlap. Finally, the average density in the measurement area (Figure 3) over the 25 seconds after the information input was calculated for each run using Pedpy [42].

Since this is an exploratory study, we used mainly descriptive statistics derived from the questionnaire data and videos to answer the research questions. In order to identify possible systematics in the direction of propagation, the angles of the communication paths were relativized to the angle of the intended information input of the respective run (right = 0°, back = 90°, left = 180°, front = 270°). Of course, the actual input did not always occur at this ideal angle, but was slightly different depending on the spatial positioning of the experimenter and the addressed participant. However, we assumed that the participants mainly noticed the information coming, for example, from behind and that the exact relation in space was not important. Lastly, we determined the time at which a certain number of participants received the respective information as well as the speed of propagation and calculated a linear multiple regression with the type of information (message vs. command), the sequence of the runs (i.e., first, second, third,...), and the respective density as predictors. The propagation speed in each run was defined as ratio of the total number of participants who were informed and the time span between the input of the information and the last person receiving it. Thus, the value indicates how long it took on average to inform one person, i.e., a higher value means a slower speed, and is therefore referred to as "transmission time".

For the analysis of the Wuppertal experiment, two message runs had to be excluded: In one run, the first participant walked around the group and informed several others; in the other run, one or two participants apparently spoke very loudly while sharing the message since suddenly several others who were actually standing farther away seemed to be informed. These behaviors were indeed nice and fostered a quick information propagation. However, they were not in accordance with the experiment instruction and therefore made the runs less comparable. Consequently, in addition to the five surprise runs, 28 interruption runs (13 message runs and 15 command runs) were taken into account.

3 Results

3.1 Wuppertal

Overall, the answers given in the questionnaires corresponded almost always to the observations in the videos (e.g., who received or executed a command). Nevertheless, there were minor discrepancies, such as a person who claimed not to have known about the message, but who was clearly spoken to. Also, there were some questionnaire responses that did not match (e.g., someone stated not having received the message, but then heard it 1-2x) or responses were missing. Since it was not possible to differentiate which information was correct, we kept all data as it was. This resulted in small deviations in the frequency data within the questionnaires as well as between the questionnaires and the observations in the videos, with minor discrepancies tolerated, while major discrepancies were addressed at the respective point.

Descriptive Statistics

Surprise Run. Due to the small number of the surprise runs ($n = 5$), only a descriptive analysis was reasonable. Table 3 shows how many people received the information per run (according to video observation) and how many seconds after the input the transmission of the information stopped (= duration of transmission). Please note that in one run of the relevant condition, one participant stated to know about the information, but it was not clear from the video how that person could have known. In another relevant run, however, one participant was clearly spoken to, but claimed not to have known. The mean values between questionnaire and video data across the relevant condition were therefore the same, even if there was a small discrepancy between single runs. Results showed that, on average, the information propagated similarly in the relevant and the irrelevant condition with a similar amount of people informed and a similar duration of transmission.

Runs 3-8. According to the questionnaire data, 95.1% of the participants received the message across all message runs. Of those, 43.0% heard it one or two times, 30.6% heard it three or four times, and 26.5% heard it more than four times. A large percentage of participants also reported passing the message on (86.7%). Likewise, in the command runs, a high proportion (85.0%) received the command. Of those, about two-thirds both saw and heard the command (65.7%), 24.8% only heard it and 9.5% only saw it. When the command was heard, it was received one or

		Number (percentage) of informed participants	Duration of information transmission (in sec)
Relevant Condition ("We have a technical problem.")	Group 1	8 (24.2)	9
	Group 3	10 (27.0)	16
	Group 5	10 (31.3)	19
	Mean	9.3 (27.5)	14.7
Irrelevant Condition ("The color is blue.")	Group 2	14 (46.7)	18
	Group 4	6 (15.8)	20
	Mean	10 (29.4)	19

Table 3: Descriptive data on information propagation in the surprise runs

two times in 44.7% of the cases, three or four times in 37.9% of the cases and more than four times in 17.5% of the cases. For observing the execution of the command, an inverse pattern emerged: A minority observed the behavior with only one or two persons (20.5%), 31.5% reported seeing it with three or four persons, and nearly half (47.9%) saw it with more than four persons. Although the absolute number of people who answered this question was about 35% higher than the number who indicated to have seen the behavior at all, the pattern was the same when only taking this subgroup into account. The majority of those who received the command also executed it (90.3%), and many passed it on (70.8%). When asked what triggered the performance, 82.3% stated performing the behavior because they heard it and only 17.7% because they saw it. When asked whether the command was executed after one or more observations (if it was executed after seeing), more than half of the responses were from participants who actually stated that they executed the command after hearing. When analyzing solely the data of participants who claimed executing the behavior upon observation, 71.4% performed it only after seeing it several times (instead of once).

According to video observations, in the message runs, 95.8% of the participants knew about the message and, in the command runs, 86.1% of the participants knew about the command. 92.6% of these also executed the command. Over all runs, the information traveled an average of 34 seconds.

Direction of Information Transmission

As stated in section 2.4, we focused on verbal communication for the identification of the direction of information transmission. Since there was one run in which the command propagated almost exclusively visually, this run was excluded from this analysis. Thus, 32 runs were included (5 surprise and 27 interruption runs), in

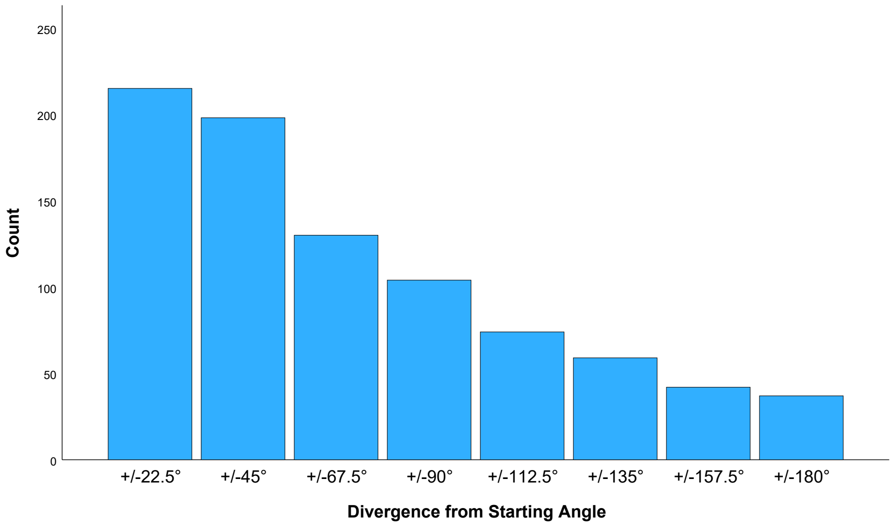

Figure 4: Frequency distribution of the divergence of the communication direction compared to the input direction (in angular degree). Almost 50% of the communication paths lie within a range of +/- 45° around the (intended) starting angle.

which – after deducting the 32 input directions given by the experimenters – 859 communication paths were identified. This number also includes information paths where participants were informed twice or more times. The frequency distribution of the communication directions relative to the angle of the information input is shown in Figure 4. In summary, the information often traveled more or less in the direction it came from, as 48.1% of the communication paths lay within a range of +/- 45° around the (intended) starting angle and 63.2% within a range of +/- 67.5°. If the range was broadening to +/- 90° around the start angle, it included already three quarters of all communication paths (i.e., a change of direction of max. 90°) and only in 24.7% of the cases the direction was completely changed.

Overall, the direction of communication did not change that often, what cannot be seen in the diagram but can be seen in the pictures with the directional arrows (Figures 2 & 7). In other words, it happened that information was passed on in a (completely) different direction than the one it came from. However, this new direction was then also temporarily maintained. Sometimes, these changes in direction occurred for no apparent reason, while other times they happened because someone shared the information with two or more participants (with only one standing in the previous direction) or because no one was left in the previous direction (i.e., at the border of a group). Following this observation, it is generally compelling to investigate what happens to information when reaching the border of a group, as it is then impossible to continue propagating in the previous direction. This is, of course, particularly important for smaller groups. Unsystematic observations of our data revealed that, on the one hand, the aforementioned changes in direction occurred. On the other hand, this was often the point at which the information

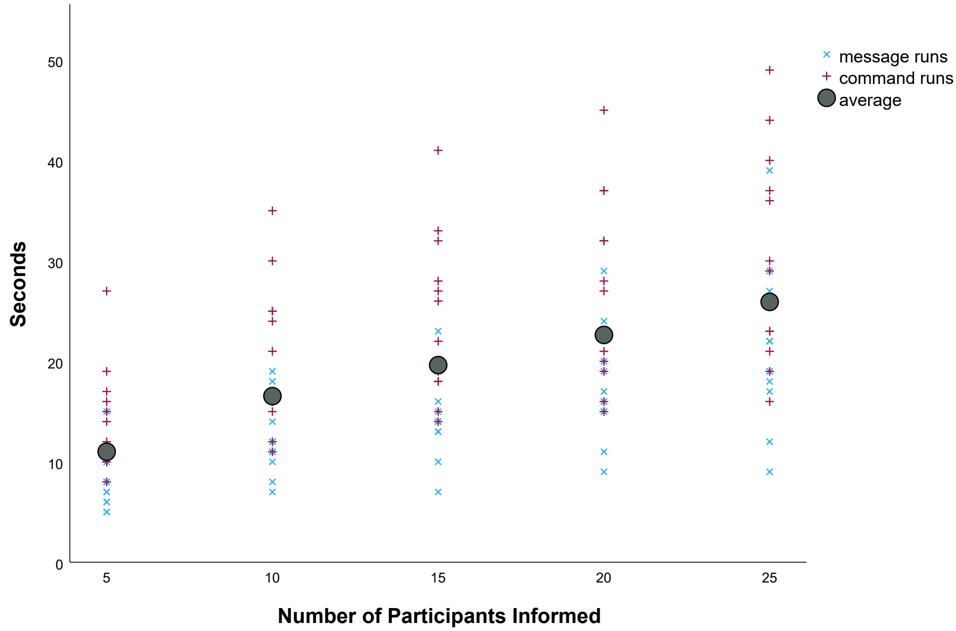

Figure 5: Time (in sec) until 5/10/15/20/25 participants received the information. Within one category, each cross represents one run. Each run is depicted in each category indicating the time needed in this run to inform the respective number of people. Overall, a linear relationship between the number of people informed and the time needed is shown.

subsided.

Speed of Information Transmission

As the surprise runs were not comparable with the other runs in terms of speed, only the message and command runs were taken into account for these analyses. To explore how long it took for a certain number of participants to be informed, Figure 5 shows – color-coded for message and command runs – after how many seconds a specified number of participants received the information as well as the mean value per category. Since only in 24 of 28 runs at least 25 participants received the information, four runs (one message and three command runs) were completely removed from this diagram. Due to a correlation between the transmission time and the probability of informing 25 people ($r = -.67$, $p < .001$), keeping these runs would have led to a distorted representation otherwise. In other words, the categories with the higher numbers of informed people would have appeared comparatively faster – although they may not have been – simply due to the dropout of the slower groups. All in all, Figure 5 shows a more or less linear relationship between the number of informed participants and the time needed, meaning that on average, it took about the same amount of time to inform five people, regardless of how many already

	B	95% CI	β	t	p
(Intercept)	2.92	[1.28, 4.57]		3.67	.001
Kind of information	0.38	[-0.07, 0.84]	0.29	1.73	.096
Density	-0.35	[-0.77, 0.07]	-0.29	-1.70	.102
Sequence of runs	-0.15	[-0.29, -0.01]	-0.39	-2.27	.033

Table 4: Regression results using kind of information, density and sequence of runs as predictors. Only the sequence of the runs significantly predicted the transmission time.

knew. Furthermore, it looks like the information traveled faster in the message runs than in the command runs as slightly more pink than blue crosses are placed above the average.

Determinants of Faster Propagation

For the following calculations, all 28 interruption runs were again considered. In order to get a first impression of the relationships between the transmission time and the kind of information, the sequence of the runs as well as the density, the correlations are depicted in Figure 6. The results showed a significantly negative relationship between transmission time and sequence of runs as well as transmission time and density, meaning the information traveled faster in later runs and in runs where it was denser, but not between transmission time and kind of information. The linear multiple regression model was also significant, $F(3, 24) = 4.41$, $p = .013$, $R^2 = .36$. However, only the sequence of the runs was a significant predictor of the transmission time. The predictors kind of information and density were not significant (Table 4).

Distortion of Information

So far, all analyses only considered whether a person has received information at all or whether a command has been executed altogether. This approach was chosen to determine general patterns of information propagation independent of the content. Nonetheless, it should not be ignored that information can also change when it is passed on. In our experiment, in two message runs and in four command runs, at least one person received the wrong information and/or executed the wrong command. For example, "the color is yellow" became "the color is green", the incorrectly colored dot was removed, or, the most common error, people did not tap themselves, but others on the shoulder. This resulted in 5.1% (message runs) and 9.9% (command runs) of the participants who received the information being misinformed. Moreover, 9.9% of those who executed a command did so incorrectly.

