NBER WORKING PAPER SERIES

SHIFTING PATTERNS OF SOCIAL INTERACTION: EXPLORING THE SOCIAL LIFE OF URBAN SPACES THROUGH A.I.

Arianna Salazar-Miranda
Zhuangyuan Fan
Michael B. Baick
Keith N. Hampton
Fabio Duarte
Becky P.Y. Loo
Edward L. Glaeser
Carlo Ratti

Working Paper 33185
http://www.nber.org/papers/w33185

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2024

The authors thank the AMS Institute, Dubai Future Foundation, Toyota, UnipolTech, Consiglio per la Ricerca in Agricoltura e l'Analisi dell'Economia Agraria, Tele2, Volkswagen Group America, FAE Technology, Fédération Internationale de l'Automobile, Rio de Janeiro, Emilia-Romagna, Stockholm, Amsterdam, Kyoto University, SMART - Singapore-MIT Alliance for Research and Technology, KAIST Center for Advanced Urban Systems, the Project for Public Spaces, and the members of the MIT Senseable City Lab Consortium for supporting this research. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w33185

Shifting Patterns of Social Interaction: Exploring the Social Life of Urban Spaces Through A.I.
Arianna Salazar-Miranda, Zhuangyuan Fan, Michael B. Baick, Keith N. Hampton, Fabio Duarte, Becky P.Y. Loo, Edward L. Glaeser, and Carlo Ratti
NBER Working Paper No. 33185
November 2024
JEL No. R41, Z13, Z30

ABSTRACT

We analyze changes in pedestrian behavior over a 30-year period in four urban public spaces located in New York, Boston, and Philadelphia. Building on William Whyte's observational work from 1980, where he manually recorded pedestrian behaviors, we employ computer vision and deep learning techniques to examine video footage from 1979-80 and 2008-10. Our analysis measures changes in walking speed, lingering behavior, group sizes, and group formation. We find that the average walking speed has increased by 15%, while the time spent lingering in these spaces has halved across all locations. Although the percentage of pedestrians walking alone remained relatively stable (from 67% to 68%), the frequency of group encounters declined, indicating fewer interactions in public spaces. This shift suggests that urban residents increasingly view streets as thoroughfares rather than as social spaces, which has important implications for the role of public spaces in fostering social engagement.

Arianna Salazar-Miranda
School of the Environment
Yale University
195 Prospect St
New Haven, Conn 06511
arianna.salazarmiranda@yale.edu

Zhuangyuan Fan
The University of Hong Kong
yuanzf@connect.hku.hk

Michael B. Baick
Carlo Ratti Associati
491 Laurel St
Longmeadow, MA 01106
michaelbaick1@gmail.com

Keith N. Hampton
Michigan State University
404 Wilson Rd
Room 409
East Lansing, MI 48824
akhampton@msu.edu

Fabio Duarte
77 Massachusetts Avenue
Cambridge, MA 02139
United States
fduarte@mit.edu

Becky P.Y. Loo
The University of Hong Kong
Rm 1034, 10/F Jockey Club Tower, Geography
Pokfulam Road
Hong Kong
bpyloo@hku.hk

Edward L. Glaeser
Department of Economics
315A Littauer Center
Harvard University
Cambridge, MA 02138
and NBER
eglaeser@harvard.edu

Carlo Ratti Senseable
City Lab
ratti@mit.edu

A data appendix is available at www.nber.org/data-appendix/w33185

1 Introduction

Do city sidewalks serve primarily for pedestrian mobility, or do they also serve as social spaces to connect people? Has their function changed over time, either because increasing urban wages have raised the opportunity cost of lingering on the street or because the demand for social interaction has gone up? In this paper, we observe pedestrian behaviors in four urban public spaces, analyzing videos taken 30 years apart. Using these data, we measure changes in walking speed and social interactions.

We use computer vision and deep learning to analyze archival video footage collected by William Whyte from 1979-80 (hereafter called 1980) and contemporary footage from the 2008-10 (hereafter called 2010), focusing on public spaces in New York, Boston, and Philadelphia. This approach builds on the visual observation traditions of urbanists like Lynch (1960), Jacobs (1961), and Gehl (2011), and Whyte (1980) [1, 2, 3, 4]. The pioneering observational work of Whyte (1980), who relied on time-lapse videos, is the direct intellectual ancestor of this paper. While we look at the same spaces and even the same videos as Whyte (1980), our analysis is essentially automated, which reduces the labor-intensive and subjective nature of manual video analysis.

In our application, we document the evolution of social dynamics in two public spaces in New York (outside of Bryant Park and the Metropolitan Museum), Boston (Downtown Crossing), and Philadelphia (Chestnut Street). These sites were initially analyzed by Whyte in 1980, and we had to use these locations in order to anchor our observations of changes over 30 years. We use these films to measure changes in group behaviors, focusing on walking speed, lingering behavior, group sizes, and group formation.

To calculate walking speed and lingering behavior, we first measure the speed for everyone in the film. We then define individuals who are moving at less than 0.5 meters per second as lingerers. Everyone else is defined as a walker. We find that in all four locations, the share of lingerers decreased by 14 percent, and walking speeds increased by 15 percent, which suggests that urbanites are using streets more as walkways than as social spaces.

