Urban Planning
2024 • Volume 9 • Article 6932
https://doi.org/10.17645/up.6932

ARTICLE

Open Access Journal

An Empirical Test of Pedestrian Activity Theories Within Informal Settlements

Yael Borofsky [1], **Stephanie Briers** [2], **and Isabel Günther** [1]

[1] Development Economics Group, ETH Zürich, Switzerland
[2] Transdisciplinarity Lab (TdLab), ETH Zürich, Switzerland

Correspondence: Yael Borofsky (yael.borofsky@nadel.ethz.ch)

Submitted: 22 March 2023 **Accepted:** 22 August 2023 **Published:** 7 March 2024

Issue: This article is part of the issue "Urban Borderlands: Difference, Inequality, and Spatio-Temporal In-Betweenness in Cities" edited by Deljana Iossifova (University of Manchester) and David Kostenwein (ETH Zurich), fully open access at https://doi.org/10.17645/up.i312

Abstract

Pedestrian activity is often measured in the formal parts of cities, yet it has rarely been studied in informal settlements, although they are typically adjacent to formal areas and residents participate in formal urban life. Route optimization and space syntax are two pedestrian activity theories that can be applied to predict path usage in urban areas. These theories have been tested in formal cities, but are they applicable in understudied informal settings? Using motion sensors, we measure pedestrian activity in a Cape Town informal settlement in the early morning and evening hours and test which theory best explains the sensor measurements. Route optimization is weakly correlated with average pedestrian activity, while space syntax performs even more poorly in predicting pedestrian activity. The predictive power of both theoretical calculations further varies by time of day. We find that both theories perform worst at the entrances/exits of the informal settlement—that is, the border between informal and formal. These results indicate that daily movement patterns in informal settlements may differ from formal areas and that the connection between the formal and informal city requires further study to better understand how pedestrian activity links these two types of areas. A new theory of route selection based on such an understanding, which also better incorporates the specific characteristics of informal urban settlements—such as high density, narrow, and constantly changing streets primarily used by residents—may be necessary to understand the needs of pedestrians within informal settlements as compared to formal areas.

Keywords

informal settlement; nighttime activity; pedestrian; route optimization; sensor; space syntax

1. Introduction

Pedestrian mobility in cities is well-studied, but the focus is almost always on the formal parts of cities. Little is known about differences in pedestrian life between formal and informal areas (Anciaes et al., 2017; Cutini et al., 2020; Hillier et al., 2000; Mehta, 2008; Salon & Gulyani, 2010; Sletto & Palmer, 2017) and even fewer studies quantitatively analyze pedestrian motion inside informal settlements (Hidayati et al., 2020; Hillier et al., 2000; Mohamed, 2016). Meanwhile, more than a billion people, every seventh person, live in informal settlements worldwide (Dodman et al., 2018). In many low- and middle-income countries, these neighborhoods are increasingly the urban norm, not the exception (Simone & Pieterse, 2017), as they expand in size and number. Extreme density, unpaved roads, reliance on shared infrastructure, vulnerability to crime, frequent house fires, and natural disasters, as well as the need for easier connection to the formal economy together suggest that understanding how pedestrians navigate in informal settlements can inform efforts to address these issues and improve well-being (Mutyambizi et al., 2020; Samper & Liao, 2023; Walls et al., 2018).

We address this knowledge gap using pedestrian motion sensors installed throughout the path network in one informal settlement in Cape Town, South Africa. We measured pedestrian activity in the evenings and early mornings from October to November 2019. Using motion data and the network structure, we analyze movement patterns in the context of two prevailing theories about pedestrian route choice based on network characteristics—derived from and tested with movement patterns in high-income, formal cities—to understand how well they explain the empirical data.

The first theory—route optimization—posits that pedestrians optimize for a goal, such as the shortest or fastest route (Willis et al., 2004). This theory is operationalized by calculating the shortest metric or time-distance route. The shortest path is the dominant heuristic because it theoretically maximizes pedestrian utility (Zhu & Levinson, 2015). As critics argue, however, the theory assumes pedestrians optimize accurately (Law & Traunmueller, 2018; Salazar Miranda et al., 2021). The second theory—space syntax—asserts that the urban network configuration is a core determinant of pedestrian activity and that measures derived from the network's topology and geometry can explain pedestrian route choice (Hillier, 2007; Hillier & Iida, 2005; van Nes & Yamu, 2021; Willis et al., 2004; Yamu et al., 2021). Hillier called this phenomenon the law of natural movement, arguing that these parameters apply in nearly any urban space, from cities to buildings (Hillier, 2007; Hillier et al., 1993).

There is an ongoing academic debate about which theory better explains pedestrian motion (Shatu et al., 2019). Various studies empirically test these theories in formal cities, mainly in the Northern Hemisphere (Bongiorno et al., 2021; Hillier & Iida, 2005; Sharmin & Kamruzzaman, 2018; Shatu et al., 2019), but few empirically study whether they apply in dense, oft-changing informal neighborhoods (Hidayati et al., 2020; Hillier et al., 2000; Mohamed, 2016).

While generalizing differences between informal and formal areas is a source of academic debate (Samper & Liao, 2023), there are several reasons why movement in informal settlements might be different in many cases. First, the path network is not formally planned, emerging from individual (e.g., where to build houses) and community-level decisions (e.g., to block throughways to enhance safety; Samper & Liao, 2023). As a result, the network constantly evolves. Second, activities considered private in most formal areas (e.g., accessing shared sanitation) often require informal settlement residents to enter public space. For example, Mutyambizi et al. (2020) conducted a representative survey of informal settlements in South Africa, finding 50% share a toilet with other households. Informal settlement paths can be described as "spaces in between" (Cutini et al., 2020) or "liminal" (Sletto & Palmer, 2017) because expectations about what is public vs. private space differ from formal areas. Third, informal settlements tend to be extremely dense (Visagie & Turok, 2020), making it difficult to assess distances or other route characteristics by sight alone. Fourth, informal settlements are often not well integrated into the broader formal urban network, creating a de facto border between the two types of spaces and limiting who enters and uses informal spaces (Karimi & Parham, 2012). Hence, most pedestrians are likely to be residents. Finally, informal settlements tend to have minimal/no streetlighting, radically changing the experience of navigation at night (Kretzer, 2020).