3.2 Duesseldorf

The propagation of the command diverged moderately in the two runs of the Duesseldorf experiment (Figure 7), but the results overall supported the findings from the

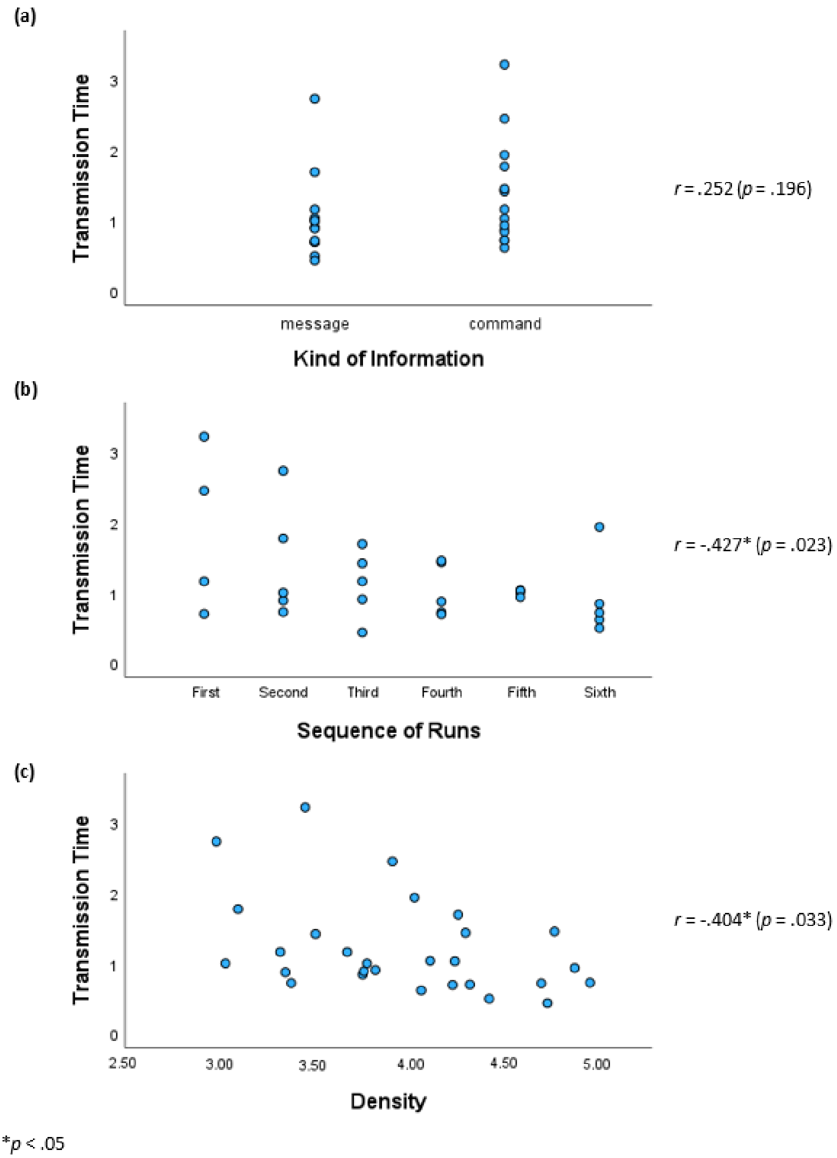

Figure 6: Correlation between transmission time (in sec) and (a) kind of information, (b) sequence of runs, and (c) density (in p/m^2). Each dot represents one run. There is a significantly negative correlation between the transmission time and the sequence of runs as well as the density, but not between the transmission time and the kind of information.

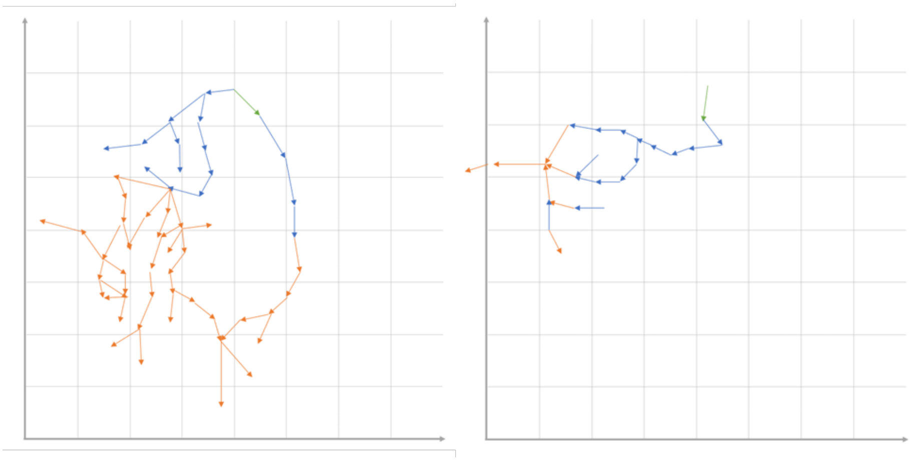

Figure 7: Direction of information transmission in the two runs of the Duesseldorf experiment. Figures are oriented so that the bottleneck is on the left-hand side (similar to Figure 2). The communication paths are indicated by arrows (green = information input, blue = first 15 communication paths indicating the initial propagation, orange = all further communication paths). The coordinate system is for visual orientation only and has no further meaning.

Wuppertal experiment. In both cases, the command traveled about 70 seconds until it subsided, but in one run 70 of 97 participants received it within this time (72.2%) and in the other run only 21 of 88 (23.9%). In the former, most of the participants followed the instruction correctly and put their hand on their own shoulder, but most of them also left it there. Seven participants first executed the instruction incorrectly and tapped on another person's shoulder, but later corrected themselves. In the latter, only the first person tapped on his/her own shoulder, whereas the remaining 20 tapped on the shoulder of another person (but removed it after a few seconds). The runs also differed with regard to the direction of information propagation. In the one where many people were informed, the command mainly traveled in the direction it started from (i.e., from right to left). From the 57 identified communication paths, 43.9% were within a range of +/-22.5° and two-third were within +/-45°, whereas the direction distribution of the 22 communications paths in the other run was less clear. Here, the orientation turned by 90° right at the beginning of the run (i.e., new orientation from the back to the front). Accordingly, almost two-third of the paths ranged from +/-45° to +/-135° while the others were distributed among the other directions.

4 Discussion

With our exploratory analysis, we were able to gain valuable insights on information propagation in crowds for crowd management and crowd research. First of all, our results showed that information propagation worked much better when people were aware of what their task is. Simply adding the instruction "please pass it on" when inputting the information (as it was done by the experimenter in the surprise runs) did not work properly, not even in the relevant condition. In fact, in our experiment, the relevant information was passed on no better than the irrelevant information, even though the participants had already waited a minute without any information, did not know what was going on, and the possible explanation came from an authorized person (blue vest). The promoting effect of relevance suggested by the simulation of Hoogendoorn et al. [38] could therefore not be shown in our experiments. Overall, the fact that information propagation worked equally (poorly) in all surprise runs was an unexpected result. Indeed, the data suggests slightly that the propagation in the relevant condition was more constant over the three runs (both speed and number of participants informed), as in the irrelevant condition once many participants were informed quickly and once only a few participants were informed slowly. However, since there were only three and two runs, this result should not be overinterpreted and further investigation is needed.

In the runs where participants knew about their task, in contrast, they really wanted to fulfill it. This is evident not only from the high proportion of informed people in general, but also from the analysis of the communication paths indicating that some participants were informed several times and that some also shared the information more than once. Further, the videos show that participants actively looked around after receiving the information to check who might not have been informed yet. Nevertheless, it must be mentioned that there were of course also participants who did not comply and did not pass the message (13.3%) or the command (29.2%) on, nor executed it (9.7%). Moreover, participants got better during the course of the experiment. This was reflected in faster propagation.

Interestingly, even though not statistically significant, the message seemed to propagate faster than the command, for which two ways of propagation would have been possible (i.e., verbally and visually). This finding was contrary to our hypothesis. A reason could be that, according to the questionnaire, many participants only executed the command after hearing instead of seeing it (a fact that could also be verified by the videos) and the phrase to pass on tended to be longer in the command runs than in the message runs (e.g., "Peel the blue sticky dot off your shoulder." vs. "The color is purple."). Although many participants indicated to have both heard and seen the command (65.7%), the order could not be deduced from our data. So, it could either be that people only became aware of the command because they were approached by others or that they felt more secure about executing the behavior after hearing the spoken word (even though they might have seen the behavior before). The latter explanation also fits the finding that most participants executed the command only after several observations rather than after one. However, it must also be considered that the given instruction of playing "whisper down the lane" might have left the participants feeling unsure, at least in the first runs, whether they were allowed to perform a behavior only on the basis of sight. Even so, latest after the first command run questionnaire, it should have been clear that this was a possibility since it was explicitly asked whether the command was executed upon sight. Additionally, in almost every run, there were participants who did so, even if they seemed to be more hesitant and looked around more to ensure themselves whether they were really doing the right thing. Only in one run, the command was transmitted almost exclusively visually, which resulted in one of the fastest command runs at 0.84 sec per informed person. This indicates that visual transmission might accelerate propagation but, in our experiment, it played a smaller role than expected.

A further finding was that the information changed in some of the runs (message as well as command). The command "tap yourself on the shoulder" seemed to be most susceptible since in two of the Wuppertal and in both Duesseldorf runs it became "tap others on the shoulder". Additionally, the participants in one of the Duesseldorf runs who correctly tapped on their own shoulder mostly left their hand there for quite some time which was considered correct for our analysis but, speaking strictly, it did not correspond 100% to the original command. This command apparently being the one with the most room for interpretation can also be seen in the run in which it propagated mainly visually. In this case, all participants executed the command correctly but, in the questionnaire, they stated, e.g., "tap on green" instead of "tap yourself on the shoulder". While all information presumably can change (e.g., in one run yellow became green), since some things sound similar or could be misunderstood, it is important to formulate messages/instructions with as little room for interpretation as possible. If the execution is additionally clearly visible to others (and perhaps remains visible for a longer time), there is a chance that individuals who were mistaken correct themselves or are being corrected by others (as in one of the Duesseldorf runs or in individual cases in Wuppertal, where participants first wanted to remove the green sticky dot when it should actually have been the blue one). In some cases, the participants even asked again or checked visually with others if they had understood the information correctly.

In general, participants did not seem to have a consistent preferred direction of communication, it was rather context-dependent. This is curious because it feels more effortful to communicate with the people at the front than with the people at the back, as the attention of the people at the front has to be gained first in order to be able to talk to them. In contrast, one only has to turn around to talk to the people at the back. In our experiments, however, it was mainly the case that the information traveled on in the direction it came from. Most of the time, this was the direction of the information input (i.e., if the information was given from behind, it tended to travel from the back to the front). However, if the direction of communication was changed for some reason (e.g., at the edge of the group), the new direction was also maintained for a longer time. This pattern was evident for the smaller groups in the Wuppertal experiment, but also for the larger group in the Duesseldorf experiment.

Concerning density, we saw at least on a correlative level what Zou and Chen [35] already assumed in their model: a higher density led to a faster propagation. Nevertheless, the maximum density reached in the Wuppertal experiment was about 5 people/m^2, which is only about half of the local densities reached in critical crowd situations (e.g., 10 people/m^2 in the Mina crowd disaster (Mecca, Saudi Arabia, 2006) [43]). As stated at the beginning of this paper, we assume an inverted u-shaped relationship between density and speed of information propagation, since reports of real life scenarios indicate an impeded propagation in very dense crowds [17, 18]. In this regard, however, further research on the exact mechanisms is needed.

Besides the fact that human senses might become more limited the closer people stand to each other, it also becomes more difficult to turn the upper body and people's attention in these situations may be focused more on themselves than on their surroundings (e.g., out of fear of losing contact with the ground).

Overall, based on the results of this study, the idea that a piece of information (e.g., a rumor) can trigger a crowd accident must at least be questioned. If people were not clearly instructed to inform others, propagation was slow and only a few received the information at all. If they knew about their task, on the other hand, the transmission worked much better, but nevertheless the information often subsided without everyone receiving it. The imitation of behavior also seems to have a smaller influence than initially thought, as most participants executed the command only because they were told to, not because they saw it in others. Moreover, if someone performed the behavior because s/he saw it, the performance was also usually more hesitant. Interestingly, this all happened despite the experimental setting in which the participants should have known that this kind of special behavior could be part of the run – at the latest after the first command run.

At this point, however, the difference between experiment and reality must be pointed out, since a potentially life-saving behavior in a life-threatening situation (such as climbing stairs or poles) probably offers a greater incentive for imitation than pulling a colored dot off the shoulder. In general, the relevance of information in real life scenarios – especially in critical ones – is clearly higher than e.g., technical problems in an experiment. Thus, the incentive to inform others may also be higher. However, in these situations, other constraints could impede the information transmission, such as background noise. Although it is unlikely that people in an emergency will pass on the information quietly, but rather shout, as at the The Who concert [17], there is also much more noise at a concert than in our experiment. In our case, raising the voice was very effective, as shown by the run that was excluded from the analysis, in which suddenly many participants received the message at the same time, presumably because someone had spoken louder. But even though we tried to break the silence by playing some crowd noise, it is probably so loud at large-scale events that even shouted information can only be heard and understood by the nearest neighbors. From a psychological perspective, a further constraint comes into play: In order for people to pass on information or imitate a behavior, they must perceive the source of information as reliable and assume that this person knows more than they do. This is also consistent with the finding of Haghani and Sarvi [44], which showed that people are more likely to follow others in an evacuation scenario when they themselves have less information about the exits. In our experiment, this should indeed hold true, since the participants knew that others might receive information at some point, whereas in reality this is not necessarily the case.