We do not, however, find any change in the share of pedestrians who are traveling alone. In 1980, 67 percent of pedestrians were walking by themselves, and this number only changed to 68 percent in 2010. That seeming stasis, however, masks two significant changes in individual sites. Bryant Park experienced an increase in the share of dyads walking (15 percentage points), likely due to its development as a more prominent destination. Conversely, Downtown Crossing showed a decline in dyads (10 percentage points), possibly reflecting its more utilitarian role focused on shopping and quick errands rather than social or leisurely activities.

Finally, we look at changes in group formation. Do people run into people on the street and then interact with them? We find a substantial decrease in the amount of group formation, meaning that there is less spontaneous interaction and/or less use of urban spaces as prearranged meeting spots.

We find that the nature of interactions has shifted significantly over time. People are now spending less time in public spaces and moving through them at a faster pace. These shifts reflect a change in how public spaces are used, potentially affecting their role in facilitating social engagement and everyday social connections.

2 Sidewalks, Transportation and Entertainment

City streets have long been places of movement, entertainment, and social connection. People watching, window shopping, and running into friends have long been pleasures created by a stroll through a dense urban area. But what determines the balance between simply passing through and lingering in space, and what would cause that balance to shift over time?

The function of the sidewalk is shaped by both users' demand for speed and the inherent interest in the space. In our spaces, the Metropolitan Museum of Art (Met), in New York, stands out as a particularly unusual space, one which typically attracts large numbers of tourists who are typically not in a hurry for whom the space is particularly novel. We expect to see more lingering and lower speeds at the Met than at our other locales.

Downtown Crossing in Boston, is both a busy transportation hub, where people shift from walking to public transit and an area that is well endowed with retail shops with somewhat interesting windows (precisely because it has so many pedestrians in transit). Consequently, it is unclear whether lingering increased or decreased in Downtown Crossing between 1980 and 2010.

Chestnut Street in Philadelphia serves as an "in-between" space, functioning primarily as a pedestrian transit corridor between destinations. We anticipate higher walking speeds and lower lingering in this location, similar to the transit-focused aspect of Downtown Crossing.

One plausible explanation for changes over time is that significant increases in urban incomes have led to a greater demand for speed because time has become more valuable. According to County Business Patterns data for 2010, average annual earnings per employee were over $70,000$ in the zip codes associated with both Downtown Crossing and Bryant Park, suggesting these areas might exhibit particularly fast walking speeds. In contrast, per-employee earnings were closer to $50,000$ in the other two zip codes associated with Chestnut Street and the Metropolitan Museum of Art, possibly indicating lower walking speeds and more lingering in these areas.

An alternative hypothesis is that an increasingly prosperous and safer urban environment has enhanced the appeal of public spaces, making them more engaging and attractive for pedestrians. Of our four areas, Bryant Park stands out as a location that has seen a marked improvement in safety over the 30-year period, potentially contributing to changes in how it is used [5]. Chestnut Street has also gentrified significantly over the 30-year period [6].

3 Analyzing Video Footage of Public Spaces

Our dataset combines archival and contemporary video footage. For the archival component, we sourced 1980 videos from the Project for Public Spaces (PPS). This non-profit organization was founded by Fred Kent, who worked closely with the renowned urbanist William "Holly" Whyte[4]. This footage, consisting of Super 8 films captured by Whyte and his students from PPS, was the basis of Whyte's pioneering work analyzing cities with timelapse videos. They were made available to us in a digital form thanks to the work done by Keith Hampton and his students.

Beginning in 2008, Hampton digitized parts of the Whyte archive and took new footage of the same sites during similar times of the day to guarantee comparability[7]. The archival footage was taken from rooftops and building windows. The 2010 footage was taken from a lower altitude but still at an elevated view. Figure 1 shows a sample of the 1980 and 2010 videos for the four sites in our analysis and highlights the location of these public spaces.

Summary statistics describing the duration, time of day, and date for these videos are in Table 1. Our analysis draws on 638 total minutes of footage. We only included frames where the camera was stationary and unobstructed. To ensure comparability across videos, we normalized all results to the frame level since the length of individual videos varied from 30 minutes to 3 hours. Most of the videos were filmed at midday, except for Chestnut Street, which was captured in the morning.

3.1 Measuring group behaviors using computer vision and deep learning

We combined computer vision and deep learning methods to detect and analyze group behaviors. Our methodology consists of three components: 1) pedestrian tracking and classification, 2) group detection, and 3) identification of encounters.

