Space Syntax Visibility Graph Analysis is not robust to changes in spatial and temporal resolution

Jonathan D Ericson
Bentley University, USA

Elizabeth R Chrastil
University of California, Santa Barbara, USA

William H Warren
Brown University, USA

Corresponding author:

Jonathan D Ericson
Department of Information Design and Corporate Communication
Bentley University
175 Forest Street
Waltham, MA 02452, USA.
Email: jericson@bentley.edu

Abstract

Space Syntax is an influential framework for quantifying the relationship between environmental geometry and human behavior. Although many studies report high syntactic-behavioral correlations, previous pedestrian data were collected at low spatiotemporal resolutions, and data transformations and sampling strategies vary widely; here, we systematically test the robustness of Space Syntax's predictive strength by examining how these factors impact correlations. We used virtual reality and motion-tracking to correlate 30 syntactic measures with high resolution walking trajectories downsampled at ten grid resolutions, and subjected to various log transformations. Overall, correlations declined with increasing grid resolution and were sensitive to data transformations. Moreover, simulations revealed spuriously high correlations (e.g. $R^2 = 1$) with sparsely sampled data (< 23 locations). These results strongly suggest that syntactic-behavioral correlations are not robust to changes in spatiotemporal resolution, and that high correlations obtained in previous studies could be inflated due to transformations, data resolution, or sampling strategies.

Keywords: space syntax, virtual reality, spatial behavior, simulation, architecture

Introduction

Space syntax (Hillier, 1999a; Hillier et al., 1996; Hillier and Hanson, 1984) is a prominent framework for examining the relationship between human behavior and the geometry of the built environment. Syntactic measures are frequently used to quantify configurational properties of urban and architectural environments, and many measures have been shown to correlate with patterns of human movement and usage in a variety of settings, including museums (Batty, 2001; Turner and Penn, 2002), malls (Okamoto et al., 2013; Omer and Goldblatt, 2017), conference halls (Mashhadi et al., 2016), hospitals (Haq and Luo, 2012), and large-scale urban spaces (Turner, 2003). However, previous studies have examined relatively coarse pedestrian flow data collected at *low* spatial and temporal resolutions. As a result, the predictive strength of syntactic measures has not been systematically examined across a variety of spatial and temporal scales. If space syntax is to serve as a robust predictive framework, the strength of its correlations with pedestrian movement patterns should be robust to changes in the spatial and temporal resolution at which correlations are computed. To address this critical gap in the literature, the present study investigated the robustness of correlations between space syntax measures and *high* spatial and temporal resolution human walking data over a wide range of spatiotemporal resolutions.

Overview of space syntax methods

This section provides an overview of space syntax methods and terminology with a focus on a subset of issues that are relevant to the present study (for more comprehensive introductions to space syntax concepts and terminology, see Bafna, 2003, and Klarqvist, 1993).

Visibility graph analysis (VGA). In the present study, we focus on syntactic measures that can be computed from “visibility graphs” (Turner, 2003; Turner et al., 2001). A visibility graph is generated by superimposing a grid on a top-down view of a space (e.g., from a CAD drawing) using space syntax software. Syntactic measures (e.g., *isovist area*) can then be computed for each grid cell using topological methods for quantifying the proximity and inter-visibility of grid cells (Turner, 2003). The resulting analyses and visualizations are frequently used to examine configurational properties of spaces that may be relevant for predicting human behavior. *Integration*, a commonly cited visibility graph measure, is examined in detail in the present study. The Integration value for a cell is obtained by computing the average depth (i.e., topological distance) of that cell to neighboring cells within a specified topological distance, effectively ranking cells "from the most integrated to the most segregated" (Klarqvist, 1993). A large number of other syntactic measures have been derived from visibility graphs, yet many of these measures have not been systematically investigated in a single study.

Pedestrian data. Previous studies have generally collected *low* spatiotemporal resolution pedestrian data, often manually. Typical methods employed in these studies include (Conroy, 2001): (1) *gate counts*, "the cumulative number of people passing over a specified 'threshold' within a given timeframe"; (2) *spot counts*, "the approximate location of all people present in any room or space at a given moment in time"; (3) *occupancy numbers*, "the total numbers of people present in a single space/room during specific time intervals throughout the day"; and (4) *movement traces*, most often obtained by researchers transcribing sketched approximations of paths taken as pedestrians walk through a space. The primary contribution of the present study is to examine correlations with continuous *high* spatiotemporal resolution movement data. A key advantage of virtual reality systems leveraged in the present study is the ability to automate collection of high resolution movement trajectories using electronic motion tracking systems.

Criticisms of space syntax

Space syntax remains the subject of lively debate. A number of thorough criticisms (Batty, 2001; Jguirim et al., 2014; Kostakos, 2010; Montello, 2007; Netto, 2016; Pafka et al., 2018; Peponis et al., 1997; Ratti, 2004a, 2004b; Turner et al., 2001), rebuttals (e.g., Hillier & Penn, 2004), and counter-rebuttals (e.g., Ratti, 2004a) to these criticisms have been advanced. Prior empirical studies investigating the relationship between space syntax measures and human movement report a range of correlation values from ranging from weak (e.g., .142; Mora, Astudillo, & Bravo, 2014) to strong (e.g., .98; Penn, Hillier, Banister, & Xu, 1998). As a result, the predictive value of space syntax across a variety of spatial scales is in question, and a review of the literature reveals a number of important criticisms and limitations of prior research. The present study focuses on evaluating a subset of criticisms of space syntax with a focus on methodological limitations of previous empirical studies. In the following sections, we discuss each of these limitations in turn, and describe our experimental approach.

Data transformations and correlation methods. Table 2 (see Supplemental Material: Tables) summarizes the results from a selection of previous studies (de Arruda Campos, 1997; Desyllas and Duxbury, 2001; Hillier et al., 1996; Mora et al., 2014; Okamoto et al., 2013; Penn et al., 1998; Turner, 2003; Turner and Penn, 1999). Taken together, these studies report a wide range of syntactic-behavioral correlations, yet many studies do not report whether data transformations were applied to pedestrian or configurational data, or do not report how correlations were computed. As a result, replicating previous studies has been hindered, casting doubt on whether high correlations may be a function of methodological decisions. In an unpublished space syntax software manual, Turner (2014) notes that taking the natural logarithm of observed pedestrian data is accepted practice when the data are not normally distributed, and that correlations of $R^2 = .40$ between space syntax measures and log-transformed (ln) observed variables are “expected in space syntax theory,” though the reasons for this expectation are not made clear. In contrast, Penn, Hillier, Banister, and Xu (1998) investigated correlations between configurational properties of street networks and flow rates (both pedestrian and vehicular) and used a data transformation not often used in other studies (taking the fourth root of flow rates). Sometimes, data transformations are also applied to syntactic measures themselves. For example, Turner (2003) averaged gate counts from a previous study and applied a log transformation to agent simulation data. While data transformations are sometimes justified, inconsistent application of transformations is an important limitation of previous work. Finally, some authors (e.g., Turner, 2003) report R^2 as well as significance values for correlations, while others do not. The present study aimed to address these inconsistencies by systematically investigating the effects of various data transformations on correlations for many visibility graph measures computed for a single space.