We first empirically measure movement between 6:00 pm and 8:00 am with sensors. Second, we apply a shortest path analysis from route optimization and a space syntax analysis of choice to the informal settlement path network to theoretically predict movement patterns. From the empirical data, we find activity is highest in the evening and early morning, and that there is more movement on weekends compared to weekdays. Correlating path usage from the two theory-based predictions with empirical data, we find the shortest path (route optimization) prediction weakly correlates with observed evening, but not early morning, activity. It performs equally well on weekdays and weekends. In contrast, we find no significant correlation between space syntax choice predictions and empirically measured activity.

This article makes three contributions. First, we test how well two theories about pedestrian route choice derived from formal cities in high-income countries predict pedestrian movement in low-income informal settlements. The second is the testing of pedestrian motion sensors for studying informal settlement route network usage. To our knowledge, this method has never been used before in this context, though it is frequently used to study path usage in formal areas. Third, we measure pedestrian activity in informal settlements at high granularity in the early mornings and evenings when paths are used most, which can shape our understanding of path usage in informal settlements.

2. Context

In Cape Town, South Africa, there are more than 600 informal settlements (City of Cape Town, 2019). We study an approximately 30-year-old informal settlement that is about 38,200 square meters (Ndifuna Ukwasi et al., n.d.), located in Khayelitsha, on the outskirts of Cape Town. Known as a "pocket informal settlement," the neighborhood is surrounded by formal houses and streets, creating a porous border between the formal and informal experience. We began studying this site as part of a broader project on public lighting in informal settlements because the community leadership was open to collaborating on transdisciplinary research.

Like many informal settlements (Kamalipour & Dovey, 2019), the path network was previously unmapped. To map it, we started with a 2018 satellite image of the informal settlement (City of Cape Town, 2018) and worked with local leaders to draw all of the network features. Figure 1 shows the informal settlement map in August 2019, just before installing sensors.

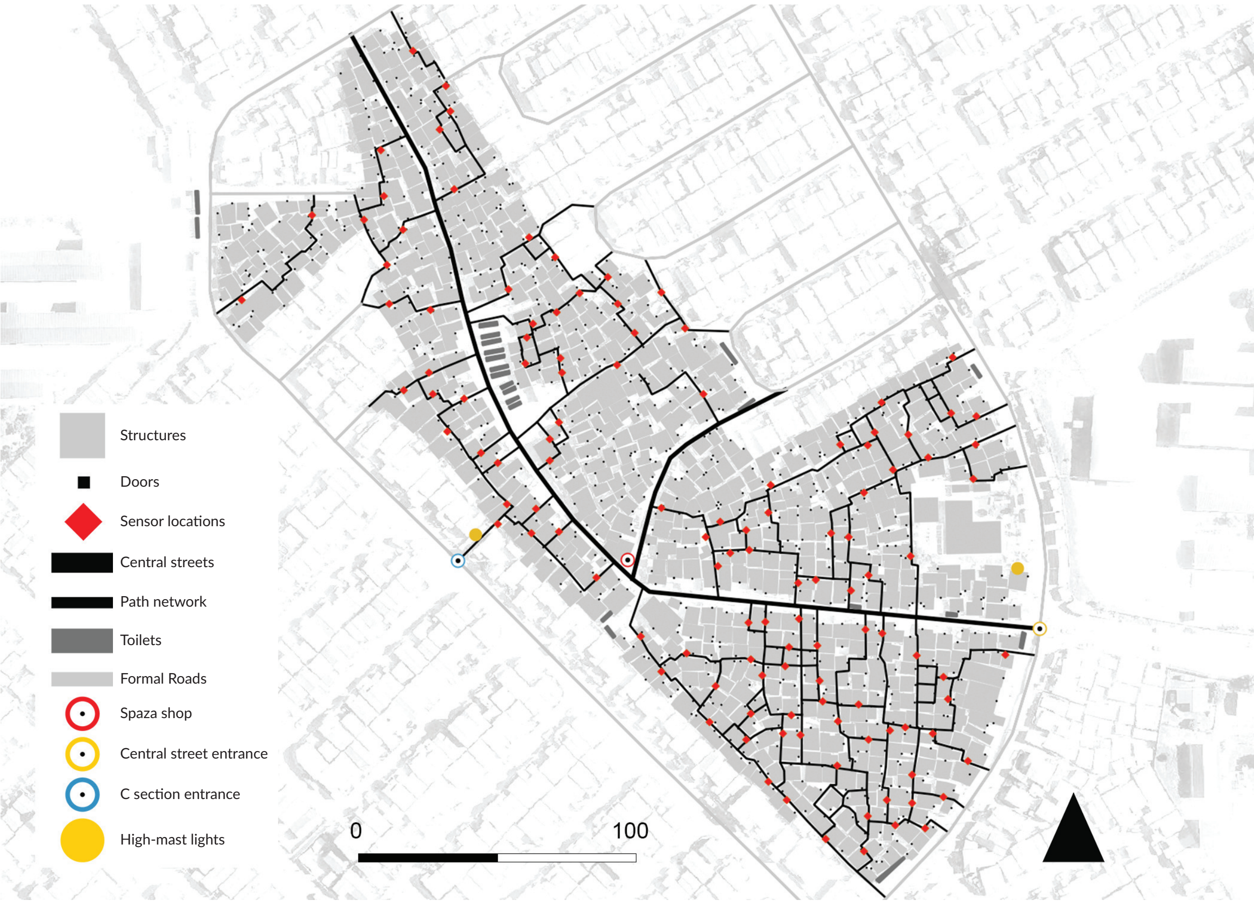

Figure 1. Study area and path network map.

We classify the path network into three categories:

1. *Central streets* (bold black lines) bisect the neighborhood. They are sand covered, but passable with a vehicle. These streets are included in the theory-based calculations because they are major connection points to formal areas, which could influence movement within the settlement, but they were not monitored with sensors because vehicles would have biased sensor measurements.
2. *Formal streets* (gray lines) are paved and surround the neighborhood. They have both vehicular and pedestrian traffic. They link the neighborhood to the surrounding economy and are included in the theory-based calculations because they influence movement within the settlement. Formal streets were not included in the sensor study because they are outside the informal settlement. Furthermore, the sensors were not designed to accurately measure wide, formal streets and vehicles using these streets would also have biased sensor measurements.
3. *Paths* (thin black lines) are pedestrian walkways. Sensors were only placed in paths. Our empirical analysis of pedestrian activity and the comparison between empirical data and theoretical predictions are based on path movement only.