Pedestrian tracking To analyze the movement of pedestrians, we track individuals' trajectories across frames. We first detect pedestrians using a Convolutional Neural Network

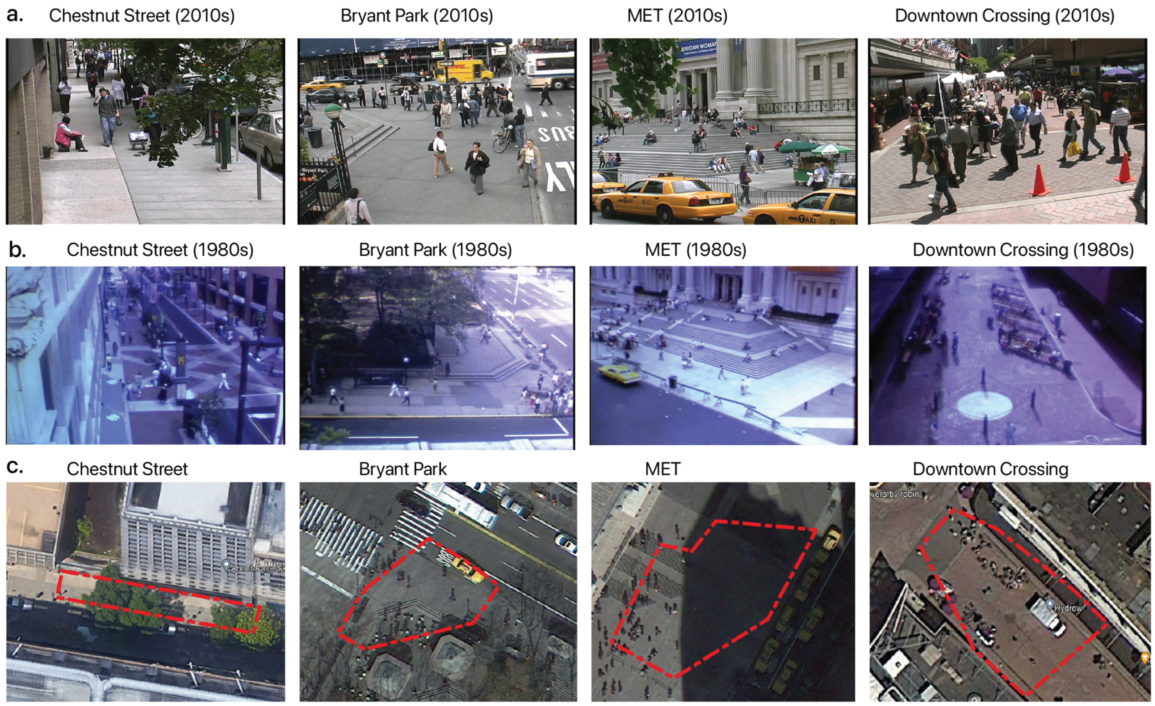

FIGURE 1: HISTORICAL AND CONTEMPORARY VIDEOS FOR THE FOUR PUBLIC SPACES. a frame extracted from videos in 2010. b frames extracted from videos in 1980. c location of the study sites and their corresponding boundary. Boundaries are defined using the common field of view in both the archival and the contemporary videos in each location.

TABLE 1: Summary statistics of the selected sites

	(I) Bryant Park		(II) Chestnut Street		(III) Downtown Crossing		(IV) MET	
City	New York City		Philadelphia		Boston		New York City	
Type	Crosswalk		Street		Street plaza		Grand stairs	
Date	1980-4-10	2008-12-05	1979-8-30	2010-05-19	1979-9-28	2010-05-21	1979-6-2	2010-08-08
Time of day	13:17:00	14:19:44	10:30	8:33:43	9:50	11:57:55	12:00	12:01:18
Total duration (mins)	101	77	31	202	43	63	85	36

Table reports the summary statistics of the video footage for all locations (column 1), Bryant Park (column 2), Downtown Crossing (column 3), and the MET (column 4). Bryant Park videos were recorded at the corner of 45th Street and 6th Avenue. Chestnut Street videos were recorded between 10th and 11th Streets. Downtown Crossing videos were recorded next to Macy's. MET videos correspond to the main stairs of the Metropolitan Museum of Art. The reported video length here only contains the time when the camera is stable and no obstruction in front of the camera.

(CNN)-based object detection model called YOLO (version 5) [8]. We then track individuals across frames using a Deep-SORT framework [9]. This allows us to trace pedestrians' trajectories and compute their speed. The YOLO model detects pedestrians in each frame, while Deep-SORT maintains the continuity of pedestrians across frames. This combination has also been previously used for pedestrian tracking [10, 11, 12].

YOLO is trained using the COCO 2017 dataset [13], which contains over 1.5 million classified object instances. The pre-trained YOLO model achieves a state-of-art mean average precision of 0.68 at IoU=0.5 (mAP@0.5) for pedestrian identification in the 2010 videos. But it was less accurate when analyzing 1980 videos with lower resolutions. To improve the model's accuracy, we fine-tuned it with 260 labeled images (with 4,792 labeled pedestrians) from the 1980 videos. This process improved the model's performance, increasing its mAP@0.5 from 0.435 to 0.781, precision from 0.673 to 0.902, and recall from 0.84 to 0.93 in the test data (See section 7.1 for details).