Spatiotemporal resolution of pedestrian and configurational data. As Desyllas & Duxbury (2001) note, “one of the methodological issues when using VGA is the effect of changing the parameter of sampling grid resolution.” Yet the spatial and temporal resolutions at which human walking data have been collected, as well as the spatial resolution at which configurational data have been computed, have not been comprehensively and consistently reported across studies. In general, pedestrian and configurational data have been examined at the level of meters and minutes rather than at the level of millimeters and milliseconds, limiting our understanding of space syntax’s predictive capabilities at high spatiotemporal resolution. The present study is the first to examine correlations between space syntax measures and human walking data collected at high temporal (~60Hz) and spatial (0.75mm, 0.05º) resolution—on the order of milliseconds and millimeters—and to explore how correlations depend on spatial and temporal binning.

Sampling Strategies. Past studies also vary widely in sampling methods used. For example, Turner and Penn (1999) calculated “mean isovist integration value[s] of nodes within a 1.5m buffer of each gate location.” And as Silva (2013) notes, when studies are conducted in professional contexts, modeling of pedestrian and vehicular traffic patterns is frequently done on an intuitive basis, “following tradition, rather than a statistically sound methodology.” The present study addresses these limitations by clearly reporting how pedestrian and configurational data were collected, by more systematically evaluating the robustness of correlations across a variety of spatial scales and sampling resolutions, and by using a simulation approach to quantify the impact of sampling strategies.

Characteristics of spaces examined. The underlying mathematical formulations of Space Syntax theory embody an implicit hypothesis that human behavior is causally linked to the geometry of the environment. Previous research on the predictive strength of syntactic measures has primarily examined human movement patterns in rooms and urban spaces. If human movement patterns are constrained by the geometry of the environment, one would expect that movement patterns in narrow spaces (e.g., hallways or maze corridors, where movement is relatively constrained) should be highly predictable relative to movement patterns in more open spaces (e.g., interconnected rooms, where movement is relatively unconstrained). Testing the predictive capabilities of syntactic measures by asking participants to navigate restricted maze corridors provides a strong test of space syntax predictions in enclosed spaces.

The present study

In sum, the present study contributes to the literature on space syntax by addressing three key criticisms and limitations of previous work. First, data collection methods are not always comprehensively reported, and details regarding data transformations applied to syntactic and pedestrian data are inconsistently reported; the present study addresses this limitation by (a) thoroughly documenting how the data were collected and transformed, and (b) assessing the impact of data transformations on correlations. We asked whether (Q1) correlations are sensitive to data transformations. Second, pedestrian data have generally been sampled at relatively *low* spatial and temporal resolutions, and prior studies have not systematically examined the impact of sampling grid resolution on syntactic-behavioral correlations; we address this by (a) sampling continuous walking trajectories at high spatial and temporal resolution using a motion tracking system, and (b) examining whether correlations vary as the spatial resolution of the sampling grid is increased. We asked whether (Q2) correlations depend on the spatial resolution of the sampling grid. Third, sampling strategies—in particular, the number of "gates"—used in previous studies vary widely; to address this, we examined how correlations change when an increasing number of randomly selected grid locations are used as the basis for computing correlations. We asked whether (Q3) a small sample of spatial locations would yield spuriously high correlations.

Methods

The present study was conducted at [removed for peer review].

Participants

A total of 36 participants were included in the analysis (18M, 18F). The mean age of participants was 20.8 years (*SD* = 4.35 years). Participants provided written informed consent to participate in the experiment in accordance with [name of university removed for peer review] Institutional Review Board (IRB) requirements, and were paid ($10/hour) for their participation.

Apparatus

Participants walked freely within a 10.5m x 12.5m area in the lab while a tracking system (InterSense IS-900, 1mm linear and 1º angular RMS error, 60 Hz sampling rate) recorded head position and orientation. Stereoscopic images of the virtual environment were presented via a head-mounted display (HMD, Rockwell-Collins SR80A, 1280x1024 pixels, 63° H x 53° V field of view for each eye) calibrated to each participant's inter-ocular distance. Displays were generated on a Dell XPS 730X desktop computer (50ms total latency).

Displays

The virtual environment (Figure 1A) was a hedge maze containing a central "home" location, eight unique target objects located at the ends of maze hallways, and four paintings that provided local landmarks. The environment was created in 3DS Max (Autodesk) and presented to the participant using Vizard (WorldViz, Version 4.0).

[Insert Figure 1]

Procedure

Participants were brought to the center of a virtual hedge maze, instructed to learn the locations of the objects in maze while freely exploring for 10 minutes, and informed that they would be tested on their knowledge of the object locations later in the experiment. When they walked up to an object, an audio file played telling them the name of that object (e.g., "bookcase"). If participants left the 10.5 x 12.5m maze, virtual brick walls appeared to prevent collisions with the lab's walls. Background noise (night sounds) was played over headphones, and a black cloth covered the HMD to block the view of the lab.

Data Analysis

Here, we present an overview of the analytical approaches used to examine each of our primary research questions (for additional detail on each approach, see Supplemental Material: Methods). Both syntactic measures and human walking data were binned at 10 discrete spatial scales in .01m increments to produce grids where the edge lengths of the cells ranged from 0.01m – 1.00m. Walking data were aggregated across all participants, yielding 1.5 million positional data points (x, y, time); raw walking data values (W) were obtained by counting the number of points in each grid cell, and raw syntactic measure (S) values were computed for corresponding grid cells using depthmapX (Version 0.5b; Varoudis, 2015b). To examine whether correlations are sensitive to data transformations (Q1), we applied 11 different data transformations (see Table 1) to all 30 syntactic measures across 10 different spatial resolutions (bin sizes), and compared the performance of each transformation. To examine whether correlations decline with increasing spatial resolution (Q2), we used a regression analysis and plotted R^2 values against bin size (see Figures 2 and 3). To examine whether spurious correlations would be obtained when few relatively few sampling locations (N_{GATES}) within the overall sampling grid are used to compute correlations (Q3), we used a simulation approach in which an increasing number of randomly positioned sampling locations (N_{GATES}) were used to compute correlations for each measure and bin size using the most frequently effective data transformation [$\log_{10}$(W)>0 vs. S] (see Figure 4, and Supplemental Material: Figure 7).