4.1 Practical Implications

On a practical level, our results show that person-to-person information sharing could be used for crowd management, as a supplement or when other technical information systems fail. However, this strategy cannot be taken for granted and some issues need to be considered. For example, people do not necessarily share information just because they are asked to do so in the situation. It is important to clearly declare the passing on as people's task right from the beginning in order to promote commitment and thus the likelihood of compliance. This could be achieved, for example, by addressing all visitors at the beginning of an event and informing them that person-to-person communication might be used in case of emergency. Such a global announcement would also be beneficial because people are then more likely to perceive others as a reliable source, as it is plausible that they have more information. At the same time, it must be ensured though that this mechanism is not abused and people deliberately distribute misinformation. Besides the risk of intentional propagation of false information, information can also easily change when it is passed on. However, this can at least be reduced by using short and unambiguous phrases. Generally, it would also be useful to train visitors in passing on information, which will probably not be possible at a single event. But if this strategy of information transmission becomes standard in crowd management, it is possible that the entirety of event visitors will improve over time.

4.2 Limitations and Further Research

Overall, we had a large number of people participating in our experiments ($N = 375$), but the individual experimental groups were rather small. In the main experiment in Wuppertal, they comprised only 33 to 41 people, and even if our main results were confirmed by the two runs with a larger group (Duesseldorf experiment), approx. 100 people is also a small number compared to large-scale events. It therefore remains to be investigated whether the findings on the systematics of information propagation obtained here apply to crowds of several thousands as well. In addition, we also had comparatively few runs. With only five surprise runs, 28 runs with a small group, and two runs with a large group, a purely descriptive analysis was more reasonable in most cases, and the conducted statistical tests must be treated with caution. Even though we found a consistent significant result (i.e., the influence of the sequence of the runs), other influences were inconclusive and only tendencies could be shown (e.g., message vs. command runs). It remains to be seen if the findings will be confirmed in a larger number of runs and if the tendencies prove to be true. To further examine the relationship between density and information propagation, runs with more variation are also necessary since the density in our experiment ranged only between 3 and 5 people/m^2. In addition, the investigation of other influencing variables such as the side of the information input on the propagation speed would be interesting for further research, as in our study we had too few runs per side to calculate a reliable analysis.

Another limitation is the usage of paper and pencil questionnaires in the time-limited situation between runs which may have affected the quality of our data. This can be seen in missing data, discrepancies between questionnaire data and video observations, or even inconsequential answers within individual questionnaires (e.g., command only heard, but then also seen in three to four persons). With computer-based survey methods, at least this inconsequence could have been prevented, since certain answer combinations could have been excluded per default. However, this type of survey was not feasible in our experimental setting. Some parts of the questionnaire may also have been too complicated (e.g., if the behavior was performed based on seeing it, was it due to one or several observations) or it was generally difficult to reflect the own behavior in retrospect (e.g., did someone perform a behavior upon seeing or hearing). The questionnaire methodology should therefore be improved in a subsequent study. However, as our analysis included not only the questionnaire but also the video data, we were able to support our main findings with observations, which is at the same time a strength of our study.

Lastly, we investigated the question of information propagation in crowds solely in an artificial, laboratory environment. However, as discussed above, the general conditions at a large-scale event might be quite different from those in an experiment (e.g., higher relevance of the information, but noisier). Therefore, field studies are needed to generalize our results and see if our advice proves helpful for crowd management in real-life situations.

5 Conclusion

To the best of our knowledge, this study is the first to systematically investigate information propagation in crowds. Our initial findings show tendencies that there is no preferred direction of communication, but that information travels mainly in the direction it comes from, that the spoken word has greater influence than a seen behavior, and that a higher density contributes to a faster propagation. Above all, it has been shown that it is not that easy to propagate behavior or messages in a crowd, but that it is possible to train people and increase the likelihood of sharing.

Acknowledgements

We would like to thank Juliane Adrian, Sina Feldmann and Armin Seyfried for their support in planning and implementing the experiments. We are also grateful to Alica Kandler, Maik Boltes, Ann Katrin Boomers, Tobias Schrödter, Mira Küpper and Mira Beermann for their help in setting up and conducting the experiments and curating the data. Finally, we would like to thank Marc Pelzer for his help with the data preparation.

Ethics statement

Full ethical clearance for both experiments was granted by ethical review committee of the University of Wuppertal, Germany. Written and informed consent was obtained from all participants who took part in the experiments.

Funding

Both experiments were conducted within the project CrowdDNA which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 899739.

Data Availability

The video recordings and head trajectories for both experiments as well as the questionnaire data for the Wuppertal experiment are available through the Pedestrian Dynamics Data Archive hosted by the Forschungszentrum Juelich, `https://doi.org/10.34735/ped.2022.5`

References

[1] Komi Afassinou. "Analysis of the impact of education rate on the rumor spreading mechanism". In: *Physica A: Statistical Mechanics and its Applications* 414 (2014), pp. 43–52. DOI: 10.1016/j.physa.2014.07.041.

[2] Virginia Giorno and Serena Spina. "Rumor spreading models with random denials". In: *Physica A: Statistical Mechanics and its Applications* 461 (2016), pp. 569–576. DOI: 10.1016/j.physa.2016.06.070.

[3] Valerie Isham, Simon Harden, and Maziar Nekovee. "Stochastic epidemics and rumours on finite random networks". In: *Physica A: Statistical Mechanics and its Applications* 389 (2010), pp. 561–576. DOI: 10.1016/j.physa.2009.10.001.

[4] M. Nekovee, Y. Moreno, G. Bianconi, and M. Marsili. "Theory of rumour spreading in complex social networks". In: *Physica A: Statistical Mechanics and its Applications* 374 (2007), pp. 457–470. DOI: 10.1016/j.physa.2006.07.017.

[5] Ray Watson. "On the size of a rumour". In: *Stochastic Processes and their Applications* 27 (1987), pp. 141–149. DOI: 10.1016/0304-4149(87)90010-X.

[6] Damián H. Zanette. "Dynamics of rumor propagation on small-world networks". In: *Physical Review E* 65 (2002), p. 041908. DOI: 10.1103/PhysRevE.65.041908.

[7] Laijun Zhao, Qin Wang, Jingjing Cheng, Yucheng Chen, Jiajia Wang, and Wei Huang. "Rumor spreading model with consideration of forgetting mechanism: A case of online blogging LiveJournal". In: *Physica A: Statistical Mechanics and its Applications* 390 (2011), pp. 2619–2625. DOI: 10.1016/j.physa.2011.03.010.

[8] Aidan Sudbury. "The proportion of the population never hearing a rumour". In: *Journal of Applied Probability* 22 (1985), pp. 443–446. DOI: 10.2307/3213787.

[9] Claude Elwood Shannon and Warren Weaver. *The mathematical theory of communication.* Urbana: University of Illinois Press, 1949.

[10] Maria Moitinho de Almeida and Johan von Schreeb. "Human Stampedes: An Updated Review of Current Literature". In: *Prehospital and Disaster Medicine* 34 (2019), pp. 82–88. DOI: 10.1017/S1049023X18001073.

[11] Ka Ming Ngai, Frederick M Burkle, Anthony Hsu, and Edbert B Hsu. "Human Stampedes: A Systematic Review of Historical and Peer-Reviewed Sources". In: *Disaster Medicine and Public Health Preparedness* 3 (2009), pp. 191–195. DOI: 10.1097/DMP.0b013e3181c5b494.

[12] DNA Web Team. *Patna stampede: Rumours may have a hand in tragedy; Bihar CM Jiten Ram Manjhi promises action, trashes 'Dummy CM' tag.* 2014-10-05. URL: https://www.dnaindia.com/india/report-patna-stampede-rumours-may-have-a-hand-in-tragedy-bihar-cm-jiten-ram-manjhi-promises-action-trashes-dummy-cm-tag-2023339 (visited on 09/15/2023).

[13] BBC News. *Iraq stampede deaths near 1,000.* 2005-08-31. URL: http://news.bbc.co.uk/2/hi/middle_east/4199618.stm (visited on 09/15/2023).

[14] John Drury, David Novelli, and Clifford Stott. "Psychological disaster myths in the perception and management of mass emergencies". In: *Journal of Applied Social Psychology* 43 (2013), pp. 2259–2270. DOI: 10.1111/jasp.12176.

[15] Helena Lügering, Dilek Tepeli, and Anna Sieben. "It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms". In: *Safety Science* 163 (2023), p. 106123. DOI: 10.1016/j.ssci.2023.106123.

[16] Dirk Helbing and Pratik Mukerji. "Crowd disasters as systemic failures: Analysis of the love parade disaster". In: *EPJ Data Science* 1 (2012), pp. 1–40. DOI: 10.1140/epjds7.

[17] Norris R. Johnson. "Panic at "The Who Concert Stampede": An Empirical Assessment". In: *Social Problems* 34 (1987), pp. 362–373. DOI: 10.2307/800813.

[18] Anna Sieben and Armin Seyfried. "Inside a life-threatening crowd: Analysis of the Love Parade disaster from the perspective of eyewitnesses". In: *Safety Science* 166 (2023), p. 106229. DOI: doi.org/10.1016/j.ssci.2023.106229.

[19] Guylene Proulx and Jonathan D Sime. "To Prevent 'Panic' in an Underground Emergency: Why Not Tell People the Truth?" In: *Fire Safety Science: Third International Symposium.* Ed. by Geoffrey Cox and Brian Langford. London: Elsevier Applied Science, 1991, pp. 843–852.

[20] David Purser. "Comparisons of Evacuation Efficiency and Pre-travel Activity Times in Response to a Sounder and Two Different Voice Alarm Messages". In: *Pedestrian and Evacuation Dynamics 2008.* Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 121–134. DOI: 10.1007/978-3-642-04504-2_9.

[21] C. Natalie van der Wal, Mark A. Robinson, Wändi Bruine de Bruin, and Steven Gwynne. "Evacuation behaviors and emergency communications: An analysis of real-world incident videos". In: *Safety Science* 136 (2021). DOI: 10.1016/j.ssci.2020.105121.

[22] Anne Templeton, Anezka Beamish Leskovcova, Gareth R. Clegg, Adam Lloyd, Elvira Fraser, Sentia Keyulong, and Sayaka Hinata. *Effective strategies in emergencies: First responders' views on communicating and coordinating with the public.* 2023. URL: https://www.research.ed.ac.uk/en/publications/effective-strategies-in-emergencies-first-responders-views-on-com (visited on 10/08/2023).

[23] Erik Auf der Heide. "Common Misconceptions about Disasters: Panic, the "Disaster Syndrome," and Looting". In: *The First 72 Hours: A Community Approach to Disaster Preparedness.* Ed. by M. O'Leary. Lincoln: iUniverse Publishing, 2004, pp. 340–380.

[24] Chris Cocking and John Drury. "Talking about Hillsborough: 'Panic' as Discourse in Survivors' Accounts of the 1989 Football Stadium Disaster". In: *Journal of Community & Applied Social Psychology* 24 (2014), pp. 86–99. DOI: 10.1002/casp.2153.

[25] Jennifer Cole, Montine Walters, and Mark Lynch. "Part of the solution, not the problem: the crowd's role in emergency response". In: *Contemporary Social Science* 6 (2011), pp. 361–375. DOI: 10.1080/21582041.2011.609332.

[26] John Drury. "Collective resilience in mass emergencies and disasters: A social identity model". In: *The Social Cure: Identity, Health and Well-Being.* Ed. by J. Jetten, C. Haslam, and S. A. Haslam. Hove, UK: Psychology Press, 2012, pp. 195–216. DOI: 10.4324/9780203813195.

[27] Robert Boyd and Peter J. Richerson. *The origin and evolution of cultures.* Oxford, New York: Oxford University Press, 2005.

[28] Gerd Gigerenzer. "Why Heuristics Work". In: *Perspectives on Psychological Science* 3 (2008), pp. 20–29. DOI: 10.1111/j.1745-6916.2008.00058.x.

[29] Milad Haghani, Emiliano Cristiani, Nikolai W.F. Bode, Maik Boltes, and Alessandro Corbetta. "Panic, Irrationality, and Herding: Three Ambiguous Terms in Crowd Dynamics Research". In: *Journal of Advanced Transportation* (Hindawi Limited 2019). DOI: 10.1155/2019/9267643.

[30] Andrew C. Gallup, Joseph J. Hale, David J.T. Sumpter, Simon Garnier, Alex Kacelnik, John R. Krebs, and Iain D. Couzin. "Visual attention and the acquisition of information in human crowds". In: *Proceedings of the National Academy of Sciences of the United States of America* 109 (2012), pp. 7245–7250. DOI: 10.1073/pnas.1116141109.

