All Fun and Games: Obtaining Critical Pedestrian Behavior Data from an Online Simulation

Kai Holländer
LMU Munich
kai.hollaender@ifi.lmu.de

Changkun Ou
LMU Munich
changkun.ou@ifi.lmu.de

Luca Schellenberg
LMU Munich
luca.schellenberg@web.de

Andreas Butz
LMU Munich
andreas.butz@ifi.lmu.de

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
http://dx.doi.org/10.1145/3334480.3382797

Abstract
Automated cars will need to observe pedestrians and react adequately to their behavior when driving in urban areas. Judging pedestrian behavior, however, is hard. When approaching it by machine learning methods, large amounts of training data is needed, which is costly and difficult to obtain, especially for critical situations. In order to provide such data, we have developed an online game inspired by *Frogger*, in which players have to cross streets. Accidents and critical situations are a natural part of the data produced in such a way without anybody getting hurt in reality. We present the design of our game and an analysis of the resulting data and its match to real world behavior observed in previous work. We found that behavior patterns in real and virtual environments correlated and argue that game data could be used to train machine learning algorithms for predicting real pedestrians' walking trajectories when crossing a road. This approach could be used in future automated vehicles to increase pedestrian safety.

Author Keywords
Games with a purpose, pedestrian safety, machine learning, automated vehicles, user behavior.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);** User studies;

Figure 1: TWD start panel.

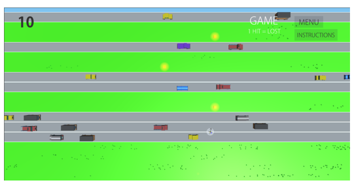

Figure 2: TWD top view.

Figure 3: TWD ego-perspective.

Introduction

Automated Vehicles will reduce the influence of human drivers and hence the likeliness of accidents caused by human error. While this is expected to lead to increased road safety [11], the technical capabilities of current automated vehicle technology might be overrated [8].

For instance, as a consequence of overtrust in technology, a driver died in 2018[1]. In the same year, an automated vehicle crash with a pedestrian ended deadly[2], because even a trained technical supervisor had overtrusted the system. Recently, Tesla's 'Summon' feature lead to various damages close to areas with pedestrians, for instance, on parking lots[3]. Such incidents could severely affect the acceptance of automated vehicles and hinder a transition from manual driving to fully automated mobility [10].

In order to create trustworthy automated cars that can truly replace human drivers, a human-like understanding of traffic situations is inevitable [15]. Hence, an adequate model of pedestrian behavior is a crucial aspect for automated vehicles to safely navigate through urban environments [13]. However, we believe that such a model should include the pedestrians' perspective by design.

The training of future machine learning based auto pilot systems requires large amounts of training data, both of regular and of critical situations. However, collecting quantified critical pedestrian behavior data remains an open challenge [13, 17]. Crossing a street in the real world depends on a wide range of potentially influencing factors, for example, the time to arrival of vehicles and the gap size between cars [3]. Previous research regarding pedestrian behavior predictions is mainly based on computer vision approaches trained from dash-cam data (i.e., the driver's perspective) [9]. However, this requires video capturing which violates general data privacy regulations in many European countries[4]. In order to overcome such legal issues and at the same time also gather quantified data from a pedestrian perspective in a controlled environment, we propose a crowd-sourcing approach.

In crowd-sourcing research, it has been shown that *games with a purpose* can produce, for example, meaningful image labels [20]. Our goal is to go beyond this type of tagging or labeling by directly collecting critical behavioral data in a simulated environment. We therefore implemented an online game based on the console game *Frogger*. Our game (called *The Walking Data* or TWD for short) is available online[5]. TWD requires players to cross as many lanes and roads as possible without causing a collision. Below, we present TWD and an initial analysis of the gathered data.

Instead of microscopic behavior modeling, we deployed the game in a pilot study on potential users, to investigate whether real-world behavior can be modeled by the gaming data. Then compared the outcome to findings of corresponding macroscopic real world observations. Our theory-driven analysis (as described in [1, 22]) suggests that the observations in the real world correspond to the game behavior. Thus, we propose a methodology for accumulating large-scale pedestrian behavior data, especially for the critical traffic situations that are difficult to capture otherwise.