Results

Here, we present results with respect to our three primary research questions. Hillier (1999a) notes that "in most studies the best performing spatial variable is Radius-3 Integration" [denoted *Visual Integration (R3)* in the present study]. In contrast, in the present study, we found that *Metric Node Count (R1)* was the best performing spatial variable. Therefore, for simplicity of presentation, aggregated results are presented for all thirty syntactic measures (see Table 1; Figures 2 and 3), while more detailed results are presented for two key measures: *Visual Integration (R3),* and Metric Node Count (R1) (see Supplemental Material).

Q1: Are correlations sensitive to data transformations?

Determining the best overall data transformation. To identify the most effective overall data transformation, syntactic-behavioral correlations were computed for 30 syntactic measures across 10 sampling grid resolutions (see Methods). For each of the 300 resulting measure-bin size pairs, a series of 11 data transformations (see Table 1) were applied to the data, yielding 3,300 (300 x 11) correlation (R^2) values. The transformation(s) yielding the highest R^2 value(s) for each measure-bin size pair were then tabulated (Table 1, n column) and converted to percentages for each data transformation (Table 1, % column). A summary of this analysis is presented in Table 1. The $log_{10}(W)>0$ *vs. S* transformation produced the highest correlation for 25.9% of the cases examined, followed by $ln(W)>0$ *vs. S* (23.7% of cases examined). The percent score indicates that these two transformations tended to yield the highest correlation values. In comparison, leaving data untransformed (W vs. S) yielded the highest correlation values in only 6.8% of cases examined. In sum, we found that (Q1) syntactic-behavioral correlations were sensitive to data transformations; in particular using logarithmically (log_{10} and *ln*) transformed walking data and untransformed syntactic values yielded the highest syntactic-behavior correlations. Therefore, to ensure that subsequent analyses were charitable toward space syntax measures (without resorting to algorithmically cherry-picking the best data transformation for each specific case, as in the previous analysis) we present $log_{10}(W)>0$ *vs. S* in several of the analyses and figures that follow.

Q2: Do correlations depend on the spatial resolution of the sampling grid?

Part 1: Best data transformation for each measure-bin size pair. For this analysis, we used whichever data transformation yielded the highest correlation in order to obtain maximal R^2 values for each measure-bin pair. For each syntactic measure, a simple linear regression was calculated to predict syntactic-behavioral correlation strength (R^2) as a function of spatial resolution (bin size). We asked whether (Q2) correlations would decline as the spatial resolution of the sampling grid is increased, corresponding to negative regression lines. Results appear in Figure 2.

[Insert Figure 2]

Significant regression equations were found ($p < .05$) for 24 of the 30 (80%) syntactic measures. Regression lines had significantly negative slopes for 22 of the 30 (73%) syntactic measures ($p < .05$), significantly positive slopes for 2 of the 30 (6.7%) syntactic measures (Metric Node Count, Gate Counts), and marginally negative ($.05 < p < .1$) for 2 measures [Angular Total Depth; Metric Mean Straight Line Distance (R2)]. Regression equations for the remaining 4 measures (13.3%) did not reach significance. The syntactic measure that yielded the highest computed correlation ($R^2 = .54$; 0.2 and 0.3m bin sizes) was Metric Node Count (R1); the slope of the regression line was significantly positive ($p < .05$); equivalent correlation values ($R^2 = .54$) were obtained using both the *ln(W+1) vs. S>0* and *log10(W+1) vs. S>0* data transformations. Several measures exhibited an apparently nonlinear trend with peaks at intermediate bin sizes [e.g., Metric Mean Shortest Path Distance (R1), Figures 2 and 3].

In sum, (Q2) correlations declined as spatial resolution was increased: regression equations for a majority (22/30 or 73%) of syntactic measures revealed a significant ($p < .05$) negative relationship between syntactic-behavioral correlation strength and spatial resolution (bin size), and a significant positive relationship was found for only a small percentage (2/30 or 6.7%) the syntactic measures.

Part 2: Correlation (R^2) vs. spatial resolution for the best overall transformation [$\log_{10}$(W)>0]. In this analysis, the best-performing *$log_{10}(W)>0$ vs. S* data transformation was used. For each syntactic measure, a simple linear regression was calculated to predict correlation strength (R^2) based on spatial resolution (bin size). Results appear in Figure 3.

[Insert Figure 3]

Significantly negative regression equations were found ($p < .05$) for 8 of the 30 (26.7%) syntactic measures, and (Q2) correlations declined approximately linearly for these measures. The syntactic measure that yielded the highest computed correlation ($R^2 = .41$ at 0.6m bin size) was Metric Node Count (R1); this is the only measure that exhibited a positive regression equation ($p < .05$) with a positive relationship between correlation strength and spatial resolution. Marginally significant ($.05 < p < .1$) regression equations were found for 10 of the 30 (33.3%) syntactic measures, and (Q2) associated regression line slopes exhibited a negative trend, indicating that correlation strength declined approximately linearly for these measures. Regression equations for the remaining 12 of 30 (40%) measures did not reach significance. In sum, (Q2) correlations declined as spatial resolution was increased.

Q3: Does a small sample of spatial locations yield spuriously high correlations?

To examine our third research question, we used a simulation approach to examine the relationship between correlation strength and the number of randomly-located gate locations (N_{GATES}) at which pedestrian data are sampled. Specifically, we ran 100 replications for each of 100 randomly-located gates and 10 bin sizes (resolution). The key result we will examine here is the change in the correlation as the spatial resolution increased. Prior to running these simulations, syntactic data and walking data were computed for each bin size using the *$log_{10}(W)>0$ vs. S* transformation; this transformation yielded the highest percentage of maximal correlations in the foregoing analysis (see Table 1). Boxplots were used to summarize simulation results for the two syntactic measures examined in the previous section: Visual Integration (R3) (Figure 4) and Metric Node Count (R1) (Supplemental Material, Figure 7).

Visual Integration (R3). At all ten spatial resolutions examined, correlations between Visual Integration (R3) and walking data decreased (ΔR^2; M = -.22, SD = .019) as N_{GATES} increased. The first value of N_{GATES} (Figure 4, x-axis) at which a significant change (Ross, 2015) in local polynomial regression (LPR) fitted R^2 values (Figure 4, y-axis, blue best fit line) was detected was N_{GATES} = 23.