[31] Helena Lügering, Ahmed Alia, and Anna Sieben. "Psychological pushing propagation in crowds—Does the observation of pushing behavior promote further intentional pushing?" In: *Frontiers in Social Psychology* 1 (2023), p. 1263953. DOI: 10.3389/frsps.2023.1263953.

[32] Trenton D Wirth, Gregory C Dachner, Kevin W Rio, and William H Warren. "Is the neighborhood of interaction in human crowds metric, topological, or visual?" In: *PNAS Nexus* 2 (2023), pgad118. DOI: 10.1093/pnasnexus/pgad118.

[33] Colin Marc Henein and Tony White. "Microscopic information processing and communication in crowd dynamics". In: *Physica A: Statistical Mechanics and its Applications* 389 (2010), pp. 4636–4653. DOI: 10.1016/j.physa.2010.05.045.

[34] Ting Dong, Yan Liu, and Lin Bian. "A Behavior Model Based On Information Transmission for Crowd Simulation". In: *Transactions on Edutainment VI*. Ed. by Zhigeng Pan, Adrian David Cheok, and Wolfgang Müller. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2011, pp. 89–98. DOI: 10.1007/978-3-642-22639-7_10.

[35] Qiling Zou and Suren Chen. "Simulation of Crowd Evacuation under Toxic Gas Incident Considering Emotion Contagion and Information Transmission". In: *American Society of Civil Engineers* 34 (2020), pp. 1–18. DOI: 10.1061/(ASCE)CP.1943-5487.0000889.

[36] Kurtulus Kullu, Uğur Güdükbay, and Dinesh Manocha. "ACMICS: an agent communication model for interacting crowd simulation". In: *Autonomous Agents and Multi-Agent Systems* 31 (2017), pp. 1403–1423. DOI: 10.1007/s10458-017-9366-8.

[37] Chen Yehua and Bai Jing. "The dynamics of pedestrians' evacuation during emergency situations". In: *Adaptive Behavior* 24 (2016), pp. 399–410. DOI: 10.1177/1059712316672517.

[38] Mark Hoogendoorn, Jan Treur, C. Natalie van der Wal, and Arlette van Wissen. "An Agent-Based Model for the Interplay of Information and Emotion in Social Diffusion". In: *2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. Vol. 2. 2010, pp. 439–444. DOI: 10.1109/WI-IAT.2010.181.

[39] Ann Katrin Boomers, Maik Boltes, Juliane Adrian, Mira Beermann, Mohcine Chraibi, Sina Feldmann, Frank Fiedrich, Niklas Frings, Arne Graf, Alica Kandler, Deniz Kilic, Krisztina Konya, Mira Küpper, Andreas Lotter, Helena Lügering, Francesca Müller, Sarah Paetzke, Anna-Katharina Raytarowski, Olga Sablik, Tobias Schrödter, Armin Seyfried, Anna Sieben, and Ezel Üsten. "Pedestrian Crowd Management Experiments: A Data Guidance Paper". In: *Collective Dynamics* 8 (2023), pp. 1–57. DOI: 10.17815/CD.2023.141.

[40] Maik Boltes, Armin Seyfried, Bernhard Steffen, and Andreas Schadschneider. "Automatic Extraction of Pedestrian Trajectories from Video Recordings". In: *Pedestrian and Evacuation Dynamics 2008*. Ed. by Wolfram Klingsch, Christian Rogsch, Andreas Schadschneider, and Michael Schreckenberg. Berlin, Heidelberg: Springer, 2010, pp. 43–54. DOI: 10.1007/978-3-642-04504-2_3.

[41] *WebPlotDigitizer - Extract data from plots, images, and maps*. URL: https://automeris.io/WebPlotDigitizer/ (visited on 09/20/2023).

[42] Schrödter, T., The PedPy Development Team. *PedPy - Pedestrian Trajectory Analyzer*. 2022. DOI: 10.5281/ZENODO.7386931.

[43] Anders Johansson, Dirk Helbing, Habib Z. Al-Abideen, and Salim Al-Bosta. "From crowd dynamics to crowd safety: a video-based analysis". In: *Advances in Complex Systems* 11 (2008), pp. 497–527. DOI: 10.1142/S0219525908001854.

[44] Milad Haghani and Majid Sarvi. "Following the crowd or avoiding it? Empirical investigation of imitative behaviour in emergency escape of human crowds". In: *Animal Behaviour* 124 (2017), pp. 47–56. DOI: 10.1016/J.ANBEHAV.2016.11.024.

PUBLICATION IV

It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms

---

Note: This article was published as Lügering, H., Tepeli, D., & Sieben, A. (2023). It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms. Safety Science, 163, 106123. https://doi.org/10.1016/j.ssci.2023.106123

Author Contributions

Conceptualization: Helena Lügering
Data curation: Helena Lügering
Formal analysis: Helena Lügering, Dilek Tepeli and Anna Sieben
Investigation: Helena Lügering, Dilek Tepeli
Methodology: Helena Lügering, Dilek Tepeli and Anna Sieben
Resources: Helena Lügering
Software: Helena Lügering
Supervision: Anna Sieben
Visualization: Helena Lügering and Anna Sieben
Writing–original draft preparation: Helena Lügering and Anna Sieben
Writing–review and editing: Helena Lügering

It's (not) just a matter of terminology: Everyday understanding of "mass panic" and alternative terms

Helena Lügering[1 2], Dilek Tepeli[1 3] and Anna Sieben[1 4]

[1] Civil Safety Research, Forschungszentrum Jülich, Jülich, Germany
[2] School of Architecture and Civil Engineering, University of Wuppertal, Wuppertal, Germany
[3] School of Social Science, Ruhr University Bochum, Bochum, Germany
[4]School of Humanities and Social Sciences, University of St. Gallen, St. Gallen, Switzerland

Abstract

Although the idea of mass panic is quite common in reports on accidents involving crowds, most experts consider it to be erroneous. In a nutshell, they argue that panic and animalistic behavior of humans are not the main causes of crowd accidents, but that it is rather an organizational issue. However, few of the existing studies have addressed the question of what lay people associate with the term. With our mixed-method study, we sought to shed light on people's underlying ideas and assumptions about mass panic. Additionally, we were interested in how these ideas change using two alternative terms, namely "mass accident" and "mass disaster". Results showed that participants in the questionnaire ($N = 282$) and interview ($N = 17$) study indeed strongly associated the term "mass panic" with irrational and selfish behavior, and less with orderly behavior. In addition to the organizers, people in the crowd were seen as responsible for such accidents. Besides, most actions judged appropriate to defuse the situation were related to the advice "Don't panic". Deviating from the concept, however, it was indicated that helping behavior can be found in critical situations. The questionnaire in which participants only saw one of the three terms revealed no change in the everyday understanding with the alternative terms. Nevertheless, interviewees found their own "mass panic explanation" insufficient but also had no alternative ideas of what causes such accidents. Therefore, replacing the problematic concept of mass panic requires not only alternative terms, but also the dissemination of scientific explanations.

Keywords: mass panic; crowd accidents; everyday understanding; pedestrian dynamics; crowd crush

1 Introduction

Accidents in crowds are rare, but in recent decades, they have become more common occurrences and have caused many fatalities and injuries [1]. Whenever such a tragedy occurs, terms like "stampede" and "mass panic" are on everyone's lips, especially in the media. Of course, these terms are common, catchy, and everyone seems to know what must have happened. But what ideas and explanations of crowd accidents actually underlie this understanding? And, practically speaking, how does this everyday understanding shape the behavior of people in crowds?

The word "panic" is widely used in our everyday speech. According to an English dictionary [2] panic is a "sudden uncontrollable fear or anxiety, often causing wildly unthinking behavior". As this study was conducted in Germany, it is important to know that the German term *Panik* has a similar meaning [3]. From a scientific point of view, however, the term is more difficult to grasp – presumably partly because it is so common in everyday language [4, 5]. In his review, Dezecache [4] considered several definitions and proposed three core components of panic, namely a negative and highly intensive affect, a situation subjectively perceived as dangerous from which escape is difficult but not impossible, and the intention to rescue oneself even by irrational means and at the expense of others.

Building upon this interpretation, a mass panic might be understood as a situation in which such individual panic has spread to an entire crowd, leading to irrational, selfish, and competitive (escape) behavior that is ultimately responsible for fatalities and injuries in a crowd accident [6]. In this article, we call this the "image of mass panic". This perception is frequently conveyed and reinforced by media and popular culture [4, 7, 8]. As several studies [4, 9] have argued, the image of mass panic can be mainly traced back to Le Bon's idea of contagion, which suggests that emotions in crowds are as contagious as microbes [10]. At this point, we would like to stress that we are only referring to accidents in crowds in which the dynamic of the crowd itself is dangerous. Other incidents that can also have dire consequences in crowds, such as terrorist attacks or natural disasters, are not considered here. These incidents (or at least the fear of them) are only mentioned when we discuss possible assumptions about the causes of crowd accidents.

The image of mass panic was nevertheless questioned quite early [11, 12], and, to this day, many researchers in the field of crowd dynamics have criticized this concept. While some reject it as completely inappropriate [6, 7, 13–16], others consider it too general to refer to very different crowd accidents [17]. In short, the main points of criticism are as follows: First, in actual crowd accidents, competitive behavior can happen, but people also go to great lengths to help each other [9, 18–20]. Additionally, while people are of course scared – even to the point of fearing death – most people do not panic in an irrational sense, i.e., they do not run around screaming and pushing each other. Instead, their behavior is often calm, appropriate, and rational, even if it does not look like this from the outside (e.g., flight can be appropriate even if it looks "panicky"). Since people often have to decide how to react within a very short time based on insufficient information, their decisions are not always objectively right. But in the respective situation, it can indeed be the best decision [4, 12, 16]. Also, the most frequent causes of death are related to high density and pressure. People fall, for example, and others accidentally trample over or fall on them, or people asphyxiate even while standing because of extreme overcrowding [17, 20–22]. However, most people do not egoistically push each other and leave others to their own fate just to save themselves. Besides all that and maybe most important, the visitors themselves are not to blame for the accident. The term "mass panic" implicitly assumes that panic is the precipitating factor for the disaster, or, in other words, that nothing would have happened if no one had panicked [23]. However, cases from real-life scenarios have shown that people's movements in extremely dense crowds are often determined by the motion of the crowd rather than by their own intentions [20]. So, although it is very difficult to identify the guilty ones in retrospect (e.g., the Love Parade trial ended without a conviction [24], victim blaming appears to be the wrong approach. Last but not least, despite all problems with the term "panic", the term "mass" must also be critically questioned. Of course, panic behavior could occur in single cases, but then it affects only individual persons and not the entire crowd [15, 18]. When talking about a mass panic or a stampede (which is an animal analogy), though, one gets quickly to the issue of herding behavior. Although studies from various disciplines are not entirely conclusive [5] and there are examples of people adopting the behavior of others [25], it is also clear that human behavior is much more complex than simple, unthinking imitation – even in cases of emergency. However, it should be mentioned that there is also work in the field of crowd dynamics, especially in the area of modeling, that uses the image of mass panic uncritically, sometimes explicitly referring to experiments with animals (for an overview, see [5]).

The image of mass panic does not only influence research, but also professionals involved in public safety [6, 26] and emergency planning [27] – potentially with fatal consequences. For instance, visitors are not or at least not adequately informed about potential hazards (e.g., fire, terrorist attacks) due to fear of mass panic [13, 27]. However, there is evidence that clear information speeds up the evacuation [28]. Furthermore, it is conceivable that the image of mass panic together with the associated advice "Don't panic (and stay calm)" affects the behavior of the visitors themselves in critical situations. But to the best of our knowledge, it has not yet been investigated whether and in what way this is the case.

Due to this significant criticism, demands have been made to replace the term [14, 15]. Potential alternatives have been suggested, including "crowd crush" [29], or – for the German term *Massenpanik* – *Massenunglück* ("mass accident") (Dirk Helbing in [30]) or *Massendesaster* ("mass disaster") (Christian Zacherle in [31]). Establishing an alternative is difficult, though, as the term (mass) panic is deeply anchored in our linguistic usage, which is, in turn, maybe even due to the lack of alternatives. This is also shown by studies in which survivors and witnesses of crowd accidents were interviewed [9, 14, 18, 19]. They frequently used terms corresponding to mass panic but, when asked more specifically, their descriptions became more differentiated (e.g., panic as a justification of extreme behavior or simply as a description for individual distress [14]). Further, Nogami [32] found, in a Japanese sample, that various behavioral and emotional responses associated with panic (e.g., shouting, trembling) were not consistently used across two different mass emergencies (i.e., earthquake, plane incident). These findings demonstrate that, although the term (mass) panic itself is very dominant, it can mean quite different things. So, there is reason to doubt that all aspects which belong to the image of mass panic (i.e., irrationality, selfishness, wild pushing etc.) are exactly reproduced when lay people are asked in detail about their understanding. Moreover, it is not clear whether the associations change with a different, more appropriate term.