3. Data and Method

3.1. Sensor Data

Many technologies can be used to measure pedestrian activity. Each approach has pros and cons. For our setting, a few criteria narrowed the options. Manual counting was unsafe during dark hours. Video surveillance was rejected to protect privacy. Although mobile phones are an increasingly common source of mobility data worldwide, a GPS tracking pilot we conducted in March 2019 showed it would not work in an informal settlement for several reasons. First, a household survey we conducted revealed not all residents own a smart phone and only 37% of respondents carry one after sunset for fear of theft, leading to sample selection bias (for more about the household survey see Supplementary File). Second, 50% of residents with a smart phone and willingness to enroll in the pilot did not have enough storage space for the tracking application. Third, GPS is inaccurate within a few meters, which is problematic when paths are narrow and houses are small, as they are in this informal settlement, because it would be difficult to determine which paths were used. Moreover, some building materials used in informal settlements reportedly interfere with GPS.

We installed 122 pedestrian motion sensors on paths in September 2019. We only use data between 6:00 pm–8:00 am because the sensor's sensitivity prevents it from accurately measuring daytime motion, which is when thermal radiation from local building materials can cause false triggers. To verify data quality, we conducted manual counts between 5:00–7:00 pm, when people are outside and it is safe to work, then compared the human-observed counts to the data from the same period. During the study, some sensors were stolen/vandalized and some malfunctioned (e.g., discharged battery). To adjust for attrition, we drop data from sensors that did not function throughout the entire study. To allow for minor, random data loss caused by Bluetooth connectivity issues, we include all sensors that had data on at least 54 days (out of 60) as long as missing days are not clumped at the end of the study period. In total, 78 sensors functioned for the entire period from October 1–November 30, 2019. Our dataset consists of 787,121 five-minute observations between 6:00 pm–8:00 am and contains the number of individuals passing the sensor in each five-minute period. The ETH Zurich Ethics Commission approved the sensor data collection (EK 2019-N-19). For more about the sensors, see Supplementary File.

We compare the sensor data with predictions from two theories. We focus on route optimization using shortest path analysis and space syntax using the choice measure because our data contain no personal or environmental characteristics.

3.2. Shortest Path Analysis

Shortest path analysis is the simplest application of route optimization theory and is frequently operationalized using Dijkstra's shortest path algorithm (Golledge, 1995; Law & Traunmueller, 2018), which calculates the shortest distance between any two nodes (origin-destination pairs) in a network using street length as edge weights (Dijkstra, 1959).

First, we chose origin-destination pairs relevant to most residents in the community, which we know based on our household survey. All the mapped doors on each structure are considered origins ($N = 869$) and three sites within the informal settlement are considered destinations: (1) the largest Spaza shop (convenience store) in the center of the settlement (near a large block of public toilets); (2) an entrance/exit near the western high-mast public light (one of two 30–40m tall sodium vapor flood lights) that residents use to access shopping and transportation in the formal area; and (3) the main eastern entrance/exit, which connects to a formal road with shops, minibus drop-off points, and a high school (see Figure 1).

We use the QGIS Shortest Path algorithm (version 3.10) to calculate the shortest path between each origin and destination in the network. We then compute the number of times each segment is part of the shortest path for each origin-destination pair to determine the most frequently used path segments for each scenario then sum the three scenarios (Figure 1 in Supplementary File) to come up with our calculation (Law & Traunmueller, 2018).

3.3. Space Syntax (NACH)

Space syntax techniques for analyzing human activity assume that "what happens in any individual space…is fundamentally influenced by the relationships between that space and the network of spaces to which it is connected" (UCL Space Syntax, 2021). The two main measures used to describe the relationship between the spatial network and pedestrian activity are *integration* and *choice*. We focus on choice because it is most comparable to the shortest path analysis (see 3.2). Space syntax choice is a measure of the frequency with which a segment is on the shortest path between all path segments within a prespecified distance, called the *radius* (*r*). Here, the "shortest path" is not the metric shortest path, but rather the path with least angular deviation. The measure can be calculated at different radii, with higher values capturing global movement flows along major arteries and smaller radii capturing local ones (van Nes & Yamu, 2021). When the *r* equals *n*, the entire study area is taken into account—the global measure.

We calculate the space syntax measure of normalized choice (NACH) to enable comparability with other studies (van Nes & Yamu, 2021). We use depthmapX (0.35b), developed by the Space Syntax Group at the University College London, to run an angular segment analysis on the informal settlement route network, plus the immediately surrounding formal roads. Angular segment analysis accounts for curved streets and other irregularities by assuming pedestrians are influenced both by network connectivity and by angles in the network and that they choose routes that minimize direction changes (angular deviation; Dalton, 2003; van Nes & Yamu, 2021). These assumptions are particularly relevant in an informal settlement, where angular deviations can be common, but unintentional.

While studies often report NACH at several radii, we focus on radii informed by our local knowledge and by the empirical data we gathered. The intent is to avoid spurious correlations by testing many radii and finding a correlation by chance—a critique of the space syntax literature (Ericson et al., 2021). First, we know from extensive observation and discussions with residents that relatively few people who do not live in this informal settlement use the smaller path segments. Therefore, to focus on movement within the settlement, we aimed to define a radius that would encompass the neighborhood. Since half the length of the horizontal central street is approximately 150 m, we use this as a heuristic to select a radius of 150 m (and therefore a diameter that captures the widest part of the area), hypothesizing that this radius will best represent "local" movement. Second, we also learned from residents that the formal streets included in our study area just outside the informal settlement are also used by residents because (a) some shared sanitation is located just outside the settlement and (b) these streets are sometimes used as an alternative to the smaller paths because some residents feel they are safer. Therefore, we also use the global measure, $r = n$, based on the notion that n radius captures movement crossing the "border" onto formal streets and helps identify edge effects (Gil, 2015).

4. Results

4.1. Sensor-Measured Path Usage

Over the study period, average five-minute motion between 6:00 pm and 8:00 am is 1.99 triggers, i.e., passers-by, or about 23.9 triggers per hour (minimum: 0.09; maximum: 7.36). Averages in the evening (6:00–9:00 pm) with 55.4 triggers and the morning (5:00–8:00 am) with 28.3 triggers per hour are higher. Movement on weekends is higher than on weekdays (Table 1).

Table 1. Summary statistics.