[Insert Figure 4]

This value was consistent across all ten of the spatial resolutions examined. Beyond 23 gates, correlations tended to stabilize (R^2_S) at a low but relatively constant value (*mean* R^2_S = .095; SD = .067). With respect to Q3, when fewer than 23 gates were used to compute correlations, perfect positive correlations (R^2 = 1) between Visual Integration (R3) and walking data were sometimes obtained; this result strongly suggests that using a small number of sampling grid locations can inflate correlations. ***Comparisons to random noise.*** To assess whether this measure correlated with walking data above chance levels, random noise was substituted for syntactic data, and correlated with walking data. Initial correlations (R^2_I) between syntactic data and walking data (*mean* R^2_I = .32, SD = .06) were 28% higher than correlations between random noise and walking data (*mean* R^2_I = .25, SD = .02), $t(10) = 3.58$, $p < .01$. Stabilized correlations (R^2_S; beyond N_{GATES} = 23) with walking data were also higher for syntactic data (M = .095, SD = .067) than random noise data (M = .022, SD = .01), $t(9) = 3.47$, $p < .01$. Thus, syntactic measures performed better than chance. However, with respect to Q3, perfect positive correlations ($R^2 = 1$) between random noise and walking data were sometimes obtained when fewer than 23 gates were used, strongly suggesting that using a small number of sampling grid locations can inflate correlations.

Discussion

In this section, we examine our results in relation to each of our three research questions, and compare our findings to the results of previous studies.

Q1: Correlations are sensitive to data transformations. We examined the relationship between (a) syntactic-behavioral correlation strength and (b) the underlying spatial resolution of both syntactic and pedestrian data. At each of the ten spatial resolutions examined, we found that correlations were highly sensitive to data transformations (see Table 1). The $log_{10}\ (W)>0$ *vs. S* data transformation yielded the highest correlation for a majority (25.5%) of measures examined, followed by the $ln(W)>0$ *vs. S* transformation (23.4%). The results strongly suggest that logarithmic ($\log_{10}$ or ln) transforming of positional data (to correct for departures from normality) and excluding zero values from the calculation yields the highest correlations between syntactic measures and continuous walking trajectories. Because data transformations are rarely reported in prior research, we strongly recommend that researchers check the distribution of their pedestrian data and clearly report whether the data were transformed.

Q2: Correlations decline as spatial resolution is increased. The results of the present study strongly suggest that correlations between syntactic measures and continuous walking trajectories are not robust to changes in scale. We found that when the most effective data transformation is used, the relationship between correlation strength and spatial resolution is significantly negative for a majority (24/30 or 80%) of syntactic measures ($p < .05$). Only one of the thirty (.03%) measures, Metric Node Count (R1), revealed a significantly positive relationship between correlation strength and spatial resolution ($p < .05$). Thus, for the vast majority of the syntactic measures examined, correlation strength declined with increasing spatial resolution.

Q3: A small sample of spatial locations can result in spuriously high correlations. When fewer than 23 locations (N_{GATES}) in the virtual environment used in the present study were used to sample syntactic and pedestrian data, the probability of obtaining spuriously high correlations increased dramatically. For each of the measures examined using the simulation approach, using fewer than 23 gates to compute correlations sometimes yielded outlying perfect positive correlations ($R^2 = 1$) between random noise and walking data. However, it is important to note that stabilized R^2 values remained higher than chance (i.e., performed better than correlations between random noise and walking data) when more than 23 gates were used, suggesting that space syntax measures can at least partially account for the variance in pedestrian movement patterns. Studies that use a small number of locations in space but show high correlations should be regarded as spurious because correlations between random noise and walking data exhibited the same behavior.

Relationship to previous research

Previous observational studies of people walking in real urban and architectural environments have found correlations between space syntax and walking data ranging from weak (e.g., .142; Mora, Astudillo, & Bravo, 2014) to strong (e.g., .98; Penn, Hillier, Banister, & Xu, 1998). Hillier (1999b) claims that "in most studies the best performing spatial variable is radius-3 integration." In comparison to previous studies, the present study used a controlled laboratory experiment and virtual reality to examine syntactic-behavioral correlations, and found that Visual

Integration (R3), reached a maximum correlation of R^2 = .30 (0.7m bin size, see both Figures 2 and 3). In the present study, the highest computed correlation with continuous walking data was found for Metric Node Count (R1) [R^2 = .54 at 0.6m, see Figure 2; R^2 = .41 at 0.6m, see Figure 3] using the $\log_{10}$ (W)>0 vs. S data transformation; other measures were generally inconsistent with this pattern of results. Results for Metric Node Count (R1) are all the more curious given that most previous research has generally focused on Visual Integration.

Desyllas and Duxbury (2001) compared the predictive capabilities of axial maps and visibility graph analysis (VGA) by sampling pedestrian flow data (N_{GATES} = 84) in a busy urban area (St. Giles Circus, London) at a rate of five minutes per hour on two non-consecutive days (Saturday and Tuesday). They computed syntactic-behavioral correlations for several axial map measures, and one local VGA measure (the natural log of Mean Visibility at 3m and 5m grid resolutions). Overall, they found that VGA significantly outperformed axial map analysis (best VGA correlation: R^2 = .625; best axial map correlation: R^2 = .429), and that the correlation between ln Mean Visibility and ln Mean Pedestrian Movement Data increased from R^2 = .456 to .625 as spatial resolution was increased from 5m to 3m. The present study differs from Desyllas and Duxbury's study in several important ways. First, while they examined several axial map measures and only one VGA measure, we systematically examined a wide variety of VGA measures. Second, they examined only two spatial resolutions (5m and 3m), and these were considerably lower than those used in the present study (1.0m – 0.1m). Our results revealed an opposite pattern of results for a majority (80%) of syntactic measures, and only one measure [Metric Node Count (R1)] exhibited an outlying pattern consistent with Desyllas & Duxbury's (2001) results. Because the present study more systematically examined a wider range of spatial resolutions, these results strongly suggest that correlations for most syntactic measures will be *strongest* when computed at *low* spatial resolution, and will *decline* at *high* spatial resolution. In addition, we evaluated a relatively large number of syntactic measures, providing baseline data for future examinations of whether syntactic measures might be grouped into classes that exhibit similar behavior across spatial scales.