1.1 The present study

To understand which underlying ideas and assumptions lay people have about the image of mass panic and how they are connected to language, we conducted a mixed-method study consisting of an online questionnaire and a semi-structured interview. Interviewees were asked to articulate their everyday understanding of and associations with all three terms. Concerning the online questionnaire, we divided our sample randomly into three groups. Basically, all questions were the same between these groups, except that they were formulated to use one of the following German terms: *Massenpanik* ("mass panic"), *Massenunglück* ("mass accident"), or *Massendesaster* ("mass disaster"). This means that each participant only saw one of the terms and was not aware of the others. For ease of presentation, in this paper, we use the English terms "mass panic" (MP), "mass accident" (MA), and "mass disaster" (MD). Based on previous demands to replace the term, our main hypothesis was that the everyday understanding of crowd accidents is different for the three terms MP, MA, and MD, or, more precisely, that the term MP evokes the image of mass panic (i.e., irrational, selfish, competitive escape behavior that spreads in a crowd and leads to fatalities) more strongly than the two alternative terms. On the other hand, we expected the term MP to be more familiar since is it often used in the context of crowd accidents whereas MD and MA are rather untypical. As far as we know, there are no previous studies investigating the everyday understanding of all three terms in detail, so we conducted additional descriptive and explorative analyses.

2 Method

Ethical approval for this research was granted by the ethics board at the University of Wuppertal, Germany. Both studies were conducted in accordance with the Declaration of Helsinki. All participants who took part in either study gave informed consent.

2.1 Questionnaire

Sample

We recruited $N = 300$ participants (convenience sample) through the social networks of the authors, different social media platforms, the website surveycircle.com, as well as the email distribution lists of universities and different soccer fan clubs. Surveycircle is a platform where people interested in research can participate in online studies to earn points. These points can be used to promote own or other studies on the platform and thus make them more attractive for other participants. Eighteen of the initial participants were excluded due to insufficient knowledge of German (B2 or lower at the Common European Framework of Reference for Languages) or because they stated afterwards that they only clicked through. The final sample was $N = 282$, divided into three groups: $n = 97$ (MP), $n = 96$ (MD) and $n = 89$ (MA). Of these participants, 193 were female, 86 male, one non-binary, and two did not specify their gender. Ages ranged from 18 to 73 ($M = 29.91$, $SD = 11.04$) and most of them (97.2%) were native German speakers or had C2 proficiency of German, whereas 2.8% had C1. Most of the sample was well educated, with 96.5% having at least a higher education entrance qualification and 57.1% currently enrolled at an institution of higher education. Of those who were not students, most were employed (36.9% of the total sample size). The participants were not paid but the Forschungszentrum Juelich donated 0.30€ for each participation to a tropical rain forest foundation.

Structure

The construction of items was inspired by current literature [6, 8, 12–15, 18, 19, 26, 27, 32–37]. Especially the question concerning the source of knowledge was closely based on studies of Nogami [26, 36]. Besides the image of mass panic, the idea of helplessness meaning that people are passive, probably in shock and in need of being rescued by others [6, 12, 13, 37] was included. The questionnaire was in German and divided into nine parts addressing general ideas about crowd accidents, perceived levels of danger (slider item), sources of danger, options for action to defuse the situation, causes of occurrence, responsible parties, associations and familiarity with the respective term, and source of knowledge about crowd accidents. Exemplary items for each block as well as the total number of items and the respective measurement scale can be found in Table 1; a complete version of the translated questionnaire is provided in the Supplemental (Table S1). The anchors of the 7–point Likert Scales were chosen according to the questions (1 = least agreement, 7 = most agreement; see Supplemental (Table S1)). The order of the thematic blocks was the same for each participant, whereas the individual items were presented randomly. At the end, participants answered demographic questions about their gender, age, educational achievement, current occupation, and knowledge of the German language. We included an additional question in which participants could state if they had taken the survey seriously. The answer to this question had no consequences for them, but helped us to ensure the quality of the data.

Procedure

The questionnaire was conducted via the online questionnaire tool SoSci Survey (https://www.soscisurvey.de/), which is compliant with the German data privacy laws. The online format allowed us to reach various participants all over Germany. Answering the questionnaire took 10 – 15 min. On the first page, participants were informed about the topic of the study (behavior at large-scale events) as well as the procedure, their right of withdrawal without consequences, the anonymity of the questionnaire, the use of the collected data, and the donation. Additionally, contact information was provided for any questions. Once they confirmed their participation, they were randomly allocated to one of the three experimental conditions. The exact instruction on the following page was then dependent on the condition: "When answering the following questions, please imagine that you are at a concert (or a similar event with a lot of people) and a *mass panic / accident / disaster* occurs." The questionnaire was not about personal experiences but only about participants' general understanding of crowd accidents. However, an effort was made to protect those who might have had traumatic experiences at such large events by suggesting that those who had been affected should think carefully about participating in the survey and reminding them that they could abort the questionnaire at any time if necessary. After completion, the participants were thanked for their participation and had the possibility to read a brief explanation of the exact study purpose. This included a short summary of why the term mass panic is problematic and an outline of this study's aim of comparing the associations of the three different terms. For further questions or interest in the results, the contact information was mentioned again.

2.2 Semi-Structured interviews

To find out more openly and without pre-formulated statements (i.e., items) what lay people associate with the concept of mass panic and the alternative terms, we also conducted seventeen qualitative semi-structured interviews (in German) with acquaintances of colleagues and friends. Ten interviews were conducted in North-Rhine Westphalia, seven in Hamburg. They were scheduled for 60 min and informed consent was given before. Our interviewees were students, academics, employees, or self-employed, and between twenty-five and forty years old. Ten of them were male, seven female. Our interview guide consisted of an initial open-ended question about the experiences the interviewees had in crowds. In the demand section, they were asked what they associate with the term MP and how people should ideally behave in the event of a MP. Finally, they were asked what they understand by the alternative terms MD and MA. The interviewees were paid 10€ for their time and effort.

Thematic block	Question	Items
1. General (27 items, 7-point Likert Scale)	In a mass panic…	... People just start running without paying attention to others. ... People are paralyzed and cannot save themselves from the situation. ... People are able to deal with the situation reasonably and act sensibly despite the circumstances. ... People help the people they know.
2. Slider Danger (1 item, slider 0-100%)	How dangerous do you think a mass panic is?	- Harmless – life-threatening
3. Dangers (7 items, 7-point Likert Scale)	A mass panic is dangerous because…	... People stumble and fall. ... People no longer show consideration for others.
4. Defusing (12 items, 7-point Likert Scale)	In order to avert the danger in the event of a mass panic or to prevent the situation from getting worse, I should..	... Not let myself be infected by the feelings of the people around me. ... Call loudly for help and draw attention to the danger.
5. Causes (10 items, 7-point Likert Scale)	A mass panic can occur when…	... Things are not moving fast enough during an evacuation, resulting in a traffic jam. ... Information is passed on that something dangerous could happen (e.g., terrorist attack).
6. Responsibility (4 items, 7-point Likert Scale)	To what extent are the following groups of people responsible for causing a mass panic at a large event?	- Organizers of the event - Visitors
7. Association (12 items, 7-point Likert Scale)	How strongly do you associate mass panic with…	... Panic ... Help behavior / helping
8. Familiarity (7 items, 7-point semantic differential)	How do you feel about the term mass panic in the context of large events?	- Inappropriate – appropriate - Vague – clear
9. Source of Knowledge (10 items, multiple selection)	Where does your knowledge about mass panics come from?	- Media (e.g., (online) newspaper, news, radio) - Hearsay

Table 1: Thematic Blocks of the Questionnaire with Exemplary Items

3 Results

3.1 Questionnaire data

All statistical analyses were performed using IBM SPSS Statistics 27. The bar plots were created using a Python code. Mean values, standard deviations and, if applicable, factor loadings for all items can be found in the Supplemental (Table S1). The significance level was set at p <.05 for all statistical tests.

Familiarity of terms

To investigate the familiarity of the terms, we averaged the seven items separately for each questionnaire and calculated a one-way ANOVA ($F(2, 279) = 5.00$, $p = .007$, $\eta^2 = 0.03$) with following Games- Howell post–hoc tests. Results depicted in Figure 1 show that the term MP was slightly more common than MD, as expected, but not as MA.

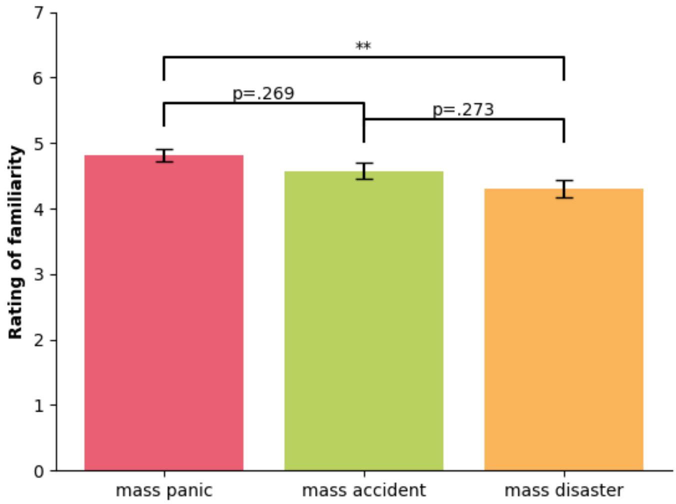

Figure 1: Comparison of Familiarity between the Three Terms. Mean values of familiarity ratings for each condition separately. Error bars represent standard error of the mean. $^{**}p < .01$

Differences between the three conditions

To test our main hypothesis that the MP questionnaire would evoke more associations in favor of the image of mass panic, we calculated one-way MANOVAs for the thematic blocks “General”, “Dangers”, “Defusing”, “Causes”, “Responsibility”, and

"Association". Results showed significant differences between the questionnaires on the combined dependent variables only for the block "Causes" ($F(20, 540) = 1.70$, p =.029, partial $\eta^2 = 0.06$, Wilk's $\Lambda = 0.89$). All other MANOVA models were not significant ("General", $F(54, 506) = 1.18$, p =.190, partial $\eta^2 = 0.11$, Wilk's $\Lambda = 0.79$; "Dangers", $F(14, 546) = 0.98$, p =.473, partial $\eta^2 = 0.03$, Wilk's $\Lambda = 0.95$; "Defusing", $F(24, 536) = 1.00$, p =.468, partial $\eta^2 = 0.04$, Wilk's $\Lambda = 0.92$; "Responsibility", $F(8, 552) = 1.76$, p =.082, partial $\eta^2 = 0.03$, Wilk's $\Lambda = 0.95$; "Association", $F(24, 536) = 1.50$, p =.061, partial $\eta^2 = 0.06$, Wilk's $\Lambda = 0.88$). For the block "Causes", the results of the following one-way ANOVAs and post-hoc Games-Howell tests can be found in Table 2. For reasons of clarity, only items for which the ANOVA was significant are presented here.

In addition to the six MANOVAs, we calculated a one-way ANOVA for the "Slider Danger", but here we found no significant difference between the groups ($F(2, 279) = 0.28$, p =.754). So, all in all, our main hypothesis was not confirmed. For almost all items, we found no difference between the three conditions. Only one of them differed significantly: The idea of contagious panic was more prevalent in MD than in MP. This difference was not in favor of our hypothesis and also rather small. Altogether, the ideas associated with the terms MP, MD, and MA were thus very similar, and we decided to combine the three questionnaires for all further descriptive and explorative analyses.

Block	Item	One-way ANOVA			Games-Howell Test		
		$F(2, 279)$	p	partial η^2	MP – MD M_{Diff} (p)	MP – MA M_{Diff} (p)	MD – MA M_{Diff} (p)
Causes	One person/a small group of people panic and other people are infected by that panic.	3.70	.026	.03	-0.39* (.022)	-0.27 (.209)	0.13 (.650)
	Information is passed on that something dangerous could happen (e.g., terrorist attack).	3.06	.048	.02	0.38 (.050)	0.10 (.769)	-0.28 (.236)

Table 2: One-way ANOVAs with Post-hoc Tests for Selected Items of the Block "Causes". *$p < .05$

Further descriptive and explorative findings

Even though our idea of different associations with different terms could not be confirmed, it is very important to gain a deeper understanding of the everyday understanding of crowd accidents. For this purpose, we first calculated two Principal Component Analyses (PCAs) for the blocks "General" and "Association" and determined how strongly participants agreed with the identified concepts. Furthermore, we investigated their assessment of what is dangerous in a crowd accident, which behaviors might help to defuse a critical situation, why a crowd accident occurs, who is responsible, and where they got their knowledge from.

Factor analysis. The general procedure for both PCAs was as follows: First, we checked for missing data but there was none. Then, the correlation matrix was analyzed. In order to better interpret the resulting structure of the factor loadings, we applied a varimax rotation. For factor extraction, we used the Kaiser's criteria, meaning eigenvalues ≥ 1 [38, 39], as a first step but also took theoretical considerations and interpretability of the factors into account. The scree plot for each PCA can be found in the Supplemental (Figure S1).

For the block "General", six factors were extracted when considering all factors with eigenvalues ≥ 1, which accounted for 56.27% of the total variance. However, this many factors were not suitable. Therefore, further solutions with two to five factors were tried out, and, in the end, a varimax-rotated four-factor solution was chosen (47.83% of total variance). This provided the most practical classification in terms of content, and most items loaded highly on just one of the four factors. As a result, it turned out that all items theoretically connected to the image of mass panic loaded on one common factor. In addition, the item "people are scared" (albeit less clear) as well as items connected to the idea of helplessness could be assigned to this factor. So, it was called "Mass Panic". The second factor, named "Orderly Behavior", included all items dealing with the idea that people stay calm, behave rationally, and that, if any, only single persons panic. The third ("Affiliation Behavior") and fourth factor ("Heterogenous Behavior") contained only two items, and the item "people help strangers" could not be clearly assigned to any factor. However, due to theoretical considerations, it was added to the factor "Affiliation Behavior".