Our study distinguishes between pedestrians who are in transit and those who are stationary. We classify pedestrians in these categories based on their speed. A pedestrian is labeled as "lingering" if their movement speed falls below 0.5 m/s for more than a 5-seconds (results are also reported for different time thresholds in Figure S5, Panel b). Our choice of the 0.5 m/s speed threshold to mean stationary is based on a mean comfortable gait speed ranging from 1.2 m/s to 1.47 m/s[14, 15]. This choice is also supported by the distribution of speed in our dataset, which is bimodal and has a clear mass of people (the people lingering) with speeds below 0.5 m/s. (see Figure S5, Panel a in the Supplemental Materials for speed distribution). This definition differs from previous work by Hampton et al. [7], who manually coded people as lingerers if they are in the same place in two consecutive frames separated by 15 seconds. Their more stringent definition results in a lower share of lingerers.

Group detection Our group detection algorithm is executed in three steps. First, we employ the DBSCAN[16] algorithm for every frame to cluster individuals who are in close proximity. Second, we calculate the vector representing the trajectory of each individual, comparing their movements between the current frame and the preceding one. The final step classifies individuals into groups. Two individuals are considered part of a group if they are positioned no more than 1.9 meters apart for more than 2 seconds and (I) their trajectory vectors are positively correlated at the time, indicating similar movement directions; or (II) both maintain speeds below 0.5 meters per second, suggesting that they are stationary.

The selection of the 1.9-meter distance threshold is based on Edward T. Hall's sociological theory on personal distance. Hall argued that individuals within 1.2 meters are in their "personal spaces" [17]. The 1.9-meter threshold we use accounts for the fact that the width of each person's shoulders is typically 35 cm.

Our method assigns a unique identifier to each pedestrian within each time frame and site when they are classified as part of a group. This approach also tracks changes in group membership as pedestrians join, leave, or stay within the group over the duration of the video.

To validate the group detection methodology, we randomly selected 80 frames (with over 3000 pedestrians) from each location from 1980 and 2010 (See Supplemental Table S4 for details). A researcher reviewed each frame to determine whether individuals were alone or in groups. The performance of our approach is in line with results reported in previous studies [18, 19], with an overall accuracy rate (compared to the codings made by the researcher) of 87.91%.

Identifying group encounters Our analysis of group formation distinguishes between pre-formed groups (i.e., those established prior to entering the public site) and group encounters, which may be either spontaneous encounters or preplanned interactions (i.e., groups that form during the observation period as people meet through the frames). We classify a group encounter as occurring if its members entered the site alone at least 5 seconds before they formed the group. The 5-second threshold accounts for the possibility that people in a pre-formed group may enter the video at either a slightly different time. Figure S4 provides a visual representation of this classification.

4 Results

This section describes the three main results from our analysis.

Result 1: people linger less in the public spaces analyzed and walk through these sites faster than in 1980. Figure 2, Panel a reports the share of people lingering per frame for all four sites. We estimate that 43% of the people in these public sites were classified as lingering in 1980 based on their temporal speed. This share has since declined to about 26%, indicating that people are lingering less. Figure S5, Panel b in the Supplemental Materials shows that using a higher time threshold to define lingering (above 5 seconds at a speed below < 0.5 m/s) results in a lower number of lingering pedestrians but continues to show a decline in their share over time.

The decline in people's propensity to linger in public spaces is sizable and visible across all locations analyzed. For example, 54% of people in Downtown Crossing in 1980 lingered there, while only 14% lingered in 2010.

The decline in people's propensity to linger is also visible for singles (see Figure 2, Panel b), which shows that the observed decline is not simply a function of changes in group composition over time. Single pedestrians lingered 39% of the time in 1980 (as compared to a lingering rate of 54% for all people). Their lingering rate declined to 24% in the recent 2010 videos, which is a similar decline to that seen for all groups.