Research Question & Hypotheses

The underlying question behind this work is: "Could real-world behavior be replicated and extracted through an on-

[1] Tesla accident; *last accessed: Nov 2019*
[2] Uber's fatal self-driving crash; *last accessed: Nov 2019*
[3] Tesla's Smart Summon feature; *last accessed: Nov 2019*
[4] Dashcams - permissible or prohibited? *last accessed: Nov 2019*
[5] The Walking Data; *last accessed: Jan 2020*

Data stored each time a gap is accepted or not accepted, a yellow block is collected, a lane is crossed or a collision occurs, or the player is facing the seventh lane:

- Player name & ID
- Gender (f/m/o)
- Age (int)
- Score (int)
- City and country of player
- Arrangement of roads and lanes
- Player position (x/y/z)
- Velocity (x/y/z) of game objects
- Position, distance and speed of appr. car
- Width of appr. vehicle
- Waiting & walking times of player
- Viewing angle of player
- Size and lane of gap
- Size and lane of far-side gaps
- Distance, speed and width of appr. vehicles of far-side gaps
- Reason which triggered saving data

line game?". The basis for the comparison of the behavior patterns are previous scientific observations of jaywalking. Those findings inform the hypotheses of our work and we investigate whether data from our game shows the same phenomena as those real-world observations. For example, Wang et al. [21] state that many people do not consider the far-side gap and thus wait at the middle of two-lane roads.

Furthermore, we were interested in the effects of the player's perspective. Therefore, The Walking Data was initially published with a top view (Figure 2), and later changed to a pedestrian ego perspective (Figure 3). We investigate the following hypotheses:

H1: Crossing decisions (accepted gaps between vehicles) depend on the distance of the vehicles rather than on their speed [12, 23].

H2: Pedestrians prefer safer over shorter paths and always look out for oncoming vehicles [24].

In addition to the real world observations, we also consider the perspective and explore which one is more suitable to gather realistic behavior.

Research Approach

Our independent variables are the environment (real-world / game) and the view (ego perspective, top-view). We identified relevant dependent variables through a literature review (see sidebar for all collected variables) [5, 14, 16, 19, 23]. Then, we developed design sketches on paper, and a test version of the game with Unity[6]. We invited three volunteers to a think-aloud test session. Afterwards, we adjusted the game according to insights from the think-aloud protocol. For example, we replaced the control keys for the game and extended the duration of the initial tutorial mode.

[6] Unity 3-D Development Platform; *last accessed: Nov 2019*

Game Concept

The Walking Data is based on *Frogger*[7] and *Crossy Road*[8]. These games require players to move a virtual character across road lanes on the screen without causing collisions. This choice was made two main reasons: First, both games are documented to be entertaining (*Crossy Road* has 4.3 million downloads and an average user rating of 4.6 / 5[8]; *Frogger* was sold over 20 million times[9]). An entertaining gameplay in turn is an essential aspect for our game, because this is the only thing we offer volunteers for their participation. Second, the concept of road crossings and the mental model of the game behavior are comparable to their counterparts in real world crossing scenarios. Our study therefore matches the mapping principles [1, 22].

The technical setup consists of an HTML / Javascript front-end exported from Unity, and the game communicates to a REST API server and stores the recorded behavior data in a MySQL database on the back-end. The whole web service is hosted on a publicly accessible server.

Game Design

At first, visitors of the website see a splash screen with the logo of the game. Subsequently, players are asked to submit a unique user name, their age and gender. We optimized the user interface design for speed and simplicity and thus request only these three inputs to reduce decision time and complexity [7]. The time needed to select a target correlates with its distance and size [6]. Hence, buttons and input fields are located close to each other in the middle of the screen, see Figure 1. The game is preset to either a top view or the ego-perspective. Each session starts with

[7] Frogger Arcade Game; *last accessed: Nov 2019*
[8] Crossy Road; *last accessed: Nov 2019*
[9] Konami's Frogger; *last accessed: Nov 2019*

TWD Settings

The width of each track and green area unit is four meters [4]. The resolution of the game is 1200×600 px, the horizontal field of view is set to 50 degrees in ego-perspective and 60 degrees in top view to avoid perspective distortion. The size of the vehicles is taken from real values[a], depending on the vehicle type. The avatar has a view height of 1.7 m. Players move continuously and not in steps. The maximum speed of the avatar is $1.42\,\mathrm{m\,s^{-1}}$, (average speed of men and women for usual walking [2, p. 15]). The pedestrian's acceleration is set to $1.69\,\mathrm{m\,s^{-1}}$, (average acceleration of men and women for usual walking [2, p. 21]). A displacement of the cars relative to the middle of the roadway is set within $\pm 0, 6$ m via normally distributed random values.