Statistic	Mean	St. Dev.	Min	Max	Hourly Mean
Sensor					
Avg. 5-min Motion	1.99	1.15	0.09	7.36	23.95
Avg. Evening 5-min Motion (6–9 pm)	4.62	2.06	0.16	9.99	55.42
Avg. Morning 5-min Motion (5–8 am)	2.36	1.79	0.16	8.10	28.30
Avg. Weekday 5-min Motion (Mon–Fri)	1.92	1.17	0.09	7.83	23.00
Avg. Weekend 5-min Motion (Sat/Sun)	2.20	1.19	0.10	6.11	26.39

Note: $N = 78$ sensors/path segments.

Figure 2 maps the path-level five-minute averages for each time interval: all hours (6:00 pm–8:00 am), evening (6:00–9:00 pm), morning (5:00–8:00 am), weekdays (Mon–Fri), and weekends (Sat/Sun). Since the theory-driven calculations are not directly translatable to our empirical measures, for all maps we split the data into tertiles—low (blue), medium (yellow), and high (red)—each with an equal number of path segments. Path segments without a sensor are black if they are within the informal settlement and gray if they are in the formal area.

The highest-activity paths vary across time intervals. Morning and evening high-use paths are noticeably different, not just because there are fewer in the morning, but also because several paths that are high use in the morning are not in the evening, possibly because different activities take place. For example, social activities may cause more night triggers and work-related movements more morning triggers. When we calculate the correlation coefficient between morning and evening activity, we find a correlation coefficient of only 0.42 ($p < 0.01$, $R^2 = 0.17$). Weekday and weekend activity are not hugely different. There is a strong correlation between weekend and weekday path usage (correlation coefficient: 0.92; $p < 0.01$, $R^2 = 0.84$), however, there are more high-use paths on the weekend since more people are home. Entrances to the informal settlement seem to always be in high use, reflecting interaction between this neighborhood and its formal surrounds.

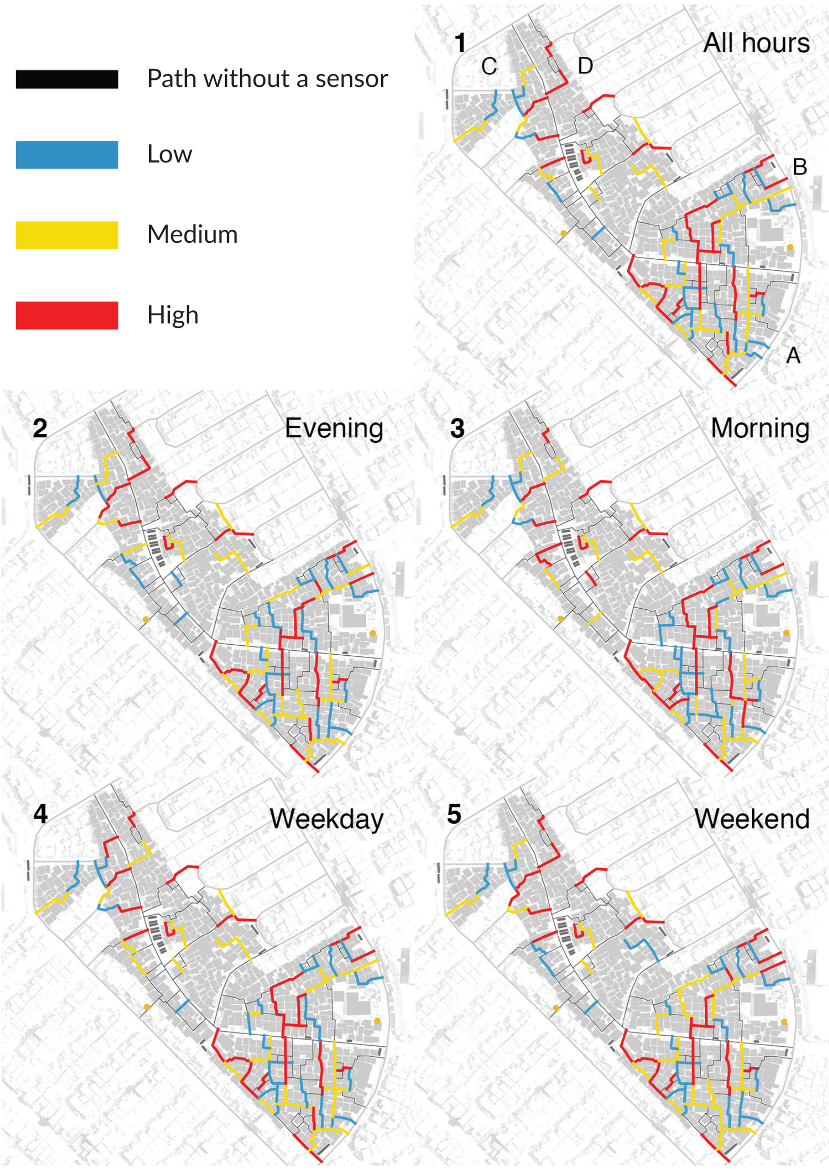

Figure 2. Path usage by time period. For all five maps, the data is split into tertiles—low use (blue), medium use (yellow), and high use (red)—each with an equal number of path segments.

4.2. Path Usage Predictions From Shortest Path and Space Syntax (NACH) Analysis

Figure 3 compares the empirical data (Panel 1) with the theory-derived predictions (Panels 2, 3, and 4). Comparing the shortest path analysis (Panel 2) to the sensor measurements (Panel 1), the predictions are reasonably close to the sensor data in sections A and B, where the network is complicated, but slightly underpredict activity in C and D sections, where the network is sparse. The space syntax NACH calculation (r = 150 m, Panel 3), however, overpredicts activity in A and B and underpredicts in D section. Both theory-driven calculations fall short at the entrances, which are the gateways to the formal area. All entrances into the informal settlement from the formal areas are either medium (yellow) or high (red) usage according to sensor measurements, but neither theory-driven calculation predicts these path segments to be high use. In Panel 4, the NACH (r = n) map does predict high usage for all entrances, but overestimates usage at several entrances, likely contributing to the fact that space syntax NACH ($r = n$) is also not correlated with the empirical data.

Figure 3 maps the two theory-driven predictions, but excludes unmonitored paths for which we have no sensor data. For the predictions including all paths see Figure 1 in Supplementary File and for summary statistics of model predictions see Table 1 in Supplementary File.