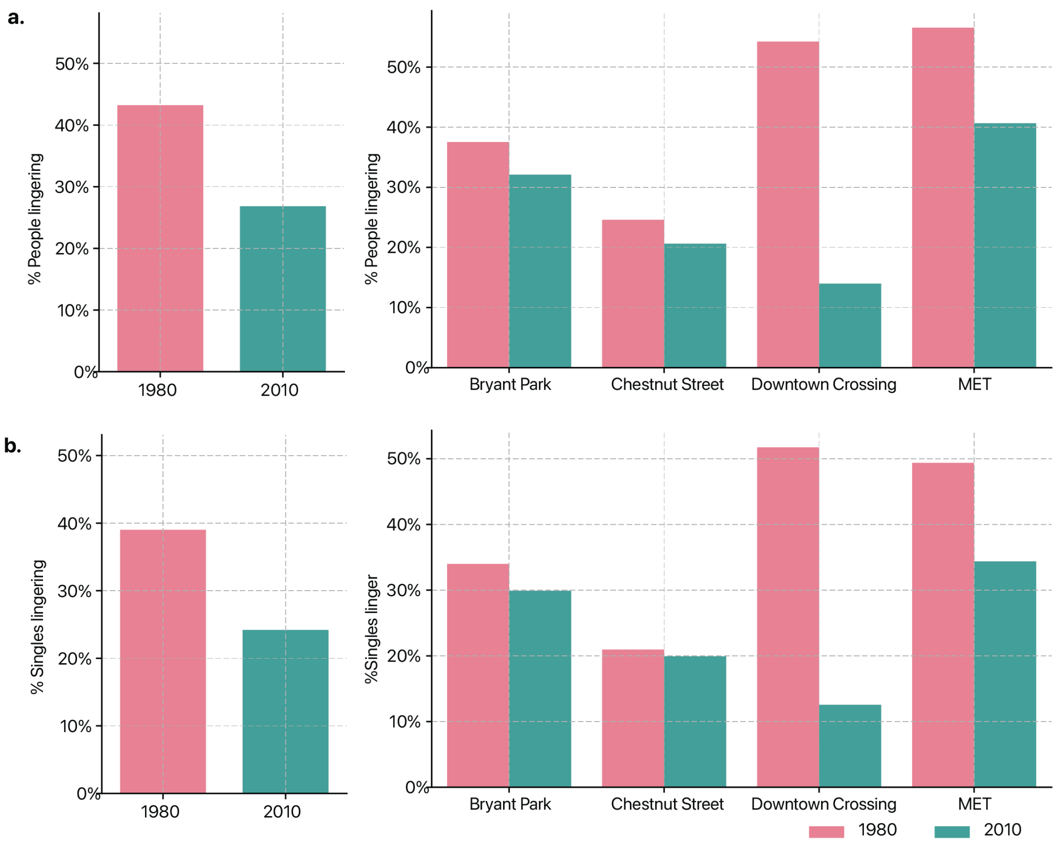

FIGURE 2: LINGERING BEHAVIOR IN PUBLIC SPACES. a percentage of people lingering per frame for all four sites (left) and broken down by location (right). b percentage of singles lingering per frame for all four sites (left) and by location (right). See Supplemental Materials Table S6 for more details.

The decline in lingering time was accompanied by an increase in walking speed among those who do not linger. Figure 3, Panel a reports statistics for average and median speed in 1980 and 2010, pooling all sites. These statistics are computed only for people who are classified as movers (moving at speeds above 0.5 m/s). In 1980, movers traversed the public sites in our sample at an average speed of 1.25 m/s (median speed of 1.22 m/s). These speed levels are comparable to those reported in the Nationwide Personal Transportation Survey from 1977. The survey asked respondents to report the distance and time spent on recent walking trips. There is a total of 80 trips reported for Boston, NY, and Philadelphia of more than 5 minutes and where people traveled 2 kilometers or less, which should be roughly comparable to the trips in our video sample. The average speed (computed as distance over time) for these trips is 1.34 m/s (median 1.24 m/s).

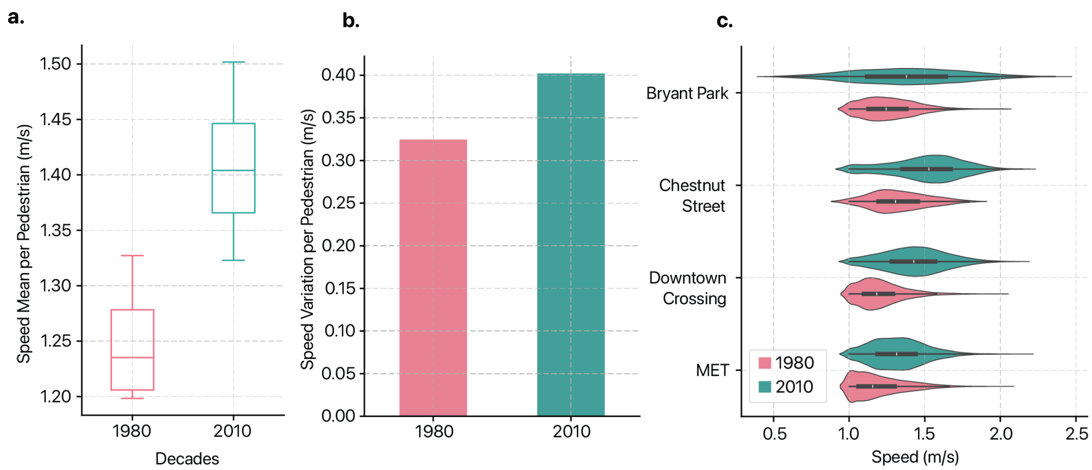

FIGURE 3: SPEED IN PUBLIC SPACES. a walking speed of pedestrians for all sites in 1980 and 2010. b variance of speed across people for all sites in 1980 and 2010. c walking speed variance across people by location.

More importantly, comparing these historical numbers to the ones estimated from recent videos on the same sites, we find that, by 2010, speed increased by 15% to 1.41 m/s (median speed 1.41 m/s). This increase in speed was also accompanied by an increase in the variance of speed across people (see Panel b), with the standard deviation in speed increasing from 0.32 in 1980 to 0.4 m/s in 2010. This change is visible across all the public sites in our sample (see Panel c). Figure 3, Panel c also shows that the increase in walking speed and its dispersion rose in all four public sites by a similar amount. Among the four sites, pedestrians on Chestnut Street had the highest average walking speed in both 1980 and 2010 (1.33 m/s and 1.50 m/s, respectively). On the contrary, the Met has seen the lowest average walking speed in both years (1.20 m/s and 1.32 m/s).