[a] Dimensions of Vehicles; *last accessed: Nov 2019*

a tutorial mode which shows an overlay explaining the control and how to gain points. In the tutorial mode, a collision has no consequences. Players can become familiar with the controls and environment. The aforementioned game *Crossy Road* inspired the graphical design of TWD. Objects are abstracted and colors appear with a bright, high contrast look, which is common for this type of games.

Through a think-aloud session we learned that the controls should be as easy as possible while still allowing the game character to move everywhere. Therefore, players select a position with the mouse cursor and can walk with the 'w'-key. When the key is released, the player stops. The goal for players is to reach as many points as possible by crossing roads and collecting yellow blocks.

The environment consists of roads, trees, green areas, clouds, plants, and yellow blocks. Objects either serve as reference points to ease speed estimations or to guide players. Yellow blocks yield extra points and represent points of interest. The first three yellow blocks appear at fixed positions. Since we believe that in the real world, points of interest affect crossing decisions, the idea of yellow blocks is to influence the chosen path of a player with a precise goal.

When a player reaches 45 points in 'tutorial' mode, it switches to 'game' with a seven seconds countdown. In 'game' mode, a collision leads to a full reset of the score and position. Points can be earned by either crossing a lane successfully (10 points) or by collecting yellow blocks (15 points). Whenever seven lanes were crossed in game mode, the speed of moving objects increases steadily, in order to increase the difficulty and challenge for players.

Game Properties

Only the first six lanes are evaluated; all subsequent lanes are only there to improve the game experience. The first three roads always consist of one, two, and three lanes, presented in random order. On each lane, vehicles can approach either from the right or from the left. The parameter of direction is randomly configured when the game starts. On two lane roads, driving directions are always opposing.

The driving characteristics of the cars differ, and neither the acceleration nor the driving speed are static. After loading the environment, vehicles are placed in the game world and accelerated with a random value between 3.4m/s^2 and 7m/s^2 to a speed of about 8.33m/s$(\approx 30$km/h$)$. The final speed has a deviation of $\pm 33\%$ to be as realistic as possible. In addition, cars with different speeds and distances are relevant for our data analysis.

The distance between cars is a random value between the minimum braking distance of the car behind and 70m. The distance is determined with a normally distributed random function, and the braking distance is $(0.1 \cdot v_c)^2$ where v_c is the current velocity. The Sidebar on this page contains further environment settings and attributes.

User Study

A total of 78 participants contributed in 89 games (see Table 1). The first three roads include one, two and three lanes with differing lane arrangements (two lanes, followed by one lane, followed by three lanes; three lanes first, followed by one lane and then two lanes etc.). Games in which less than six lanes were crossed were excluded from analysis to automatically exclude erroneous behavior such as walking in zig-zag paths, constant reciprocating, and walking to the borders of the game world, which did not serve the objective of the game. Players were recruited via digital channels and accessed TWD from Germany, Austria, France, Italy, and the Netherlands. We collected two main data sets for the two perspectives during one week each.

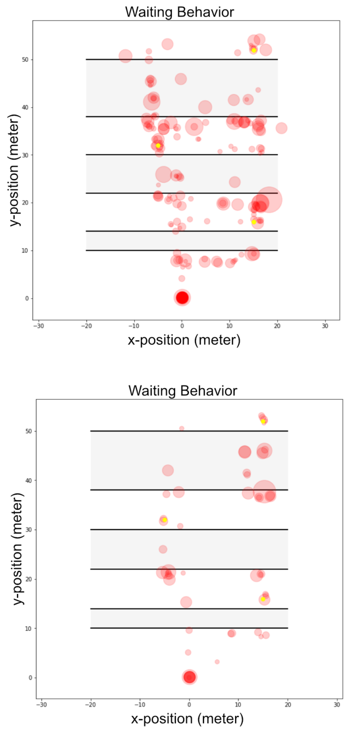

Figure 4: Waiting behavior data from ego-perspective (top) and top view (bottom) recordings, showing waiting positions (red circles) and positions of yellow blocks (yellow circles). The longer a player waited, the bigger the circle. If circles overlap, color saturation increases. Gray bars represent roads with one, two, or three lanes.

Table 1: Distribution of Participants.