For the block "Association", the extraction of factors according to Kaiser's criteria resulted in a model with three factors which accounted for 55.45% of total variance. Nevertheless, it turned out that a varimax-rotated two-factor solution (45.92% of total variance) was better in terms of interpretability and, here again, most items loaded highly on only one of both factors in this solution. The classification was divided into a factor "Mass Panic" (including "danger to life") and a factor "Orderly and Affiliation Behavior". The item "passivity / state of shock" could not be clearly assigned to either of the two factors, although, interestingly, it loaded somewhat higher on the factor "Orderly and Affiliation Behavior".

Underlying ideas and assumptions about crowd accidents. Based on the results of the PCAs, we recoded items if necessary and built new scales according to the extracted factors. For each scale, the number of items as well as the (averaged) inter-item-correlations and Cronbach's Alpha-coefficient can be found in the

Supplemental (Tables S2 – S8). Subsequent analysis showed that most scales differed significantly (Figure 2). For the block "General", a repeated measures ANOVA with a Greenhouse-Geisser correction was statistically significant ($F(2.77, 778.40) = 495.19$, $p < .001$, partial $\eta^2 = 0.64$). Bonferroni-adjusted post-hoc analysis revealed significant differences between all scales except for the difference between "Heterogenous Behavior" and "Mass Panic". This suggested that participants agreed more to "Mass Panic" and "Heterogenous Behavior" than to the other two scales. A paired t-test for block "Association" indicated a significantly higher agreement with "Mass Panic" than with "Orderly and Affiliation Behavior", as well.

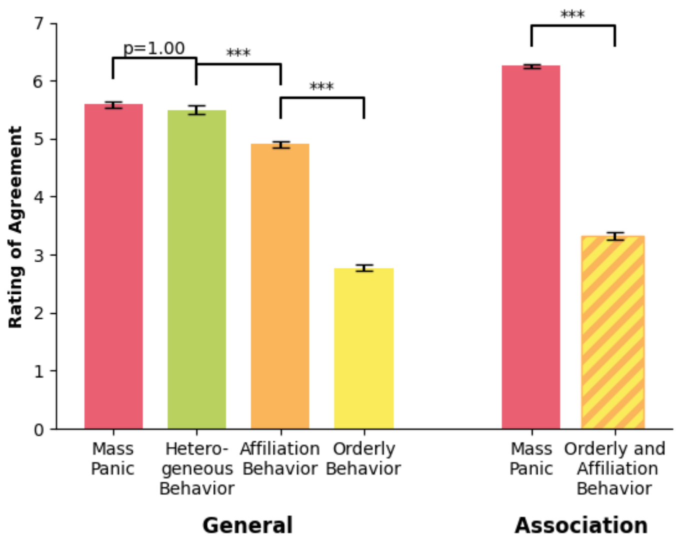

Figure 2: Comparison of Agreement between the New Scales of Blocks "General" and "Association". Mean values of the scales built after PCAs for blocks "General" and "Association" separately. Error bars represent standard error of the mean. ***$p < .001$

Perceived danger. Looking at the "Slider Danger" it became obvious that crowd accidents were perceived as something very dangerous ($M = 85.37$, $SD = 13.47$). The mean values of all items from the block "Dangers" ($M = 5.72$ to $M = 6.49$) further indicated that participants evaluated every danger as very plausible. However, a Greenhouse-Geisser corrected repeated measures ANOVA revealed even significant differences between them ($F(4.41, 1240.26) = 34.77$, p $<.001$, partial $\eta^2 = 0.11$). Bonferroni-adjusted post-hoc analysis showed that two dangers were considered especially plausible and differed significantly (all $p < .001$) from all others. These were "people stumble and fall" ($M = 6.40$, $SD = 0.83$) and "people on the ground are trampled upon by others" ($M = 6.49$, $SD = 0.84$). However, another danger that leads to fatalities in actual crowd accidents, namely "people can no longer breathe and, in the worst case, suffocate" ($M = 5.80$, $SD = 1.37$) was one of the dangers rated slightly less plausible.

Option for Defusing. Indicated by the mean values, participants divided all given options for action clearly into appropriate (all M greater than 5) and less appropriate (all M less than 4). According to participants' assessment, appropriate options were: remain calm and do not panic, keep your fear to yourself and remain calm, do not let yourself be infected by the feelings of the people around, look for a way out and instruct others to move there, help weaker people, and be fully informed about the circumstances to be able to make an informed decision. Whereas less appropriate options were: call loudly for help and draw attention to the danger, make it clear to others around you that you are afraid, do nothing at all and wait until someone rescues you, get out of the situation as quickly as possible – i.e., look for a way out, push your way out –, not be given any information about the exact circumstances because that will only cause more worries, and do what the people around you are doing. Building two scales according to this classification, the difference between both was significant ($t(281) = 36.96$, $p < .001$, $d = 2.20$). Furthermore, we compared the two questions about whether participants want to be informed or not with a paired t-test, too. Results showed that full information was significantly preferred ($t(281) = 9.49$, $p < .001$, $d = 0.57$).

Causes. As in the case of the dangers, all the given causes were considered as very plausible ($M = 5.28$ to $M = 6.45$). Only the item "sometimes a mass panic occurs for no reason at all and no one is to blame" was rated with slightly lower agreement ($M = 4.88$, $SD = 1.63$). Calculating a Greenhouse-Geisser corrected repeated measures ANOVA ($F(6.31, 1771.73) = 73.55$, $p < .001$, partial $\eta^2 = 0.21$) and following Bonferroni-corrected post-hoc tests, it turned out that the differences between this item and all others were significant (all $p \leq 0.023$). The item "one person/a small group of people panic and other people are infected by that panic" was assessed as somewhere in between ($M = 6.04$, $SD = 1.04$). Regarding the two questions as to whether a crowd accident occurs when people are informed about a possible hazard or not, a paired t-test revealed that participants thought it is more dangerous if information is passed on ($t(281) = 5.81$, $p < .001$, $d = 0.35$).

Responsibility. The results concerning the responsibility of the groups of persons are illustrated in Figure 3. A repeated measures ANOVA with Greenhouse-Geisser correction revealed how differently responsibility was perceived ($F(2.25, 633.01) = 93.28$, $p < .001$, partial $\eta^2 = 0.25$) with all differences being significant (Bonferroni-adjusted).

Source of knowledge. On average, three out of nine options for the source of knowledge were selected. Only two participants (0.7%) stated that they could not say. Most of the others acquired their information about crowd accidents from the media (94.3%). The proportions of participants who selected the other sources are, in descending order, social media (59.2%), non-fictional works (42.6%), hearsay (39.7%), experiences of other people from the social environment (25.9%), fictional works (24.5%), own experience (23.0%), and profession (3.2%). Further sources, each mentioned by one person, were emergency management, the Love Parade disaster in Germany, online games, and training courses as a fan representative.

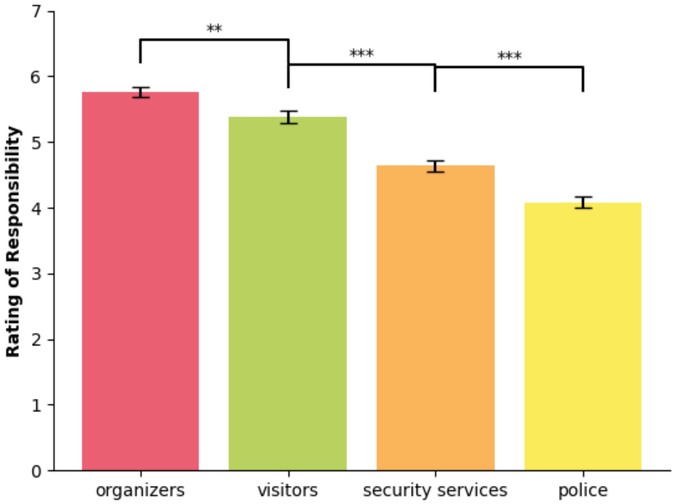

Figure 3: Comparison of Responsibility between the Groups of Persons. Mean values of responsibility ratings for each group of persons separately. Error bars represent standard error of the mean. $^{**}p < .01$; $^{***}p < .001$

3.2 Qualitative interview data

The qualitative interviews were audiotaped, and the parts on MP, MA, and MD selected and transcribed. The material was analyzed with qualitative content analysis [40]. In a first step, the everyday understanding of MP was identified in each interview and then compared with the other interviews and typologized. In a second step, the associative fields with MP were analyzed. In a third step, the suggestions on how to act in case of MP were extracted, as well as the associations with the terms MA and MD. Numbers in parentheses indicate in how many interviews the corresponding statements were found.

Interestingly, the individual understandings of MP differed in detail but not in general. Thus, only one type of lay theory could be reconstructed (Figure 4): In all interviews, an individual psychological state of panic was at the center of the theory. This was said to be accompanied by a fear of death and an instinct to flee. A narrowing of consciousness, decrease of rationality, the emotional expression of panic, and physical stress reactions were also attributed to the state of panic. This emotional state is triggered by different factors: The interviewees listed lack of organization, an increasing sense of crowding, an external threat such as a terrorist attack (or such a rumor), counterflow movements, and blocked exits, but also individual problems such as circulatory problems, fainting, or an initially individual panic attack. On an individual level, this condition was said to lead to a narrowing of perception, a lack of spatial orientation, and, as a consequence, irrational behavior. Irrational was the term used in interviews to describe behavior that cannot lead to the goal of escaping from the crowd and preserving one's well-being. On a social level, the state of panic was said to lead to a strong egocentrism. People focus only on themselves. In two interviews, however, it was mentioned that people would only care about their own children. Furthermore, apparently, there is a failure to observe social norms, with the result that weaker people are not protected. It was thus described that social solidarity breaks down. Only one of those interviewed mentioned that people may desire to stay with those they know. Furthermore, interviewees thought that communication in the crowd is no longer possible. In many interviews (9), it was said that the panic in the crowd intensifies as in a vicious cycle and people infect each other. However, none of the interviewees described this process of contagion in detail. These dynamics were said to lead to dangerous crowd behavior, to people running over each other, and also to the individual being rendered completely powerless. Many interviewees (10) added that people in panic are not solely responsible, however, but that mass panics can be prevented through better organization.

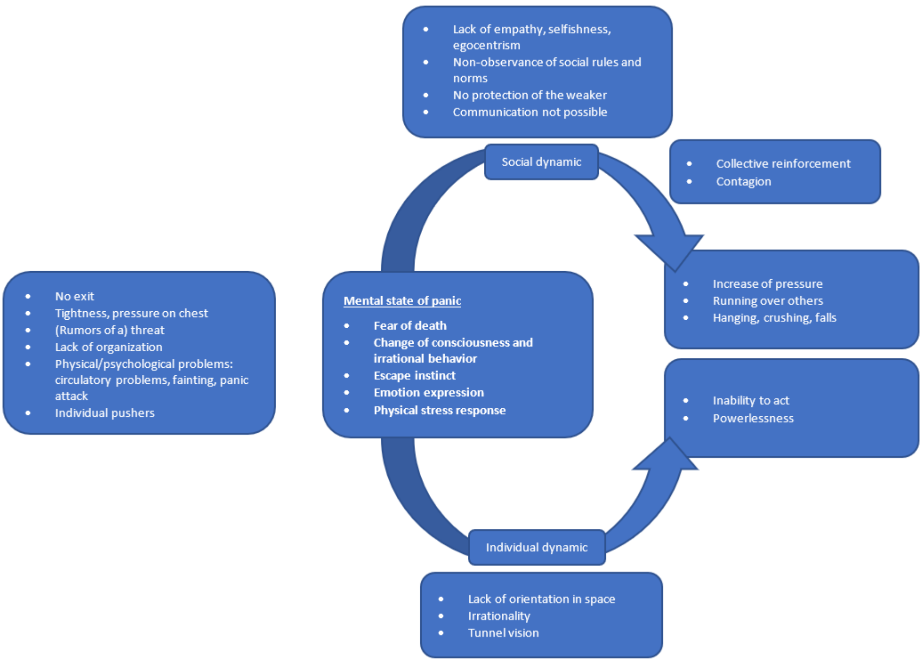

Figure 4: Lay Theory of Mass Panic Developed from Qualitative Interviews.