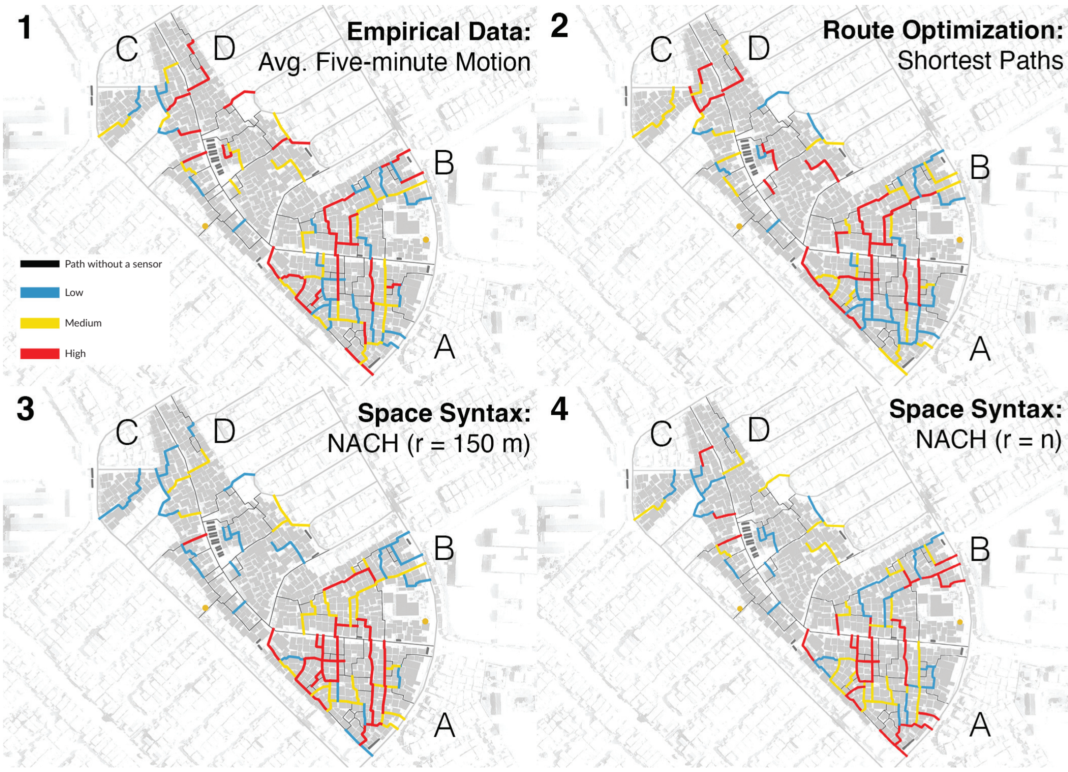

Figure 3. Visual comparison between sensor-measured path usage and theoretical predictions.

4.3. Correlation Between Predictions and Observed Path Usage

Van Nes and Yamu (2021, p. 62) argue the space syntax literature demonstrates that angular segment analysis appears to be the best predictor of pedestrian movement, while "[m]etric distance is a distant third." We test this hypothesis by comparing one element of angular segment analysis, normalized choice, and metric distance (via shortest path analysis) in the informal settlement with sensor data by calculating a Pearson pairwise correlation coefficient and the R^2, since the latter is sometimes reported in the space syntax literature and commonly used across social science disciplines. Although existing pedestrian movement theories do not differentiate between times of the day/week, we compare the theoretical predictions of pedestrian activity to the overall average (6:00 pm–8:00 am), as well as the evenings (6:00–9:00 pm), early mornings (5:00–8:00 am), weekdays (Mon–Fri), and weekends (Sat/Sun) separately, to check whether time plays a role in the theoretical predictive power.

The correlation coefficients in Table 2 confirm the impressions from Figure 3. There is generally low correlation between theory-predicted and sensor-measured usage patterns. The highest correlation is between the shortest path calculation and evening (6:00–9:00 pm) average motion ($R = 0.39$, $p < 0.01$). The correlations between the shortest path calculation and overall average motion (6:00 pm–8:00 am), weekday motion (Mon–Fri), and weekend motion (Sat/Sun) are all smaller. We find no statistically significant correlation between either space syntax NACH calculation ($r = 150$ and $r = n$) and the empirical data for any time.

Table 2. Correlation between average five-minute motion in different time intervals and theory-driven calculations.

	Avg. 5-Min Motion				
	All (1)	Evening (2)	Morning (3)	Weekday (4)	Weekend (5)
Shortest Paths					
Avg. Count	0.25/0.06 (0.03)	0.39/0.15 (0.00)	0.15/0.02 (0.20)	0.23/0.05 (0.04)	0.29/0.08 (0.01)
Space Syntax					
Norm. Avg. Choice ($r = 150$ m)	0.01/0.00 (0.92)	0.16/0.03 (0.15)	−0.05/0.00 (0.63)	−0.01/0.00 (0.93)	0.07/0.00 (0.57)
Norm. Avg. Choice ($r = n$)	−0.08/0.01 (0.50)	0.05/0.00 (0.69)	−0.11/0.01 (0.34)	−0.08/0.01 (0.48)	−0.06/0.00 (0.59)

Notes: The table reports Pearson pairwise correlation coefficient (R) on the left of the slash and the R^2 on the right; p-value in parentheses. Statistical significance is 0.05 or less.

5. Discussion

The few papers on pedestrian activity in informal settlements have used space syntax (not route optimization) to predict pedestrian flows (Hidayati et al., 2020; Mohamed, 2016). However, we find the space syntax (normalized) choice measure is not a good predictor of pedestrian activity in informal settlements. Comparing choice measures (150-m and *n* radii) and a shortest (metric) path prediction of most frequently used path segments to empirical data, we find the shortest path prediction explains more of the variation observed with sensors. Moreover, the visual comparison of the shortest path predications and sensor averages reveals that the predictions perform better in areas of the informal settlement with more complicated route configurations compared to simpler ones. Furthermore, the theory-driven measures poorly predict the level of activity recorded at the entrances of the settlement, which are important connection points at the de facto border between informal and formal urban spaces.

Our study is one of three to find weak correlations between choice predictions and empirical measures in informal settlements (see Hidayati et al., 2020; Mohamed, 2016). In formal cities, the correlation coefficient between space syntax and manual pedestrian counts has been extensively studied, but also varies (Ericson et al., 2021). While a meta-analysis of 14 studies using various space syntax measures to study pedestrian movement confirms that on average it predicts pedestrian movement in neighborhoods of high-income countries well (Sharmin & Kamruzzaman, 2018), the correlation coefficient between space syntax choice and counted pedestrian movement varied between $R = -0.141$ (Sodermalm, Stockholm) and $R = 0.885$ (Bakirköy, Istanbul).