	Ego-Perspective	**Top** View
Games	58	31
Players	47	31
Women	31 %	24 %
Men	69 %	58.5 %
Other	0	17.5 %
Mean Age (SD)	27.2 (6.89) years	27.7 (13.52) years

Theory-driven Analysis & Discussion

To answer H1, we performed a logistic regression. The logit model includes a normally distributed random effect for each player, to account for individual differences. Table 2 shows corresponding results from 58 games including 92 crossing decisions. Table 3 includes results from 31 games including 62 crossing decisions.

Distance is the only significant impact factor of the parametric coefficients on the crossing decision with $(Pr(>|z|) = 0.006)$ (ego-perspective) and $(Pr(>|z|) = 0.0002)$ (top-view). Thus, we can accept H1 and state that crossing decisions in TWD are rather based on the distance to approaching vehicles than their speed, player's gender, or age. Hence, both perspectives are in line with related work, considering the effect of distance [12, 23]. The ego-perspective replicates age-related effects additionally.

Table 2: Analysis H1 | ego-perspective.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-2.378	2.156	-1.103	0.270
distance	0.052	0.019	2.761	0.006
velocity	-0.164	0.163	-1.008	0.314
gender_m	0.415	0.517	0.803	0.422
age	0.089	0.047	1.914	0.056

Table 3: Analysis H1 | top view.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	2.483	2.509	0.000	0.999
distance	9.712	2.659	3.653	0.0002
velocity	-1.769	2.251	-0.786	0.432
gender_f	-2.507	2.509	0.000	0.999
gender_m	-2.427	2.509	0.000	0.999
age	-8.503	1.313	-0.647	0.517

Wang et al. [21] report:

> "...we found that many pedestrians cross the road regardless of the far-side gap, [...] resulting in the fact that such pedestrians wait at the middle of the road for the next possible gaps [...] to continue to cross the road." [21, p. 4].

The authors observed a two-lane road. In comparison, on two-lane streets in TWD, 35% of players waited with an ego-perspective and 31% with a top view, which can be argued to match 'many' from the cited paper. Thus, the ego-perspective leads to a higher percentage of people waiting on the road. Interestingly, the more lanes there are, the more people tend to wait on the street, see Figure 4[10].

According to Zhuang and Wu [24], pedestrians prefer safer over shorter paths. Figure 5 shows a corresponding categorization of paths. Figure 6 shows chosen paths in TWD. We can see that pedestrians also select safe routes and

[10] Due to the limited space the street layouts in presented margin figures consistently includes roads with successive one, two and three lanes. Please contact the authors if you want to retrieve the results from all possible lane combinations (six data sets).

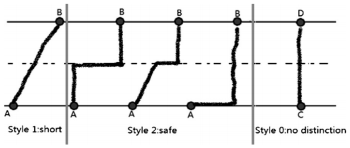

Figure 5: Classification of paths [24] as 'safe' or 'short'.

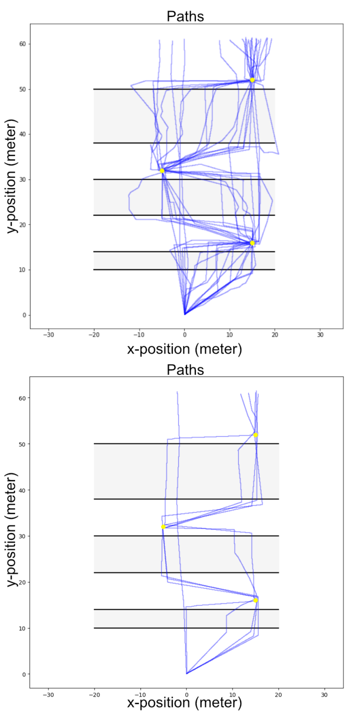

Figure 6: Top: ego-perspective paths (blue) and three yellow blocks (yellow), bottom: top-view paths. Gray bars represent roads with one, two, and three lanes.

are highly influenced by artificial points of interest (yellow blocks). In line with Zhuang and Wu [24], we observed that all pedestrians looked for cars. A 'look' was defined as a camera rotation of >37 degrees from the walking direction.

Based on our results, we conclude that the ego-perspective is better suited to reproduce real world behavior than the top view. We suspect that an ego-perspective leads to a stronger feeling of embodiment and thus provokes behavior closer to reality. Additionally, the ego-perspective motivates players to complete multiple rounds (ego-perspective: 47 players, 58 games; top view: 31 players, 31 games).