Limitations of the present study

In order to provide a strong test of syntactic predictions in enclosed spaces, we chose to examine correlations between syntactic measures and continuous, naturalistic walking trajectories in hallways (a virtual hedge maze), where locomotion is relatively constrained. Future studies should examine correlations between syntactic measures and continuous walking trajectories in a variety of large open spaces (e.g., buildings consisting of variably-sized rooms linked by hallways and rings of circulation), using GIS methods (Lee and Seo, 2013; Liu et al., 2015), and leveraging tracking data from mobile phones and smart cities to examine the robustness of space syntax to changes in spatiotemporal resolution across a wider variety of navigational modes, contexts (both real and virtual), and spatial structures. Based on the present results, we expect that correlations will be even lower in open spaces because locomotion is relatively less constrained in open spaces than in restricted corridors. Although our goal-directed task is one that humans routinely perform in everyday life, it is unknown whether the cognitive demands of goal-directed navigation match those of the more undirected tasks from previous work. Thus, future work should also examine correlations when pedestrian data are drawn from a variety of tasks, in addition to the goal-directed exploration task used in the present study. In addition, because few comparisons of walking paths in real and virtual environments have been reported, future work should directly compare environments with matched configurations.

Past research has clearly demonstrated that space syntax measures *correlate* with human movement. Although strong correlations have led some researchers to conclude that spatial geometry *causes* particular behavior patterns, we recommend caution in drawing causal conclusions from correlational data. Future studies should more rigorously explore whether the features of environmental geometry that are encoded in space syntax measures play a demonstrably *causal* role in shaping human behavior and spatial knowledge. To this end, future studies should examine the robustness of syntactic predictions to parametric manipulations of environmental geometry. For example, we found that syntactic-behavioral correlations tended to stabilize when at least 23 gates were used to sample pedestrian data—a number that roughly corresponds to the number of maze segments between junctions and corners. This suggests a tentative hypothesis: that strong syntactic-behavioral correlations in enclosed spaces will be found when the number of gates corresponds to the number of path segments. Our use of randomly positioned gates underscores that spuriously high correlations can be obtained even when the gate positions are not placed in principled locations (e.g., corresponding to major path segments or junctions), which casts doubt on the hypothesis that strong syntactic-behavioral correlations in enclosed spaces will be found when the number of gates corresponds to the number of path segments.

Finally, past research suggests that a variety of contextual factors including the presence of salient objects (e.g., paintings, Tzortzi, 2009), landmarks (Appleyard, 1970; Montello and Pick, 1993), people (Appleyard, 1970; Dalton et al., 2011; Emo et al., 2012; Peponis et al., 1990), shops (Hillier et al., 1993), and traffic (Emo et al., 2012) can serve as attractors that impact pedestrian movement patterns. In contrast to prior observational studies—which have been constrained by the configuration of existing spaces—the use of controlled laboratory experiments in conjunction with virtual reality offers researchers the ability to systematically test the relationship between contextual factors, configurational properties, and human movement patterns.

Conclusions

The goal of the present study was to systematically examine how syntactic-behavioral correlations are impacted by data transformations, data resolution, and sampling strategies. The present study contributes to the space syntax literature by clarifying the impact of data transformations, comparing the performance of syntactic measures to a baseline of random noise, and more closely examining the robustness of correlations to changes in spatiotemporal scale. In sum, we found that syntactic-behavioral correlations (Q1) are sensitive to data transformations, (Q2) decline as the spatial resolution of VGA sampling grids is increased, and (Q3) can reach spurious levels when computed for only a subset of sampling locations in a visibility graph. We also found that correlations tend to stabilize when at least 23 sampling locations are used in the calculation, for our environmental configuration. Our results also provide useful baseline data for assessing the performance of syntactic measures across a wide variety of spatial and temporal scales. Finally, our findings strongly suggest that space syntax correlations are not robust to changes in spatial or temporal scale, and that high correlations obtained in previous space syntax studies may be spuriously high due to previously unexamined effects of data transformations, data resolution, or sampling strategies. Therefore, we recommend that researchers employing space syntax methods thoroughly report—and carefully consider—how each of these factors impact syntactic-behavioral correlations.

Acknowledgments

We are grateful to VENLab colleagues and staff, research assistants, and developers (Joost de Nijs, Kurt Spindler, Neal Fulwiler).

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by NSF (BCS-0214383; BCS-0843940), the NASA RI Space Grant, and Bentley University.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Appleyard D (1970) Styles and methods for structuring a city. *Environment and Behavior* 2(1): 110–117.

Bafna S (2003) Space syntax: A brief introduction to its logic and analytical techniques. *Environment and Behavior* 35(1): 17–29. DOI: 10.1177/0013916502238863.

Batty M (2001) Exploring isovist fields: Space and shape in architectural and urban morphology. *Environment and Planning B: Planning and Design* 28(1): 123–150. DOI: 10.1068/b2725.

Conroy R (2001) *Spatial Navigation in Immersive Virtual Environments (Unpublished doctoral dissertation)*. University College London, UK. DOI: 1111.

Dalton RC, Troffa R, Zacharias J, et al. (2011) Visual information in the built environment and its effect on wayfinding and explorative behavior. In: Bonaiuto M, Bonnes M, Nenci A, et al. (eds) *Urban Diversities—Environmental and Social Issues*. Hogrefe, pp. 6–76.

de Arruda Campos MB (1997) Strategic space: patterns of use in public squares of the city of London. In: *First International Symposium on Space Syntax*, Space Syntax Laboratory, Bartlett School of Graduate Studies, University College, London, 1997.

Desyllas J and Duxbury E (2001) Axial maps and visibility graph analysis. In: *Space Syntax Third International Symposium*, 2001, pp. 27.1-27.13.

Emo B, Hölscher C, Wiener JM, et al. (2012) Wayfinding and spatial configuration: evidence from street corners. In: *Eighth International Space Syntax Symposium*, Santiago de Chile, 2012, pp. 1–16. PUC.

Haq S and Luo Y (2012) Space Syntax in Healthcare Facilities Research: A Review. *Health Environments Research & Design Journal (HERD)* 5(4): 98–117. Available at: http://ezproxy.lib.ucalgary.ca:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=79867802&site=ehost-live.

Hillier B (1999a) The Common Language of Space: a Way of Looking at the Social Economic and Enviromental Functioning of Cities in a Common Basis. *Journal of Enviromental Sciences* 11(3): 344–249. DOI: 10.1177/1470357206065333.

Hillier B (1999b) The hidden geometry of deformed grids: Or, why space syntax works, when it looks as though it shouldn't. *Environment and Planning B: Urban Analytics and City Science*. DOI: 10.1068/b260169.

Hillier B and Hanson J (1984) *The Social Logic of Space*. Cambridge, UK: Cambridge University Press. DOI: 10.1016/0169-2046(86)90038-1.