The notion that informal settlements, or unplanned settlements, as scholars in the space syntax literature also refer to them, have unique characteristics (e.g., irregularity) that might mediate the relationship between the network and pedestrian activity has already been acknowledged in the space syntax literature (Karimi & Parham, 2010, 2012). Our study points to an important insight: space syntax's choice measure may be predictive when compared to observed data from cities that *include* informal settlements, i.e., encompass the boundary between the formal and the informal, but it may lose predictive power when studying a local informal network that is influenced by the porous, but poorly understood border between formal and informal—thus, performing even worse than shortest paths analysis. To our knowledge, no previous study has compared shortest path predictions and space syntax's choice measure with empirical data in informal settlements. In the informal settlement we studied, six possible explanations come to mind why space syntax, in particular, performs poorly in predicting empirically measured pedestrian activity.

First, is the issue of extreme density in many informal settlements (Visagie & Turok, 2020). The space syntax choice measure does not account for path width, however, some path segments are so narrow that a pedestrian must turn sideways to pass. Therefore, although a segment may appear geometrically, topologically, and metrically ideal, few pedestrians may use it. Ignoring path width also explains some shortcomings of the shortest path analysis. The ability to fit through a path is not an issue in formal cities, but it is important in this context in these inherently unplanned and emergent neighborhoods.

Second, Hillier et al. (1993, p. 31) write that "configuration may affect movement, but configurational parameters cannot be affected by it." In informal settlements, that assumption is invalid. For example, as we learned through the mapping process, the residents in the informal settlement we study frequently create new cul-de-sacs, called compounds, by blockading a path to limit through-traffic and enhance their sense of safety, demonstrating that movement *can* affect configuration.

Third, the residents living in our study site seem to rely on a mental map of landmarks to navigate. Our community team reported that residents try to reach a central street or formal road as quickly as possible. This strategy makes sense if a pedestrian wants to avoid narrow paths, feels fearful in smaller paths, and/or if smaller paths are darker at night—all legitimate possibilities in the area we study. This strategy to quickly reach main arteries may counterbalance the lack of distance information, making shortest route distances a better predictor of pedestrian mobility. It could also explain why entrance paths (see Figure 3) are used more than both theories predict and highlights the complexity of pedestrian behavior at the interaction between formal and informal areas. That said, we cannot rule out that poor prediction at entrances/exits is linked to the study area size (Gil, 2015).

A fourth explanation why shortest path analysis might be more predictive could be that residents report that they typically only know the immediate area where they live and rarely use paths in other parts of the settlement. Hence, the critique that residents lack the information necessary to optimize route choice might be misplaced in this context, as most pedestrians know their immediate surroundings well. Furthermore, it may be that the majority of motion we measure is generated by those that have sufficient experience to recognize the shortest path, compared to areas in which many pedestrians lack familiarity with network (common in formal cities).

Fifth, the presence of shared essential infrastructure, as well as the differences in perceptions about the use of public and private space in informal vs. formal areas (Cutini et al., 2020; Sletto & Palmer, 2017), also may substantially shift the motion patterns measured by the sensors in ways that are not accounted for in either space syntax or shortest path analysis. Since space syntax so heavily emphasizes network configuration, this may lead to even more serious mis-estimation of pedestrian motion.

A final reason lies in the assumptions behind the calculations. Angular segment analysis accounts for direction change, whereas the shortest path analysis measures the shortest metric distance between origin-destination pairs. While emphasizing directional change makes sense in formal cities, where such features are usually intentional, in this network paths often have superfluous bends (e.g., caused by a house with ingress into the pathway) wherein a pedestrian might not perceive a meaningful direction change, but angular deviation in the network is captured in the calculation (see Figure 1). Therefore, some segments may be unduly penalized when using such a predictive approach in an informal settlement. Like Karimi's (2002) study of Iranian "organic cities," adjustments to the space syntax methodology might be necessary to be more predictive in informal settlements.

Furthermore, by comparing the theory-driven measures to average motion in the evening and morning, we highlight what might be an important variable these theories do not address: time. Correlation coefficients were highest for the evening average and lowest for the morning average, indicating that the distribution of activity is not similar. One explanation could be that evening activity in informal settlements is representative of a broader range of activities, including evening commuting, shopping, using shared sanitation, and socializing in public space (Cutini et al., 2019, 2020; Kamalipour, 2020), while morning activity may be limited in ways that create differential patterns, as people are starting their day (e.g., using the toilet, commuting, but perhaps not socializing as much).

We caveat our findings by pointing out that we do not have sensor data for all path segments, whereas the theory-driven calculations are based on the entire network no matter whether we monitored the segment or not. Although we addressed this by only including path-segment level observations from the shortest path and choice calculations for which we also have sensor data, we cannot rule out that the missing sensor data would not change correlation coefficients. Furthermore, the sensors inaccurately measure pedestrian motion during peak daylight hours, meaning we do not measure all of the time windows space syntax methodology recommends for manual measurement (van Nes & Yamu, 2021). However, since this informal settlement is primarily residential, evenings and mornings are when most people are around. In addition, we only use three scenarios to construct the shortest paths calculation, thus it is possible that considering more scenarios would affect correlation coefficients. Finally, other space syntax studies frequently analyze larger study areas. We selected this area since we focus only on pedestrian, not vehicular activity, however, it is possible that this impacts the results, since we cannot say how much activity outside the settlement influences what we measure inside.

6. Conclusion

Using pedestrian motion sensors to measure activity in the evenings and early mornings, we analyzed mobility patterns in an informal settlement and tested how two common, and in formal cities frequently tested, theory-driven approaches to predicting pedestrian route choice perform in comparison. The sensor data allow us to analyze, at high granularity over a longer period, the extent to which previous findings about these theories hold in an informal settlement. We find that space syntax choice measures are not correlated with pedestrian activity in the informal settlement we study. Route optimization theory seems to perform better. Interestingly, to our knowledge, no previous study has used route optimization to predict pedestrian activities in informal settlements. When we analyze the sensor data alone, we find evening and morning average motion are different, whereas weekend and weekday average motion are strongly correlated, though weekend activity is more intensive.