Result 2: public spaces are dominated by single pedestrians. The average share of single pedestrians across the sites studied was 67% in the historical videos from 1980 and remained high, at 68%, in the recent videos from 2010.

This pattern is visible across all sites analyzed and is shown in Figure 4, Panel a. In

Chestnut Street, for example, the share of single pedestrians remained high at 79% both in the 1980 and 2010 videos (see Panel b). Even in the Met, we estimate that 48% of pedestrians are singles, with this share remaining constant in time. To ensure that our results are not driven by our choice of time threshold, we tested a range of thresholds from 0.1 to 10 seconds to define groups. Table S3 and Figure S3 show that singles continue to make up the majority across all sites, though at the Met, group proportions slightly increase with thresholds below 2 seconds.

Most of the remaining people in public spaces are in groups of two (dyads), accounting for about 27.7% in the 1980s and 26.7% in the 2010s, while a slim minority forms triads (3.53% in the 1980s, 3.49% in the 2010s) or larger groups (1.40% in the 1980s, 1.82% in the 2010s). The high prevalence of single-group pedestrians is in line with previous findings by Hampton et al. [7]. These patterns are further summarized in Table S2, which reports the share of people in groups of different sizes separately for each site.

Result 3: decline in group encounters We now ask if groups are actually forming on these sidewalks. These groups may have either been pre-planned meetings or spontaneous encounters.

One critical challenge in measuring group formation is to separate people who were together but entered the film at slightly different times from people who genuinely met in this space. We report the share of people that form at least one group s seconds after they enter the location and vary the value of s to allow for more stringent methods of measuring true meetings rather than straggling friends.

To put different values of s in perspective, Panel a in Figure 5 plots the share of people who exit a public site (defined here as exiting the video frame) s seconds after they enter (with their entry defined by the first time they show on a frame), which we define as the churn rate. In 1980, 80% of the people entering a site exited within 20 seconds of their arrival. In 2010, this share went up to 92%. This increase in the churn rate is the natural consequence of walking faster. However, changing churn rates over time means that it is particularly important to hold the time delay constant for defining group formation.

Figure 5, Panel b reports the share of people that form at least one group s seconds after they enter the location—setting higher values for s guarantees that the groups were formed in situ and that the individuals did not arrive together at slightly different times. Using a baseline value of $s = 5$ seconds, we estimate that in 1980, 5.5% of people (one in every 20 people) entering one of our public spaces formed a group. By 2010, this percentage declined to 2% (one in every 50 people). The same decline is visible when using different thresholds