Most interviewees in our sample had no personal experience with situations of mass panic. Only two people had experienced a similar situation at a political demonstration. Most participants (13) mentioned the Love Parade disaster as an example (other examples mentioned less frequently were Mecca and the Heysel Stadium disaster in Belgium). Though none of them were present at these exemplary incidents, they heard about them primarily from the media. Due to the lack of their own experiences, the interviewees fell back on their own vague and associative ideas about MP, and it became clear that knowledge from other contexts was used to make the phenomenon plausible. We mainly identified two associative fields in which ideas about MP were embedded: The term MP was most strongly linked to biological concepts. Most interviewees (12) talked about people behaving instinctively. There was talk of the instinct of self-preservation. Another association taken from the animal kingdom was that of so-called swarm behavior. It is particularly interesting that one of the individuals interviewed used the neologism "swarm panic" (instead of swarm intelligence). Herd animals and stampedes were also mentioned. These biological ideas adhered to a social Darwinist understanding in which the instinct of self-preservation leads to the dissolution of social norms meant to protect the weaker. The crowd becomes a place where Hobbes' "war of all against all" prevails. Several interviewees (5) did use this biological framing, but at the same time made it clear that they were dissatisfied with this idea, as it actually did not correspond to their view of human beings and seemed insufficient. However, these same interviewees could not think of an alternative explanation. The second associative field of MP was dominated by psychological terms. Lay knowledge about the clinical phenomenon of a panic attack was used to imagine what happens in a mass panic. From this it was deduced that people in a state of panic are severely psychologically impaired and perceive very little, as well as being incapable of communicating with each other. A person in a panic attack was described as isolated from others. Interestingly, and contrary to our expectation, the image of "contagion" (and the associative field of spreading a disease) was hardly used in the interviews.

When asked how one should react in a mass panic situation, most individuals (11) indicated that one should remain calm, or calm others (emotion regulation). To get others out of their panic state and isolation, one should try to communicate with them. Many also said that the organization should be better in advance – but that in this situation it is no longer possible to behave properly. Some (3) suggested that one should try to face the pressure and not push. Other possible responses mentioned in individual interviews included following rules, sticking together and protecting each other, assigning a leader who gives clear commands, raising arms, and using a warning signal.

The interviewees were not familiar with the terms MA or MD and they could hardly think of anything to say about them (even less in the case of MD than of MA). Some interviewees (5) stated that there are a lot of deaths and injuries in a MA (examples: plane crash, shipwreck, natural disaster). A MA was seen as tragic and fateful – no one is to blame (however, sometimes the exact opposite was said). The term MD was interpreted by one interviewee as an ironic exaggeration (i.e., "disastrous date"). Others said that in a disaster there are clearly responsible parties. Overall, the interviewees found it difficult to articulate clear ideas about the two terms.

4 Discussion

In this mixed-method study, we investigated what underlying ideas and assumptions lay people have about crowd accidents and if these ideas change when using three different terms, namely *Massenpanik* (MP, mass panic), *Massenunglück* (MA, mass accident) and *Massendesaster* (MD, mass disaster). Thus, we have linked the criticism of the image of mass panic that has been expressed by scientists for more than 60 years [11, 12] to the public perception of the term. In the questionnaire as well as in the qualitative interviews, the image of mass panic was prevalent and answers were influenced by concepts like irrationality and selfishness. In the interviews, biological terms such as instinct, self-preservation, or "swarm panic" strongly shaped notions of mass panic. The feeling of helplessness also played a role, albeit more subordinate. In turn, the occurrence of orderly behavior was rated rather implausible in the questionnaire and also hardly mentioned by interviewees. The results for helping behavior were more ambiguous. In the questionnaire, items referring to affiliation behavior were sometimes rated higher and sometimes lower, whereas interviewees reported that, in such situations, people would only care for themselves and that social solidarity would break down. Just some of them said that people would help their own children or stay with others they know. In general, helping and staying together with people one knows seems to be perceived as more likely than helping strangers or forming a cohesive unit as a whole. Surprisingly, the image of contagion which is linked to Le Bon and often critically discussed in the scientific community was less prevalent in the interviews. Although, in the questionnaire, items concerning "contagious panic" showed relatively high levels of agreement, it became evident that the participants did not think that everyone behaves the same in these situations. Altogether, this indicated that the "mass" part is less problematic than the "panic" part.

Apart from this associative knowledge, many participants did not seem to have a clear understanding of why crowd accidents occur and what happens in these situations. Many possible causes were mentioned in the interviews (Figure 4); in the questionnaires, all provided causes were more or less affirmed. The same applies to possible dangers. Interestingly, the questions about information transfer in the two blocks "Causes" and "Defusing" were answered totally differently. On the one hand, people preferred to be fully informed so that they can make an informed decision. But on the other hand, they thought that a crowd accident is more likely to occur when information about a potential hazard is passed on. Research showed, however, that clear information speeds up the evacuation [28]. Not to mention the fact that the police and the security service cannot be everywhere [12, 41]. So, it is reasonable and also recommended by experts [9, 42, 43] to inform people – especially in case of emergency – giving them the chance to act on their own responsibility.

But in order for visitors to receive adequate information from those responsible or even from other people in the crowd, the emergency must first be recognized. In very dense crowds, people are often not aware of what is happening a few meters away and it is also extremely difficult to tell from the outside whether the situation is still normal or already dangerous [20]. Even if technical solutions for crowd management become more and more sophisticated, feedback from the crowd would be helpful in these cases. However, as our study indicates, this might be impeded by the associations that the term "mass panic" evokes. Our results show that lay people have a clear concept of how to behave in critical situations but most strategies for action were closely related to the advice "Don't panic"; actively drawing attention to oneself was mentioned only by few interviewees, and was also rated low in the questionnaire. In the interviews, it became evident as well that one main assumption is that people are panicking and they need to calm down so that nothing happens.

Apart from a possibly counterproductive recommendation for action, this notion also implies an allocation of blame. By assuming that nothing would have happened if no one had panicked, the interviewees stated implicitly that the people in the crowd are responsible for, or at least a decisive factor in, crowd accidents. In fact, this seems to be a general implication of the image of mass panic as it was also expressed in the questionnaire where visitors were ranked as the second most responsible group out of four (for further evidence, see [23]). Nevertheless, participants in both studies agreed that besides the visitors, the organizers are to blame for crowd accidents and that critical situations could be prevented by better organization in advance.

For the questionnaires, all these findings apply regardless of the term used for framing the items. In fact, we found only one difference between the conditions, which was also rather small (less than 0.5 scale point), namely that the idea of contagious panic was more prevalent in MD than in MP. Even though we do not want to ignore the fact that this difference contradicts our hypothesis, it should not be overestimated. By choosing MANOVAs as statistical method, we corrected for multiple testing as best we could. However, since we had many items and calculated numerous comparisons, this one small difference should probably be considered a false positive test result. The fact that the three terms – MP, MA, and MD – evoked such similar ideas in the questionnaire can be explained as follows: The image of mass panic is so predominant among our everyday speech that people automatically think of it whenever confronted with accidents at large events – there are no other concepts available. This can also be seen in the interviews, in which some interviewees were dissatisfied with their own explanations of mass panic, but had no alternative ones. So, potentially, the terms MD and MA did not create their own associations, but rather activated the concept MP and thereby the very same associations and ideas. In other words, it might have happened that the alternative terms were either mentally replaced by "mass panic" or that the questions were only processed figuratively, meaning participants had a picture of a "classical mass panic" in mind. Something very similar only the other way around has, in fact, happened in the interviews with survivors and witnesses of crowd accidents [9, 14, 18, 19] where participants had a different concept of the respective disasters but used the same terms associated with panic. Combined, these studies show how entangled ideas and language are and that it is difficult to change one without the other.

Interestingly, even though "mass panic" is used almost exclusively in our everyday language, the level of familiarity with the three terms was assessed similarly in the questionnaire study. The interviewees, in turn, had very few associations with the alternative terms (MD and MA) and judged them to be inappropriate for the crowd context. These divergent results can be explained by reference to the methods: In the questionnaire, the detailed items already specified what kind of misfortune was involved, whereas this additional information was not given in an open-ended question in the interview ("What do they imagine a MA to be?"). Since the familiarity question was asked toward the end of the questionnaire, participants had already been confronted several times with the alternative terms and could fill them with (familiar) associations. In practice, this means that the terms MA and MD are only associated with crowds when additional information is given. This can also be interpreted as a positive result, because it indicates how flexible the terms MA and MD are and how easily people can get used to something new.

All in all, our quantitative findings were largely in line with the qualitative ones. This is very important as the items of the questionnaire were created on a theoretical basis to determine lay people's agreement with the image of mass panic or alternative concepts discussed in the literature. Due to the large sample size, we were able to gain a good impression of the prevailing concept. However, the genuine problem with questionnaires is that the participants just have to agree or disagree with the given statements. It is hard to say whether the ideas would also have been produced spontaneously or whether completely different concepts would then have been described. Even though far fewer people were interviewed, the free descriptions again indicated a relatively clear predominance of the image of mass panic. Additionally, the findings from the interviews were partly able to explain or extend the questionnaire results. This supports the validity of our data from both studies and the general use of mixed methods. At this point, please note that we did not discuss every single item of the questionnaire in detail here (see Supplemental (Table S1) for descriptive analysis), though it would have been interesting. This paper should only give an overview about the everyday understanding and associations. Anything else would have been beyond the scope. However, the data set is open access and available to anyone who would like to conduct further analysis.

4.1 Limitations

Although, our mixed-method approach provided a powerful study design to gain valuable insights into the everyday understanding of mass panic, we do not want to omit the limitations. First of all, the non-representativeness of our sample can be criticized most. With $N = 282$ for the questionnaire study and $N = 17$ for the interviews, we had large sample sizes relative to the respective study designs. Concerning age and gender they were also reasonably diverse but overall, the participants were mostly highly educated. This is mainly due to the recruitment via the social environment of the authors and a survey platform that is primarily used by people interested in research. However, obtaining a representative sample was never aimed. Knowledge about crowd accidents is not part of the usual general education and most people get their information from (social) media – as our but also other studies [26, 37] have shown – which is equally accessible to all social classes. Therefore, there is no reason to expect different results with a more diverse sample. What is important to note, however, is that we only included German native speakers or people who speak German at a very high level. Thus, the study cannot say anything about associations of people who are just learning the language. This leads us to the next limitation, namely that the findings are of course only transferable to the German-speaking area. Nevertheless, at least in English and Japanese, the criticism of the terms "(mass) panic" or "stampede" is very similar. Therefore, our study provides first indications and may be useful for further investigations in other languages. Last but not least, we only examined two alternative terms. There might be another term that evokes completely different associations and reflects more accurately what happens in a crowd accident. Based on our results, however, we assume that the image of mass panic is so ingrained that no other term would automatically produce other, more accurate associations.

4.2 Practical implications

Our study shows that the rejection of the term "mass panic" is justified, especially since it is not just a language issue but also has fatal practical consequences. Firstly, there is the moral aspect of victim blaming after a crowd accident. According to our participants, visitors are the second responsible party for injuries and fatalities after organizers, who bear the primary responsibility. Secondly, similar to those responsible who may withhold information about the seriousness of the situation [13, 27], visitors do not seem to draw attention to the danger or themselves in critical situations. The results of our mixed-method study support the assumption that the advice "Don't panic (and stay calm)" is mainly followed. Even though crowd safety is clearly the responsibility of professionals, in some situations the expression of discomfort may help to recognize the danger at early stages. Nevertheless, further research is required to determine whether people actually behave this way and how this affects the dynamics of the situation – perhaps with the help of reconstructing past disasters. All in all, however, our findings support the calls for a new term that does not represent the old ideas, but the proposed terms MD and MA do not sufficiently fulfill this task. Simply exchanging the term is thus not enough; what is needed is a more elaborate education on the underlying theory of crowd accidents. The interviews showed that people are not satisfied with their existing explanations, so other explanations are likely to be accepted. Furthermore, it is important that the public will adopt the new term and perceived it to be appropriate and unambiguous. People appear to get used to a new term very quickly, however, it must be ensured that an alternative really evoke adequate associations. In our study, for example, the idea of contagious panic was even higher for MD. As mentioned above, this finding was unexpected, rather small and may be a false positive. However, further research is needed so that a new term could solve the problem instead of making it bigger. To advance the change, reporters and editors should be better informed about the problematic concept of mass panic, since most people gain their knowledge from (social) media. Prospectively, these channels could also be used to establish new, more realistic concepts and to disseminate scientific explanations of crowd accidents.

4.3 Conclusion

To conclude, we found out that lay people's ideas and assumptions about crowd accidents are rather vague and associative and closely related to the image of mass panic. We also learned that these ideas do not change just because a different term is introduced. Nevertheless, changing the term in combination with accurate information seems to be an important initiative. Therefore, science and (social) media should work together to develop and establish an appropriate alternative term and explanation of crowd accidents. This is the only way to overcome the classical image of mass panic with all its negative and fatal consequences.

Data Availability

The raw data of the questionnaires and the interview transcripts (in German) are available through the Pedestrian Dynamics Data Archive hosted by the Forschungszentrum Jülich, https://doi.org/10.34735/ped.2021.10.

Acknowledgements

We thank Lena Blind for writing the Python code for the figures, Jennifer Schirrmacher, Sandra Schaffarczik and Judith Kühn for transcribing the interviews and Armin Seyfried for comments on the interpretation of the interviews.

Funding

This study has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 899739 (CrowdDNA).

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssci.2023.106123.

References

[1] Claudio Feliciani, Kenichiro Shimura, and Katsuhiro Nishinari. *Introduction to Crowd Management: Managing Crowds in the Digital Era: Theory and Practice*. Cham: Springer International Publishing, 2021. DOI: 10.1007/978-3-030-90012-0.