Limitations

A limitation of our approach could be cultural differences. Our hypotheses are inspired by observations from Greece [23], Australia [12] and China [21,24]. However, via the database, any desired location could be excluded for location specific analysis. Another limitation might be the appearance of the game. We did not try to implement a photo-realistic look and feel. Other studies in the context of automated vehicles also implemented an abstract look, for example in the work of Siripanich [18]. For the future, we plan to run a comparison study with a more realistic visual game design to validate if there are significant differences in player behavior if the game appearance changes.

The outcome of a collision in the real world is worse than in TWD and therefore, the behavior might be different. We do not claim that behavior which might result in injuries (real world) matches perfectly with behavior resulting in a loss of points (virtual environment). Nevertheless, the motivation to 'survive' in the game follows a similar mental model as crossing in the real world. In both environments people aim to avoid vehicles to not face consequences. Other than in some racing games, where players can simply continue even after a high speed collision, TWD does not allow players so resume after a crash and resets all current achievements. The overlap of game and real life outputs can furthermore only be stated within the scope of our hypotheses. We do not know yet if real world observations overlay with game trajectories. Our results are interpreted as an initial indicator for some degree of matching outputs and will be compared to real world observations in a next step.

Conclusion & Future Work

The goal of this study was to find out whether strategies and behavior patterns in reality overlap with those demonstrated in *The Walking Data* (TWD). Our theory-driven analysis (according to the paradigm described in [1, 22]) indicates that both environments evoke similar behavior. Thus, gathering accurate large-scale data on pedestrian behavior could become less costly in terms of time and money through a game based on real-world parameters. For a wider evaluation of data we will include external measurements, e.g., time on street. Such information can be extracted from our data set through recorded timestamps and spatial position data (x-y-z coordinates) within TWD.

We are currently implementing an autoregressive model with a recurrent neural network (e.g., attention-based models) based on data gathered through TWD. Afterwards, we will perform a cross-validation of our data set. We plan to verify whether our model successfully predicts pedestrian trajectories in a data-driven approach. If this produces convincing results relative to the game-based data, the predicted paths will be compared to observations in the wild to finally assess the value of the outcome. We will also publish an anonymized version of the data set.

Please feel free to try the game and support our research: https://thewalkingdata.medien.ifi.lmu.de/.

REFERENCES

[1] Muhammad Aurangzeb Ahmad, Cuihua Shen, Jaideep Srivastava, and Noshir Contractor. 2014. On the Problem of Predicting Real World Characteristics from Virtual Worlds. In *Predicting Real World Behaviors from Virtual World Data*, Muhammad Aurangzeb Ahmad, Cuihua Shen, Jaideep Srivastava, and Noshir Contractor (Eds.). Springer International Publishing, Cham, 1–18.

[2] Bettina Bartels and C Erbsmehl. 2014. Bewegungsverhalten von Fußgängern im Straßenverkehr, Teil 1. *FAT-Schriftenreihe* 267 (2014).

[3] Matthias Beggiato, Claudia Witzlack, and Josef F. Krems. 2017. Gap Acceptance and Time-To-Arrival Estimates As Basis for Informal Communication Between Pedestrians and Vehicles. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)*. ACM, New York, NY, USA, 50–57. DOI:http://dx.doi.org/10.1145/3122986.3122995

[4] European Commission. 2018. Roads. (February 2018). Directorate General for Transport.

[5] Debargha Dey and Jacques Terken. 2017. Pedestrian Interaction with Vehicles: Roles of Explicit and Implicit Communication. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17)*. Association for Computing Machinery, New York, NY, USA, 109–113. DOI: http://dx.doi.org/10.1145/3122986.3123009

[6] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.

[7] William E Hick. 1952. On the rate of gain of information. *Quarterly Journal of experimental psychology* 4, 1 (1952), 11–26.

[8] Kai Holländer, Philipp Wintersberger, and Andreas Butz. 2019. Overtrust in External Cues of Automated Vehicles: An Experimental Investigation. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '19)*. ACM, New York, NY, USA. DOI: http://dx.doi.org/10.1145/3342197.3344528

[9] Julian F. P. Kooij, Fabian Flohr, Ewoud A. I. Pool, and Dariu M. Gavrila. 2019. Context-Based Path Prediction for Targets with Switching Dynamics. *International Journal of Computer Vision* 127, 3 (01 Mar 2019), 239–262. DOI: http://dx.doi.org/10.1007/s11263-018-1104-4

[10] John D Lee and Kristin Kolodge. 2018. Understanding attitudes towards self-driving vehicles: Quantitative analysis of qualitative data. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 1399–1403.