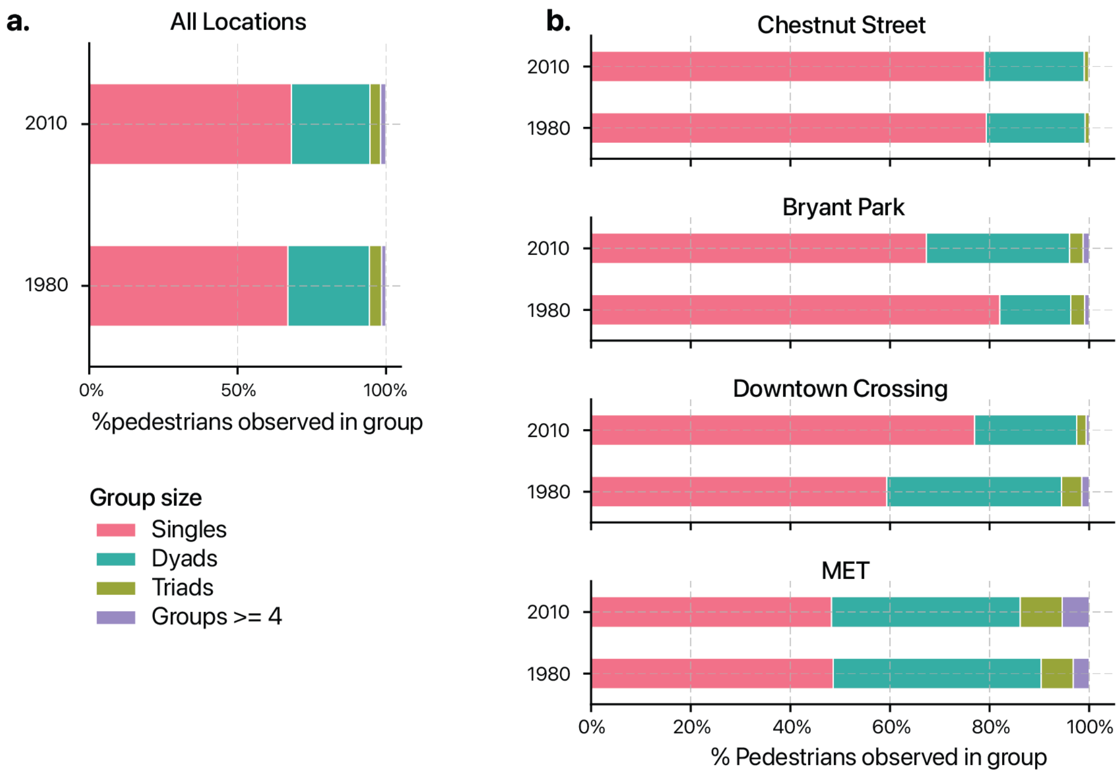

FIGURE 4: GROUP DYNAMICS OVER TIME IN PUBLIC SPACES. a percentage of different group sizes (singles, dyads, triads, and groups of four or more) observed across all locations in 1980 and 2010. See Supplemental Materials Table S2 for details. b. percentage of pedestrians by group size for each location in 1980 and 2010. We also repeat the analysis using different time thresholds to define groups. The full result are reported in Table S3.

s to define group encounters.

Figure 5, Panel c shows this decline is pervasive across all locations. The decline is particularly pronounced in the Met (from 12.5% to 7%) and in Downtown Crossing (from 8% to 1%). While this trend reflects broad changes in how these spaces are used, our data cannot distinguish whether the decline is due to fewer pre-arranged meetings or reduced spontaneous interactions.

5 CONCLUSION

This paper shows how AI methods can analyze videos to study group dynamics in public spaces. This approach builds on existing visual observation traditions in urban studies and

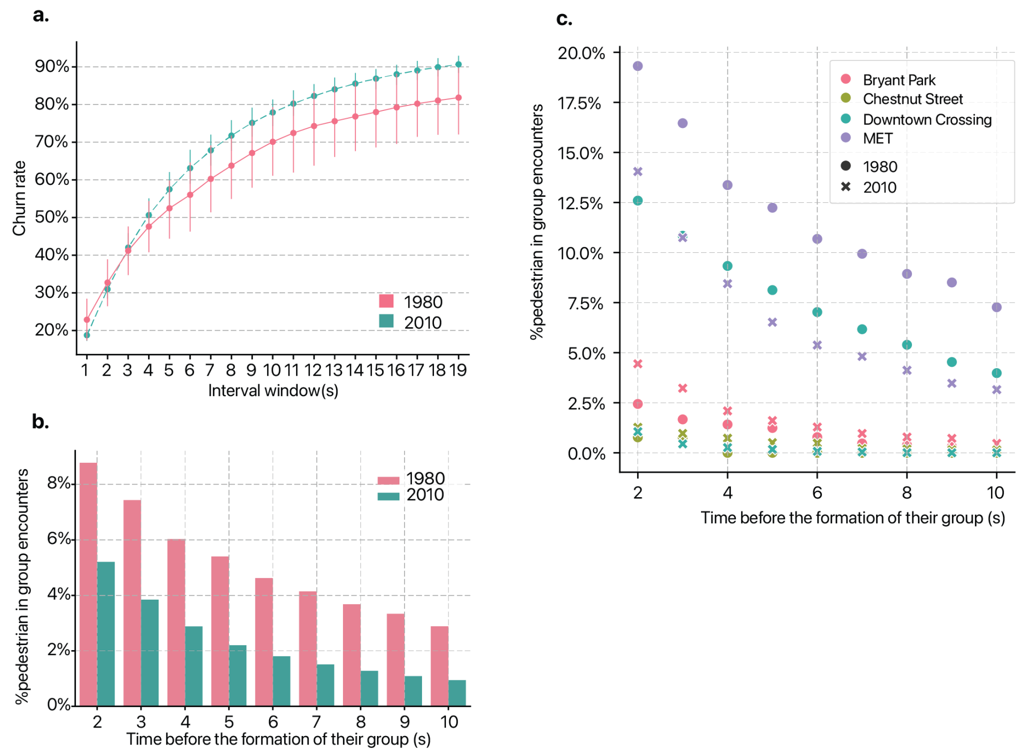

Figure 5: Churn and group encounters in public spaces. a churn rate, defined as the share of people who exit a public site, measured at different time intervals. b share of group encounters, measured at different seconds after they enter the location. c share of group encounters by location.

overcomes the challenges of manual analysis, such as labor intensity and inter-observer reliability. Our approach facilitates differentiation between various group sizes and identification of "lingering" behavior and spontaneous meetings. We apply this method to study four important public sites: Chestnut Street in Philadelphia, Downtown Crossing in Boston, Bryant Park in New York City, and the Metropolitan Museum of Art in New York City.

Our findings reveal several trends in these locations. First, public spaces appear dominated by single individuals and not so much by groups. Second, the nature of interactions has shifted significantly over time: People are now spending less time in public spaces and moving through them at a faster pace. Group formation has declined significantly, suggesting that public spaces are becoming less conducive to planned and unplanned social encounters.

These patterns are consistent with the idea that people balance their need for speed with the desire for engaging experiences in urban environments. While higher wages might be associated with increased walking speeds and reduced social interactions in so many settings[20], the Metropolitan Museum stands out as an exception, retaining high levels of lingering despite being located in a high-income neighborhood. This indicates that when urban amenities are sufficiently compelling, they can still encourage slower social behavior.