Hillier B and Penn A (2004) Rejoinder to Carlo Ratti. *Environment and Planning B: Urban Analytics and City Science* 31(4): 501–511. DOI: 10.1068/b3019a.

Hillier B, Penn A, Hanson J, et al. (1993) Natural movement; or, configuration and attraction in urban space use. *Environment and Planning B: Planning and Design* 20: 29–66.

Hillier B, Major M, Desyllas J, et al. (1996) *Tate Gallery, Millbank: A study of the existing layout and new masterplan proposal*. London. Available at: http://eprints.ucl.ac.uk/932 (accessed 10 April 2013).

Jguirim I, Brosset D and Claramunt C (2014) Functional and Structural Analysis of an Urban Space Extended from Space Syntax. In: *GeoVisual Analytics: Interactivity, Dynamics, and Scale*, Vienna, AT, 2014. International Cartographic Association Commission on GeoVisualization.

Klarqvist B (1993) A space syntax glossary. *Nordisk Arkitekturforskning* 2: 11–12.

Kostakos V (2010) Space syntax and pervasive systems. In: Jiang B and Yao X (eds) *Geospatial Analysis and Modelling of Urban Structure and Dynamics*. Springer Netherlands, pp. 31–52. DOI: 10.1007/978-90-481-8572-6_3.

Lee S and Seo KW (2013) Combining Space Syntax With Gis-Based Built Environment Measures in Pedestrian Walking Activity. In: *the 9th International Space Syntax Symposium*, Seoul, 2013, pp. 098:1–14. Sejong University.

Liu J, Wu D, Hidetosi F, et al. (2015) Investigation and Analysis of Urban Spatial Structure around the Train Stations in Kitakyushu by Using Space Syntax and GIS. *Open Journal of Civil Engineering* 5(1): 97–108.

Mashhadi A, Acer UG, Boran A, et al. (2016) Exploring space syntax on entrepreneurial opportunities with Wi-Fi analytics. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '16*, Heidelberg, Germany, 2016, pp. 658–669. ACM. DOI: 10.1145/2971648.2971745.

Montello DR (2007) The contribution of space syntax to a comprehensive theory of environmental psychology. In: *6th International Space Syntax Symposium Instanbul 2007*, Istanbul, Turkey, 2007, pp. 1–12. Istanbul Technical University. Available at: http://www.spacesyntaxistanbul.itu.edu.tr/papers%5Cinvitedpapers%5Cdaniel_montello.pdf.

Montello DR and Pick HL (1993) Integrating Knowledge of Vertically Aligned Large-Scale Spaces. *Environment and Behavior* 25(3): 457–484. DOI: 10.1177/0013916593253002.

Mora R, Astudillo H and Bravo S (2014) Looking ahead: a vision-based software for predicting pedestrian movement. *Ingeniería e Investigación* 34(1): 79–82. DOI: 10.15446/ing.investig.v34n1.42792.

Netto MV (2016) `What is space syntax not?' Reflections on space syntax as sociospatial theory. *Urban Design International* 21(1): 25–40. DOI: 10.1057/udi.2015.21.

Okamoto K, Kaneda T, Ota A, et al. (2013) Correlation analyses between underground spatial configuration and pedestrian flows by space syntax measures: a case study of underground mall complex in Nagoya Station. In: *International Seminar on Urban Form (ISUF)* (eds P Sanders, M Guaralda, and L Carroli), Brisbane, Australia, 2013, pp. 116–128. Queensland University of Technology. Available at: http://eprints.qut.edu.au/80192/1/ISUF_Papers_Master_Document_Final.pdf#page=116.

Omer I and Goldblatt R (2017) Using space syntax and Q-analysis for investigating movement patterns in buildings: The case of shopping malls. *Environment and Planning B: Urban Analytics and City Science* 44(3): 504–530. DOI: 10.1177/0265813516647061.

Pafka E, Dovey K and Aschwanden GDPA (2018) Limits of space syntax for urban design: Axiality, scale and sinuosity. *Environment and Planning B: Urban Analytics and City Science* 0(0): 1–15. DOI: 10.1177/2399808318786512.

Penn A, Hillier B, Banister D, et al. (1998) Configurational modelling of urban movement networks. *Environment and Planning B: Planning and Design*. DOI: 10.1068/b250059.

Peponis J, Zimring C and Choi YK (1990) Finding the Building in Wayfinding. *Environment and Behavior* 22(5): 555–590. DOI: 10.1177/0013916590225001.

Peponis J, Wineman J, Rashid M, et al. (1997) On the description of shape and spatial configuration inside buildings: Convex partitions and their local properties. In: *Environment and Planning B: Urban Analytics and City Science*, 1997, pp. 761–781. DOI: 10.1068/b240761.

Ratti C (2004a) Rejoinder to Hillier and Penn. *Environment and Planning B: Urban Analytics and City Science* 31(4): 513–516. DOI: 10.1068/b3019b.

Ratti C (2004b) Space syntax: Some inconsistencies. *Environment and Planning B: Urban Analytics and City Science* 31(4): 487–499. DOI: 10.1068/b3019.

Ross G (2015) Parametric and nonparametric sequential change detection in R: The cpm package. *Journal of Statistical Software* 66(3): 1–20. DOI: http://dx.doi.org/10.18637/jss.v066.i03.

Silva J (2013) Towards statistical significance of configurational models: New evidence of variance and bootstrappin. In: *11th Space Syntax Symposium #170* (eds T Heitor, M Serra, JP Silva, et al.), Lisbon, Portugal, 2013, pp. 170.1-170.16. Instituto Superior Técnico, Departamento de Engenharia Civil, Arquitetura e Georrecursos. Available at: http://www.11ssslisbon.pt/docs/book-proceedings-05072017.pdf.

Turner A (2003) Analysing the visual dynamics of spatial morphology. *Environment and Planning B: Urban Analytics and City Science* 30(5): 657–676. DOI: 10.1068/b12962.

Turner A (2014) UCL Depthmap 7: Data analysis, Version 7.12.00c [Software manual]. Available at: http://archtech.gr/varoudis/depthmapX/LearningMaterial/depthmap7data.pdf%0D.

Turner A and Penn A (1999) Making isovists syntactic: isovist integration analysis. In: 1999. Citeseer.

Turner A and Penn A (2002) Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment. *Environment and Planning B* 29(4): 473–490. DOI: 10.1068/b12850.

Turner A, Doxa M, O'Sullivan D, et al. (2001) From isovists to visibility graphs: A methodology for the analysis of architectural space. *Environment and Planning B: Urban Analytics and City Science* 28(1): 103–121. DOI: 10.1068/b2684.