These findings open important considerations for future work to understand what drives pedestrian activity and route choice in informal settlements. Empirically-measured pedestrian activity may deviate from existing theories that only account for route network characteristics for several reasons, including extreme density and narrow path width, constantly changing network configuration, pedestrian reliance on mental maps absent formal ones, measured activity may be dominated by residential activity rather than through-traffic, the need to go outside to access shared sanitation services, and finally time-related differences in activities. Therefore, there may be a limit to the explanatory power of network-focused theories, compared to formal areas, necessitating theories that directly consider the nature of pedestrian behavior in informal neighborhoods and how formal areas influence it.

Informal settlements are a large and growing share of the urban environment throughout Africa, Asia, and Latin America. Thus, understanding pedestrian activity dynamics can provide insights to improve public service accessibility. For example, fires are a common danger. With more information about pedestrian activity, communities may be better equipped to improve access and escape routes. For city officials, understanding pedestrian patterns could enable more informed upgrading decisions, such as where to site sanitation facilities or streetlights as well as where to pave pathways.

Our study also shows that sensors are a minimally invasive method for collecting high-frequency, long-term pedestrian activity data, especially at night or in dense areas where movement data cannot be collected with mobile phones. Despite the challenges we encountered deploying this technology in an informal settlement for the first time, we gathered detailed data about nighttime pedestrian activity. Future research could improve pedestrian motion sensors to more effectively find ways to improve service delivery, disaster response, and quality of life in informal settlements.

Finally, our work reveals that the de facto border between the formal and the informal in urban areas exists not only in physical reality, but also in research. Sensors have been used to measure pedestrian activity in formal areas, but until now, not in informal ones; route optimization has been tested in formal parts of cities, but not in informal settlements. This research highlights the need for a new or modified theory of route selection based on a more nuanced understanding of this physical and abstract border that also better incorporates the specific characteristics of informal urban settlements along with the characteristics of formal areas.

Acknowledgments

The authors would like to thank Yamkela Rongwana, Sibongile Mvumvu, and Jennifer Qongo for outstanding sensor data collection and maintenance, Noliyema Swartbooi and the Dora Educare Center for providing storage space, the leadership committee for collaboration on the research, and Sensen for co-developing the sensors. We would further like to thank Erwin LeFoll for excellent research assistance, Dr. Christiaan Joubert for helpful guidance on implementing the shortest path analysis in QGIS, and Dr. Subik Shrestha, Dr. Beatrix Emo, and Ruth Nelson for thoughtful comments on the space syntax analysis.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Anciaes, P. R., Nascimento, J., & Silva, S. (2017). The distribution of walkability in an African city: Praia, Cabo Verde. *Cities*, *67*, 9–20. https://doi.org/10.1016/j.cities.2017.04.008

Bongiorno, C., Zhou, Y., Kryven, M., Theurel, D., Rizzo, A., Santi, P., Tenenbaum, J., & Ratti, C. (2021). Vector-based pedestrian navigation in cities. *Nature Computational Science*, *1*(10), 678–685. https://doi.org/10.1038/s43588-021-00130-y

City of Cape Town. (2018). *Aerial imagery 2018 Feb*. https://odp.capetown.gov.za/documents/raster-imagery/explore

City of Cape Town. (2019). *Informal settlement boundaries*.

Cutini, V., Di Pinto, V., & Rinaldi, A. M. (2019). Using space syntax and geographic information systems. In S. Misra, O. Gervasi, B. Murgante, E. Stankova, V. Korkhov, C. Torre, A. M. A. C. Rocha, D. Taniar, B. O. Apduhan, & E. Tarantino (Eds.), *Computational science and its applications – ICCSA 2019* (Part III) (pp. 343–356). Springer. https://doi.org/10.1007/978-3-030-24302-9

Cutini, V., Di Pinto, V., Rinaldi, A. M., & Rossini, F. (2020). Proximal cities: Does walkability drive informal settlements? *Sustainability*, *12*(3), 1–20. https://doi.org/10.3390/su12030756

Dalton, R. C. (2003). The secret is to follow your nose: Route path selection and angularity. *Environment and Behavior*, *35*(1), 107–131. https://doi.org/10.1177/0013916502238867

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271. https://doi.org/10.1007/BF01386390

Dodman, D., Archer, D., & Mayr, M. (2018). *Addressing the most vulnerable first: Pro-poor climate action in informal settlements*. UN-Habitat. https://unhabitat.org/sites/default/files/2019/05/pro-poor_climate_action_in_informal_settlements-.pdf

Ericson, J. D., Chrastil, E. R., & Warren, W. H. (2021). Space syntax visibility graph analysis is not robust to changes in spatial and temporal resolution. *Environment and Planning B: Urban Analytics and City Science*, *48*(6), 1478–1494. https://doi.org/10.1177/2399808319897624

Gil, J. (2015). Examining "edge effects": Sensitivity of spatial network centrality analysis to boundary conditions. In K. Karimi (Ed.), *SSS 2015—10th International Space Syntax Symposium* (pp. 1–16). International Space Syntax Symposium.

Golledge, R. G. (1995). Path selection and route preference in human navigation: A progress report. In A. U. Frank & W. Kuhn (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 207–222). Springer. https://doi.org/10.1007/3-540-60392-1_14

Hidayati, I., Yamu, C., & Tan, W. (2020). Realised pedestrian accessibility of an informal settlement in Jakarta, Indonesia. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, *14*(4), 434–456. https://doi.org/10.1080/17549175.2020.1814391

Hillier, B. (2007). *Space is the machine: A configurational theory of architecture*. Space Syntax.

Hillier, B., Greene, M., & Desyllas, J. (2000). Self-generated neighbourhoods: The role of urban form in the consolidation of informal settlements. *Urban Design International*, *5*(2), 61–96. https://doi.org/10.1057/palgrave.udi.9000018

Hillier, B., & Iida, S. (2005). Network and psychological effects in urban movement. In A. G. Cohn & D. M. Mark (Eds.), *Spatial information theory* (pp. 475–490). Springer. https://doi.org/10.1007/11556114_30

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: Or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: Urban Analytics and City Science*, *20*(1), 29–66. https://doi.org/10.1068/b200029

Kamalipour, H. (2020). Improvising places: The fluidity of space in informal settlements. *Sustainability*, *12*(6), Article 2293. https://doi.org/10.3390/su12062293

Kamalipour, H., & Dovey, K. (2019). Mapping the visibility of informal settlements. *Habitat International*, *85*, 63–75. https://doi.org/10.1016/j.habitatint.2019.01.002

Karimi, K. (2002). Iranian organic cities desmystified: A unique urban experience or an organic city like others. *Built Environment*, *28*(3), 187–201.