The number of locations we surveyed is relatively limited, but this was intentional to ensure comparability with William Whyte's original 1980 study. While this constraint affects the generalizability of our findings, it enables a focused analysis of changes over 30 years in the same sites. The changes in pedestrian behavior we observe could be due to a variety of factors, including shifts in human behavior, land use, or the character of the space. Additionally, our analysis captures the co-presence of different group sizes and frequencies but does not address the quality of these interactions. For instance, our encounter measure cannot distinguish between spontaneous meetings and scheduled ones, friendly encounters versus confrontations, or strangers versus acquaintances. Some factors, such as gender and mobile phone use, remain challenging even for skilled human coders to identify [7]. Future research could expand this work with larger datasets and more nuanced categories of social interactions, potentially incorporating data from GPS, Bluetooth, and visual imagery.

References

[1] Kevin Lynch. *The image of the city.* MIT Press, 1960.

[2] Jane Jacobs. *The death and life of great American cities.* Random House, Inc., New York, 1961.

[3] Jan Gehl. *Life between buildings : using public space.* Island Press, 2011.

[4] William Hollingsworth Whyte et al. The social life of small urban spaces. 1980.

[5] Project for Public Spaces. Bryant park, 2024. Accessed: November 5, 2024.

[6] Pew Charitable Trusts. Philadelphia's changing neighborhoods: Gentrification and other shifts since 2000, 2016. Accessed: November 5, 2024.

[7] Keith N Hampton, Lauren Sessions Goulet, and Garrett Albanesius. Change in the social life of urban public spaces: The rise of mobile phones and women, and the decline of aloneness over 30 years. *Urban Studies*, 52(8):1489–1504, 2015.

[8] Glenn Jocher. YOLOv5 by Ultralytics, May 2020.

[9] Nicolai Wojke, Alex Bewley, and Dietrich Paulus. Simple online and realtime tracking with a deep association metric. In *2017 IEEE international conference on image processing (ICIP)*, pages 3645–3649. IEEE, 2017.

[10] Antonio Brunetti, Domenico Buongiorno, Gianpaolo Francesco Trotta, and Vitoantonio Bevilacqua. Computer vision and deep learning techniques for pedestrian detection and tracking: A survey. *Neurocomputing*, 300:17–33, 2018.

[11] Daniel Stadler and Jurgen Beyerer. Improving multiple pedestrian tracking by track management and occlusion handling. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10958–10967, 2021.

[12] Becky PY Loo and Zhuangyuan Fan. Social interaction in public space: Spatial edges, moveable furniture, and visual landmarks. *Environment and Planning B: Urban Analytics and City Science*, page 23998083231160549, 2023.

[13] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, pages 740–755. Springer, 2014.

[14] Richard W Bohannon. Comfortable and maximum walking speed of adults aged 20—79 years: reference values and determinants. *Age and ageing*, 26(1):15–19, 1997.

[15] Elaine M Murtagh, Jacqueline L Mair, Elroy Aguiar, Catrine Tudor-Locke, and Marie H Murphy. Outdoor walking speeds of apparently healthy adults: A systematic review and meta-analysis. *Sports Medicine*, 51:125–141, 2021.

[16] Erich Schubert, Jörg Sander, Martin Ester, Hans Peter Kriegel, and Xiaowei Xu. Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21, 2017.

[17] Edward T Hall. A system for the notation of proxemic behavior. *American anthropologist*, 65(5):1003–1026, 1963.

[18] Weina Ge, Robert T Collins, and Barry Ruback. Automatically detecting the small group structure of a crowd. In *2009 workshop on applications of computer vision (WACV)*, pages 1–8. IEEE, 2009.

[19] Qi Wang, Mulin Chen, Feiping Nie, and Xuelong Li. Detecting coherent groups in crowd scenes by multiview clustering. *IEEE transactions on pattern analysis and machine intelligence*, 42(1):46–58, 2018.

[20] Robert Putnam. Bowling alone: The collapse and revival of american community. 2000.

[21] R Devi. Geometric transformations and thresholding of images using opencv-python. *GRD Journal for Engineering*, 2(11):49–52, 2017.

[22] Zhewei Huang, Tianyuan Zhang, Wen Heng, Boxin Shi, and Shuchang Zhou. Real-time intermediate flow estimation for video frame interpolation. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022.

[23] Keith N Hampton, Oren Livio, and Lauren Sessions Goulet. The Social Life of Wireless Urban Spaces: Internet Use, Social Networks, and the Public Realm. *Journal of Communication*, 60(4):701–722, 2010.

[24] Wenhan Wu, Wenfeng Yi, Jinghai Li, Maoyin Chen, and Xiaoping Zheng. Automatic identification of human subgroups in time-dependent pedestrian flow networks. *IEEE Transactions on Multimedia*, 26:166–177, 2023.

[25] Mohamed H Zaki and Tarek Sayed. Automated analysis of pedestrian group behavior in urban settings. *IEEE Transactions on Intelligent Transportation Systems*, 19(6):1880–1889, 2017.