Tzortzi K (2009) The art museum as a city or a machine for showing art? In: *Proceedings of the 7th International Space Syntax* ... (eds D Koch, L Marcus, and J Steen), Stockholm, 2009, pp. 117:1-117:13. KTH Royal Institute of Technology. DOI: 10.1017/S1359135510000746.

Varoudis T (2015) depthmapX v0.5b [Computer software]. Available at: http://varoudis.github.io/depthmapX/%0D.

Table 1

Percentage of correlations (N = 410) for which each potential data transformation yielded the highest computed R^2 value

Data Transformation	*n*	%
W vs. S	28	6.8
W>0 vs. S>0	10	2.4
W vs. S>0	25	6.1
ln(W)>0 vs. S	97	23.7
ln(W+1) vs. S>0	35	8.5
ln(W+1) vs. S	30	7.3
ln(W)>0 vs. S>0	12	2.9
$\log_{10}$ (W)>0 vs. S	106	25.9
$\log_{10}$ (W+1) vs. S>0	30	7.3
$\log_{10}$ (W+1) vs. S	27	6.6
$\log_{10}$ (W)>0 vs. S>0	10	2.4

Note: W represents untransformed walking data, S represents untransformed syntactic measure, and parentheses indicate the order of operations applied to untransformed walking data prior to computing correlations (R^2). Overall, percentages in the % column sum to >100%: at some bin sizes, more than one data transformations produced equivalent maximal correlation values; this occurred for 110 cases, yielding a total of N = 410 (300 + 110) maximum correlations; in these cases, the counter (n) was incremented for more than one data transformation.

Figures

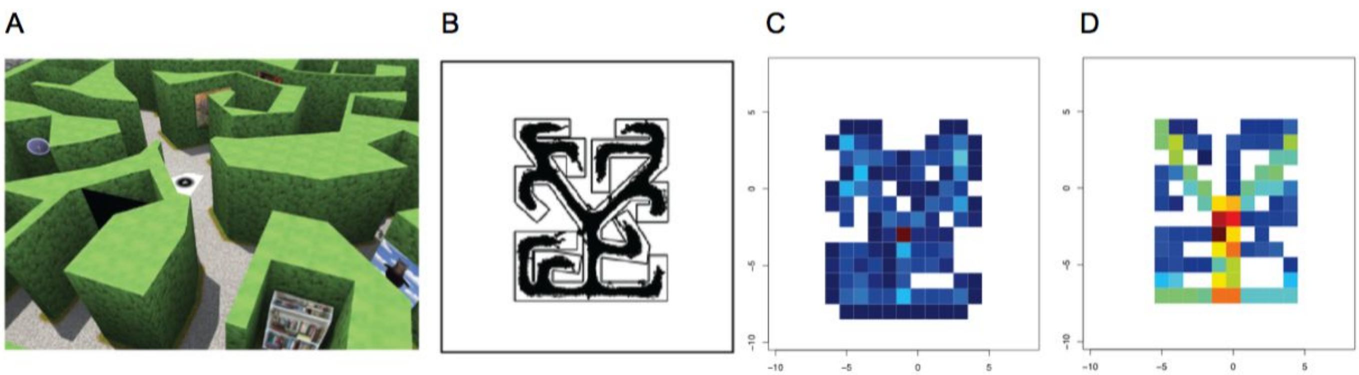

Figure 1. Displays and data collection. (A) Birds-eye-view of the virtual hedge maze. (B) Raw aggregated walking data from all study participants comprising 1.5 million data points collected at 1.5mm/0.10º spatial resolution and 60Hz temporal resolution. (C) Low-resolution binned walking data (bin size = 1.0m). (D) Example of low-resolution space syntax data (Connectivity; bin size = 1.0m) generated by depthmapX. With respect to panels C and D, note that the simulation parameter NGATES denotes the number of individual grid cells that have been randomly selected from among all of the available cells in the sampling grid to compute correlations; in the space syntax literature, this corresponds to the number of experimenters stationed to count pedestrian flows (i.e., obtain gate counts).