[2] *Meaning & Definition for UK English — panic*. URL: https://www.lexico.com/definition/panic (visited on 12/03/2021).

[3] *Duden — Panik*. URL: https://www.duden.de/rechtschreibung/Panik (visited on 12/01/2021).

[4] Guillaume Dezecache. "Human collective reactions to threat". In: *Wiley Interdisciplinary Reviews: Cognitive Science* 6 (2015), pp. 209–219. DOI: 10.1002/wcs.1344.

[5] Milad Haghani, Emiliano Cristiani, Nikolai W.F. Bode, Maik Boltes, and Alessandro Corbetta. "Panic, Irrationality, and Herding: Three Ambiguous Terms in Crowd Dynamics Research". In: *Journal of Advanced Transportation* (Hindawi Limited 2019). DOI: 10.1155/2019/9267643.

[6] John Drury, David Novelli, and Clifford Stott. "Psychological disaster myths in the perception and management of mass emergencies". In: *Journal of Applied Social Psychology* 43 (2013), pp. 2259–2270. DOI: 10.1111/jasp.12176.

[7] Lee Clarke. "Panic: Myth or Reality?" In: *Contexts* 1 (2002), pp. 21–26. DOI: 10.1525/CTX.2002.1.3.21.

[8] Rita F. Fahy, Guylène Proulx, and Lata Aiman. "Panic or not in fire: Clarifying them is conception". In: *Fire and Materials* 36 (2012), pp. 328–338. DOI: 10.1002/fam.1083.

[9] Chris Cocking, John Drury, and Steve Reicher. "The psychology of crowd behaviour in emergency evacuations: Results from two interview studies and implications for the Fire and Rescue Services". In: *Irish Journal of Psychology* 30 (2009), pp. 59–73. DOI: 10.1080/03033910.2009.10446298.

[10] Gustave Le Bon. *The crowd: A study of the popular mind.* New York: Macmillan Company, 1896.

[11] Alexander Mintz. "Non-Adaptive Group Behavior". In: *The journal of abnormal and social psychology* 46 (1951), pp. 150–159.

[12] Enrico L Quarantelli. "Images of Withdrawal Behavior in Disasters: Some Basic Misconceptions". In: *Social Problems* 8 (1960), pp. 68–79. URL: https://www.jstor.org/stable/798631.

[13] Erik Auf der Heide. "Common Misconceptions about Disasters: Panic, the "Disaster Syndrome," and Looting". In: *The First 72 Hours: A Community Approach to Disaster Preparedness.* Ed. by M. O'Leary. Lincoln: iUniverse Publishing, 2004, pp. 340–380.

[14] Chris Cocking and John Drury. "Talking about Hillsborough: 'Panic' as Discourse in Survivors' Accounts of the 1989 Football Stadium Disaster". In: *Journal of Community & Applied Social Psychology* 24 (2014), pp. 86–99. DOI: 10.1002/casp.2153.

[15] Enrico L. Quarantelli. "Panic, Sociology of". In: *International encyclopedia of the social and behavioral sciences.* Ed. by N. J. Smelser and P. B. Baltes. New York: Pergamon Press, 2001, pp. 11020–11023.

[16] Jonathan D Sime. "The concept of panic". In: *Fires and human behaviour.* Ed. by D. Canter. London: David Fulton, 1990, pp. 63–81.

[17] Dirk Helbing and Pratik Mukerji. "Crowd disasters as systemic failures: Analysis of the love parade disaster". In: *EPJ Data Science* 1 (2012), pp. 1–40. DOI: 10.1140/epjds7.

[18] John Drury, Chris Cocking, and Steve Reicher. "Everyone for themselves? A comparative study of crowd solidarity among emergency survivors". In: *British Journal of Social Psychology* 48 (2009), pp. 487–506. DOI: 10.1348/014466608X357893.

[19] John Drury, Chris Cocking, and Steve Reicher. "The Nature of Collective Resilience: Survivor Reactions to the 2005 London Bombings". In: *International Journal of Mass Emergencies and Disasters* 27 (2009), pp. 66–95.

[20] Norris R. Johnson. "Panic at "The Who Concert Stampede": An Empirical Assessment". In: *Social Problems* 34 (1987), pp. 362–373. DOI: 10.2307/800813.

[21] James R. Gill and Kristen Landi. "Traumatic asphyxial deaths due to an uncontrolled crowd". In: *American Journal of Forensic Medicine and Pathology* 25 (2004), pp. 358–361. DOI: 10.1097/01.PAF.0000147316.62883.8B.

[22] Maria Moitinho de Almeida and Johan von Schreeb. "Human Stampedes: An Updated Review of Current Literature". In: *Prehospital and Disaster Medicine* 34 (2019), pp. 82–88. DOI: 10.1017/S1049023X18001073.

[23] Luis-Manuel Garcia. "Pathological Crowds: Affect and Danger in Responses to the Love Parade Disaster at Duisburg". In: *Dancecult: Journal of Electronic Dance Music Culture* 2 (2011), pp. 6–11. DOI: 10.12801/1947-5403.2011.02.01.15.

[24] "German court halts Love Parade trial with no convictions". In: *DW News* (2020-05-04). URL: https://p.dw.com/p/3bkSi (visited on 12/08/2021).

[25] John Drury. "The role of social identity processes in mass emergency behaviour: An integrative review". In: *European Review of Social Psychology* 29 (2018), pp. 38–81. DOI: 10.1080/10463283.2018.1471948.

[26] Tatsuya Nogami. "Disaster myths among disaster response professionals and the source of such misconceptions". In: *Journal of Contingencies and Crisis Management* 26 (2018), pp. 491–498. DOI: 10.1111/1468-5973.12218.

[27] John Drury, David Novelli, and Clifford Stott. "Representing crowd behaviour in emergency planning guidance: 'mass panic' or collective resilience?" In: *Resilience* 1 (2013), pp. 18–37. DOI: 10.1080/21693293.2013.765740.

[28] Guylene Proulx and Jonathan D Sime. "To Prevent 'Panic' in an Underground Emergency: Why Not Tell People the Truth?" In: *Fire Safety Science: Third International Symposium.* Ed. by Geoffrey Cox and Brian Langford. London: Elsevier Applied Science, 1991, pp. 843–852.

[29] Ed Galea. "It's a crowd crush NOT A STAMPEDE - the difference is important!" In: *LinkedIn* (2021-04-30). URL: https://www.linkedin.com/pulse/its-crowd-crush-stampede-difference-important-ed-galea/ (visited on 12/07/2021).

[30] Gianna-Carina Grün. "Was sind crowd dynamics? - Forschungsfrage". In: *jetzt* (2011-07-22). URL: https://www.jetzt.de/forschungsfrage/was-sind-crowd-dynamics-528031 (visited on 12/07/2021).

[31] Micha Hörnle. "Letzte "Nachtschicht" in Heidelberg: "Massenpanik" ist ein Mythos". In: *Rhein-Neckar-Zeitung* (2018-01-02). URL: https://www.rnz.de/nachrichten/heidelberg_artikel,-letzte-nachtschicht-in-heidelberg-massenpanik-ist-ein-mythos-plus-fotogalerie-_arid,327609.html (visited on 12/07/2021).

[32] Tatsuya Nogami. "Who Panics and When: A Commonly Accepted Image of Disaster Panic in Japan". In: *International Perspectives in Psychology* 5 (2016), pp. 245–254. DOI: 10.1037/ipp0000050.

[33] John Drury, David Novelli, and Clifford Stott. "Managing to avert disaster: Explaining collective resilience at an outdoor music event". In: *European Journal of Social Psychology* 45 (2015), pp. 533–547. DOI: 10.1002/ejsp.2108.

[34] Laura Künzer, Robert Zinke, and Gesine Hofinger. "Mythen der Entfluchtung". In: *Tagungsband Jahresfachtagung*. Ed. by Vereinigung zur Förderung des Deutschen Brandschutzes eV. 2012, pp. 725–735.

[35] Anthony R. Mawson. "Understanding mass panic and other collective responses to threat and disaster". In: *Psychiatry* 68 (2005), pp. 95–113. DOI: 10.1521/psyc.2005.68.2.95.

[36] Tatsuya Nogami. "Negative misconceptions about disaster behaviour through availability cascades: An examination of secondhand information and the moderating effect of trait anxiety on disaster myths". In: *Journal of Community & Applied Social Psychology* 30 (2020), pp. 369–380. DOI: 10.1002/casp.2441.

[37] Tatsuya Nogami and Fujio Yoshida. "Disaster myths after the Great East Japan Disaster and the effects of information sources on belief in such myths". In: *Disasters* 38 (2014), pp. 190–205. DOI: 10.1111/disa.12073.

[38] Louis Guttman. "Some necessary conditions for common-factor analysis". In: *Psychometrika* 19 (1954), pp. 149–161. URL: https://link.springer.com/content/pdf/10.1007/BF02289162.pdf.

[39] Henry F. Kaiser. "The Application of Electronic Computers to Factor Analysis". In: *Educational and Psychological Measurement* 20 (1960), pp. 141–151. DOI: 10.1177/001316446002000116.

[40] Philipp Mayring. "Qualitative Content Analysis [28 paragraphs]. Forum Qual. Sozialforsch". In: *Forum Qual. Soc. Res. 1, Art. 20* (2000). DOI: 10.17169/fqs-1.2.1089.

[41] John Drury. "Collective resilience in mass emergencies and disasters: A social identity model". In: *The Social Cure: Identity, Health and Well-Being*. Ed. by J. Jetten, C. Haslam, and S. A. Haslam. Hove, UK: Psychology Press, 2012, pp. 195–216. DOI: 10.4324/9780203813195.

[42] Daniel Brunsch. "Publikumszentrierte Sicherheits- & Krisenkommunikation auf Veranstaltungen". In: *Eventpsychologie*. Ed. by Steffen Ronft. Wiesbaden: Springer, 2021, pp. 569–604. DOI: 10.1007/978-3-658-28888-4.

[43] Laura Künzer and Gesine Hofinger. "Eine Einführung aus Sicht der Human Factors-Psychologie". In: *Eventpsychologie,* ed. by S. Ronft. Wiesbaden: Springer, 2021, pp. 515–542. DOI: 10.1007/978-3-658-28888-4_23.

Band / Volume 52
JSC Guest Student Programme Proceedings 2021
I. Kabadshow (Ed.) (2023), ii, 82 pp
ISBN: 978-3-95806-684-7

Band / Volume 53
Applications of variational methods for quantum computers
M. S. Jattana (2023), vii, 160 pp
ISBN: 978-3-95806-700-4

Band / Volume 54
Crowd Management at Train Stations in Case of Large-Scale Emergency Events
A. L. Braun (2023), vii, 120 pp
ISBN: 978-3-95806-706-6

Band / Volume 55
Gradient-Free Optimization of Artificial and Biological Networks using Learning to Learn
A. Yeğenoğlu (2023), II, 136 pp
ISBN: 978-3-95806-719-6

Band / Volume 56
Real-time simulations of transmon systems with time-dependent Hamiltonian models
H. A. Lagemann (2023), iii, 166, XXX pp
ISBN: 978-3-95806-720-2

Band / Volume 57
Plasma Breakdown and Runaway Modelling in ITER-scale Tokamaks
J. Chew (2023), xv, 172 pp
ISBN: 978-3-95806-730-1

Band / Volume 58
Space Usage and Waiting Pedestrians at Train Station Platforms
M. Küpper (2023), ix, 95 pp
ISBN: 978-3-95806-733-2

Band / Volume 59
Quantum annealing and its variants: Application to quadratic unconstrained binary optimization
V. Mehta (2024), iii, 152 pp
ISBN: 978-3-95806-755-4

Schriften des Forschungszentrums Jülich
IAS Series

Band / Volume 60
Elements for modeling pedestrian movement
from theory to application and back
M. Chraibi (2024), vi, 279 pp
ISBN: 978-3-95806-757-8

Band / Volume 61
Artificial Intelligence Framework for Video Analytics:
Detecting Pushing in Crowds
A. Alia (2024), xviii, 151 pp
ISBN: 978-3-95806-763-9

Band / Volume 62
The Relationship between Pedestrian Density, Walking Speed
and Psychological Stress:
Examining Physiological Arousal in Crowded Situations
M. Beermann (2024), xi, 117 pp
ISBN: 978-3-95806-764-6

Band / Volume 63
Eventify Meets Heterogeneity:
Enabling Fine-Grained Task-Parallelism on GPUs
L. Morgenstern (2024), xv, 110 pp
ISBN: 978-3-95806-765-3

Band / Volume 64
Dynamic Motivation in Crowds: Insights from Experiments
and Pedestrian Models for Goal-Directed Motion
E. Üsten (2024), ix, 121 pp
ISBN: 978-3-95806-773-8

Band / Volume 65
Propagation of Stimuli in Crowds:
Empirical insights into mutual influence in human crowds
H. Lügering (2024), xi, 123 pp
ISBN: 978-3-95806-775-2

Weitere ***Schriften des Verlags im Forschungszentrum Jülich*** unter
http://wwwzb1.fz-juelich.de/verlagextern1/index.asp

IAS Series
Band / Volume 65
ISBN 978-3-95806-775-2