[11] Todd Litman. 2019. Autonomous Vehicle Implementation Predictions Implications for Transport Planning. *Victoria Transport Policy Institute* (October 2019), 40.

[12] Jennifer A. Oxley, Elfriede Ihsen, Brian N. Fildes, Judith L. Charlton, and Ross H. Day. 2005. Crossing Roads Safely: An Experimental Study of Age Differences in Gap Selection by Pedestrians. *Accident Analysis & Prevention* 37, 5 (2005), 962 – 971. DOI: http://dx.doi.org/https://doi.org/10.1016/j.aap.2005.04.017

[13] Amir Rasouli, Iuliia Kotseruba, Toni Kunic, and John K. Tsotsos. 2019. PIE: A Large-Scale Dataset and Models for Pedestrian Intention Estimation and Trajectory Prediction. In *The IEEE International Conference on Computer Vision (ICCV)*.

[14] Amir Rasouli, Iuliia Kotseruba, and John K. Tsotsos. 2018. Towards Social Autonomous Vehicles: Understanding Pedestrian-Driver Interactions. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE. DOI: http://dx.doi.org/10.1109/itsc.2018.8569324

[15] A. Rasouli, I. Kotseruba, and J. K. Tsotsos. 2018. Understanding Pedestrian Behavior in Complex Traffic Scenes. *IEEE Transactions on Intelligent Vehicles* 3, 1 (March 2018), 61–70. DOI: http://dx.doi.org/10.1109/TIV.2017.2788193

[16] Amir Rasouli, Iuliia Kotseruba, and John K. Tsotsos. 2018. Understanding Pedestrian Behavior in Complex Traffic Scenes. *IEEE Transactions on Intelligent Vehicles* 3, 1 (March 2018), 61–70. DOI: http://dx.doi.org/10.1109/tiv.2017.2788193

[17] Markus Rothmüller, Pernille Holm Rasmussen, and Signe Alexandra Vendelbo-Larsen. 2018. Designing for Interactions with Automated Vehicles: Ethnography at the Boundary of Quantitative-Data-Driven Disciplines. *Ethnographic Praxis in Industry Conference Proceedings* 18, 1 (2018), 482–517. DOI: http://dx.doi.org/10.1111/1559-8918.2018.01219

[18] Siripanich. 2017. Crossing the road in the world of autonomous cars. TEAGUE Labs. (Aug 2017). https://bit.ly/2rp8rNh

[19] Wonho Suh, Dwayne Henclewood, Aaron Greenwood, Angshuman Guin, Randall Guensler, Michael P Hunter, and Richard Fujimoto. 2013. Modeling pedestrian crossing activities in an urban environment using microscopic traffic simulation. *SIMULATION* 89, 2 (Jan. 2013), 213–224. DOI: http://dx.doi.org/10.1177/0037549712469843

[20] Luis von Ahn and Laura Dabbish. 2004. Labeling Images with a Computer Game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04)*. ACM, New York, NY, USA, 319–326. DOI: http://dx.doi.org/10.1145/985692.985733

[21] T. Wang, J. Wu, P. Zheng, and M. McDonald. 2010. Study of pedestrians' gap acceptance behavior when they jaywalk outside crossing facilities. In *13th International IEEE Conference on Intelligent Transportation Systems*. 1295–1300. DOI: http://dx.doi.org/10.1109/ITSC.2010.5625157

[22] Dmitri Williams. 2010. The Mapping Principle, and a Research Framework for Virtual Worlds. *Communication Theory* 20, 4 (2010), 451–470. DOI: http://dx.doi.org/10.1111/j.1468-2885.2010.01371.x

[23] G. Yannis, E. Papadimitriou, and A. Theofilatos. 2013. Pedestrian gap acceptance for mid-block street crossing. *Transportation Planning and Technology* 36, 5 (2013), 450–462. DOI: http://dx.doi.org/10.1080/03081060.2013.818274

[24] Xiangling Zhuang and Changxu Wu. 2011. Pedestrians' crossing behaviors and safety at unmarked roadway in China. *Accident Analysis & Prevention* 43, 6 (2011), 1927 – 1936. DOI: http://dx.doi.org/https://doi.org/10.1016/j.aap.2011.05.005