Karimi, K., & Parham, E. (2010). *Slums and informal settlements: An evidence-based approach to sustainable upgrading and development*. Space Syntax.

Karimi, K., & Parham, E. (2012). An evidence informed approach to developing an adaptable regeneration programme for declining informal settlements. In M. Greene, J. Reyes, and A. Castro (Eds.), *Proceedings: Eighth International Space Syntax Symposium* (pp. 1–27). Pontificia Universidad Católica de Chile.

Kretzer, D. M. (2020). High-mast lighting as an adequate way of lighting pedestrian paths in informal settlements? *Development Engineering*, *5*, Article 100053. https://doi.org/10.1016/j.deveng.2020.100053

Law, S., & Traunmueller, M. (2018). Off the shortest path: Betweenness on street network level to study pedestrian movement. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems* (pp. 971–976). IEEE. https://doi.org/10.1109/ITSC.2017.8317950

Mehta, V. (2008). Walkable streets: Pedestrian behavior, perceptions and attitudes. *Journal of Urbanism*, *1*(3), 217–245. https://doi.org/10.1080/17549170802529480

Mohamed, A. A. (2016). People's movement patterns in space of informal settlements in Cairo metropolitan area. *Alexandria Engineering Journal*, *55*(1), 451–465. https://doi.org/10.1016/j.aej.2015.07.018

Mutyambizi, C., Mokhele, T., Ndinda, C., & Hongoro, C. (2020). Access to and satisfaction with basic services in informal settlements: Results from a baseline assessment survey. *International Journal of Environmental Research and Public Health*, *17*(12), Article 4400. https://doi.org/10.3390/ijerph17124400

Ndifuna Ukwasi, Social Justice Coalition, IBP, & OpenUp. (n.d.). *Struggle for dignity in Cape Town's informal settlements: The facts*. https://ismaps.org.za/desktop.html

Salazar Miranda, A., Fan, Z., Duarte, F., & Ratti, C. (2021). Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment. *Computers, Environment and Urban Systems*, *86*, Article 101563. https://doi.org/10.1016/j.compenvurbsys.2020.101563

Salon, D., & Gulyani, S. (2010). Mobility, poverty, and gender: Travel "choices" of slum residents in Nairobi, Kenya. *Transport Reviews*, *30*(5), 641–657. https://doi.org/10.1080/01441640903298998

Samper, J., & Liao, W. (2023). Testing the informal development stages framework globally: Exploring self-build densification and growth in informal settlements. *Urban Science*, *7*(2), Article 50. https://doi.org/10.3390/urbansci7020050

Sharmin, S., & Kamruzzaman, M. (2018). Meta-analysis of the relationships between space syntax measures and pedestrian movement. *Transport Reviews*, *38*(4), 524–550. https://doi.org/10.1080/01441647.2017.1365101

Shatu, F., Yigitcanlar, T., & Bunker, J. (2019). Shortest path distance vs. least directional change: Empirical testing of space syntax and geographic theories concerning pedestrian route choice behaviour. *Journal of Transport Geography*, *74*, 37–52. https://doi.org/10.1016/j.jtrangeo.2018.11.005

Simone, A., & Pieterse, E. (2017). *New urban worlds: Inhabiting dissonant times*. Polity Press.

Sletto, B., & Palmer, J. (2017). The liminality of open space and rhythms of the everyday in Jallah Town, Monrovia, Liberia. *Urban Studies*, *54*(10), 2360–2375. https://doi.org/10.1177/0042098016643475

UCL Space Syntax. (2021). *Overview*. https://www.spacesyntax.online/overview-2

van Nes, A., & Yamu, C. (2021). *Introduction to space syntax in Urban Studies*. Springer. https://doi.org/10.1007/978-3-030-59140-3

Visagie, J., & Turok, I. (2020). Getting urban density to work in informal settlements in Africa. *Environment & Urbanization*, *32*(2), 351–370. https://doi.org/10.1177/0956247820907808

Walls, R. S., Eksteen, R., Kahanji, C., & Cicione, A. (2018). Appraisal of fire safety interventions and strategies for informal settlements in South Africa. *Disaster Prevention and Management*, *28*(3), 343–358. https://doi.org/10.1108/DPM-10-2018-0350

Willis, A., Gjersoe, N., Havard, C., Kerridge, J., & Kukla, R. (2004). Human movement behaviour in urban spaces: Implications for the design and modelling of effective pedestrian environments. *Environment and Planning B: Urban Analytics and City Science*, *31*(6), 805–828. https://doi.org/10.1068/b3060

Yamu, C., van Nes, A., & Garau, C. (2021). Bill Hillier's legacy: Space syntax—A synopsis of basic concepts, measures, and empirical application. *Sustainability*, *13*(6), Article 3394. https://doi.org/10.3390/su13063394

Zhu, S., & Levinson, D. (2015). Do people use the shortest path? An empirical test of Wardrop's first principle. *PLoS ONE*, *10*(8), Article e0134322. https://doi.org/10.1371/journal.pone.0134322

About the Authors

Yael Borofsky is an early-career fellow at the Collegium Helveticum and an affiliate post-doctoral researcher in the Development Economics Group at ETH Zürich where she studies infrastructure access in informal settlements and the use of sensors for social science research in data-scarce settings. Yael received her PhD from ETH Zürich in 2022. She has a dual masters' degree in city planning as well as technology and policy from the Massachusetts Institute of Technology.

Stephanie Briers is an architect by training, an urbanist by practice, and a generalist by nature, bridging research and practice in complex urban environments. She is specialised in stakeholder engagement and transdisciplinary research, where she utilises creative practices to enrich the research process. Stephanie holds a PhD in architecture from ETH Zürich, where she did an action research project on public lighting inequalities in South Africa. She is currently a post-doctoral researcher at the Transdisciplinarity Lab at ETH Zürich.

Isabel Günther is professor of development economics and academic director of NADEL at ETH Zürich. She was previously a research associate at the Faculty of Economics of the University of Göttingen and the Harvard School of Public Health. Her research focus is empirical microeconomics, particularly the measurement of poverty and inequality, population economics, technologies for poverty reduction, and evidence-based policymaking. She has carried out most of her research in Benin, Burkina Faso, Ghana, Kenya, Uganda, and South Africa.