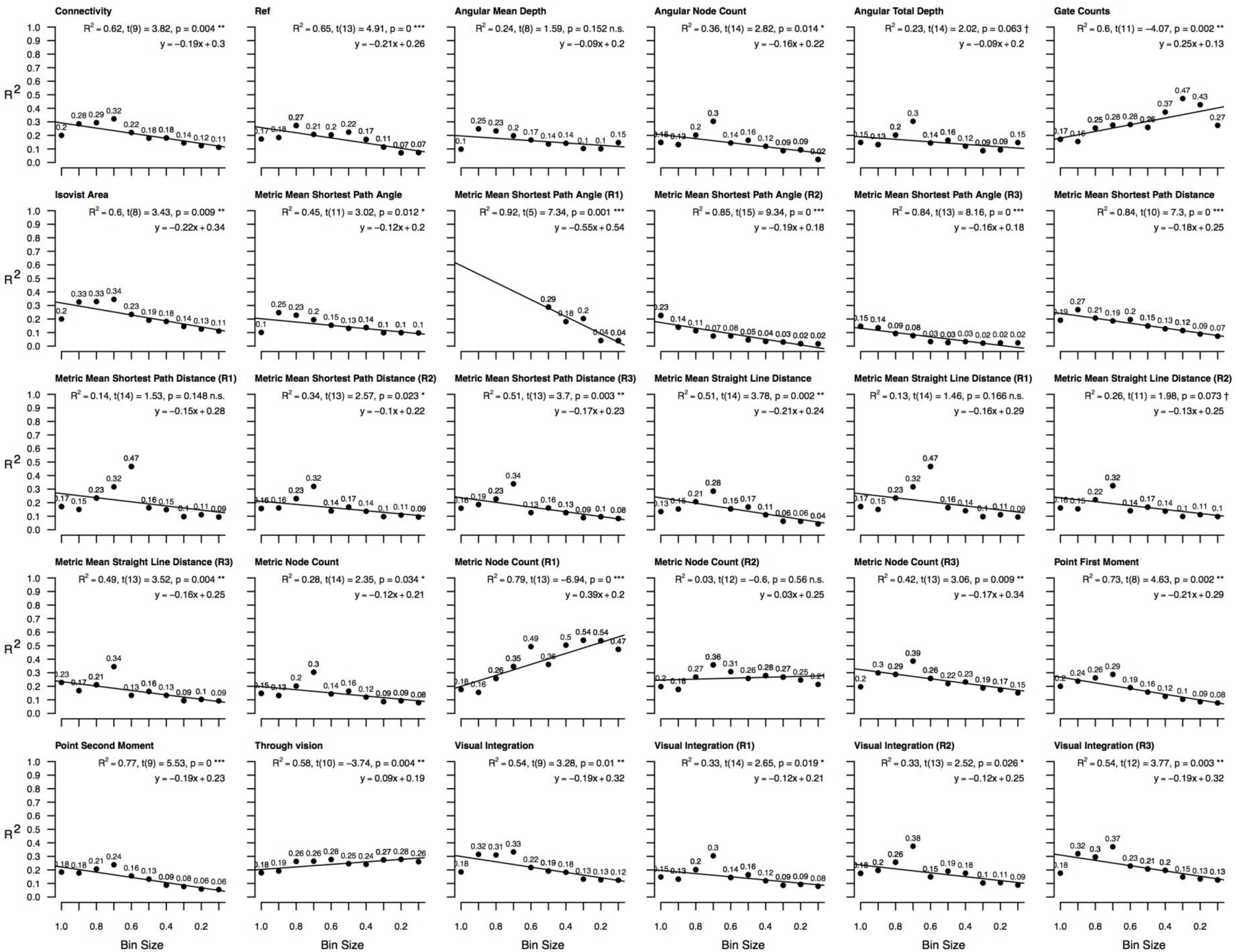

Figure 2. Correlation (R^2) vs. Spatial Resolution (Best Data Transformation at Each Bin size). In order to facilitate interpretation with respect to predictions concerning spatial resolution, bin sizes are plotted in reverse order (i.e., in order of increasing spatial resolution) from left to right. Significance tests and regression equations (y = mx + b) reported beneath facet titles (syntactic measures) indicate statistical results of linear regression. Significance levels are indicated as follows: * $p < .05$; ** $p < .01$; *** $p < .001$, † marginal ($0.5 < p < 0.1$), and results that failed to reach significance are also indicated (n.s.). Metric Mean Shortest Path Angle (R1) could not be computed for bin sizes ranging from 0.6m to 1.0m.

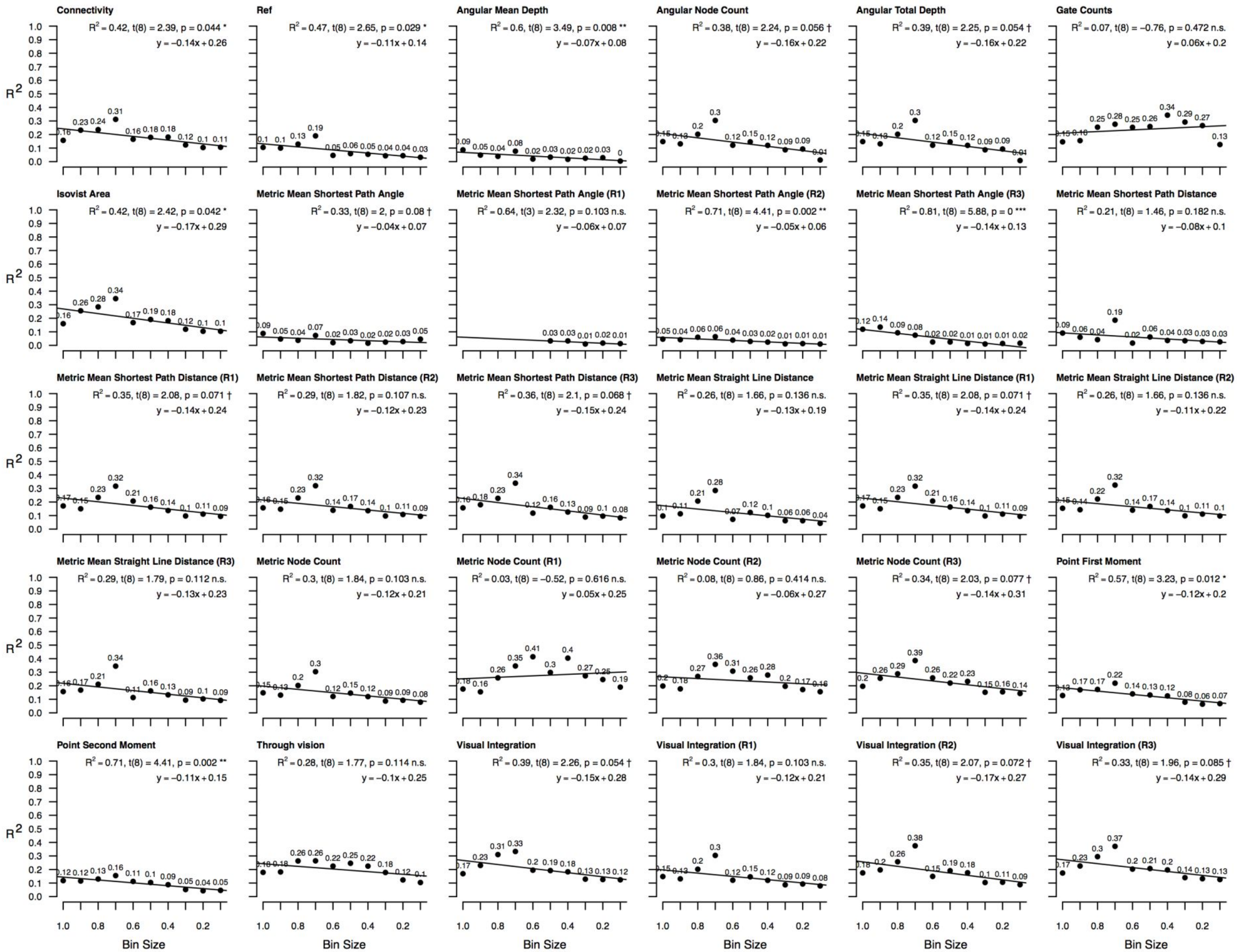

Figure 3. Correlation (R^2) vs. Spatial Resolution for Best Overall [log10 (W)>0 vs. S] Data Transformation. In order to facilitate interpretation with respect to predictions concerning spatial resolution, bin sizes are plotted in reverse order (i.e., in order of increasing spatial resolution) from left to right. Significance tests and equations (y = mx + b) reported beneath facet titles (syntactic measures) indicate statistical results of linear regression. Significance levels are indicated as follows: * p < .05; ** p < .01; *** p < .001, † marginal (0.5 < p < 0.1), and results that failed to reach significance are also indicated (n.s.). Note: The Metric Mean Shortest Path Angle (R1) calculation in depthmapX could not be computed for bin sizes ranging from 0.6m to 1.0m.

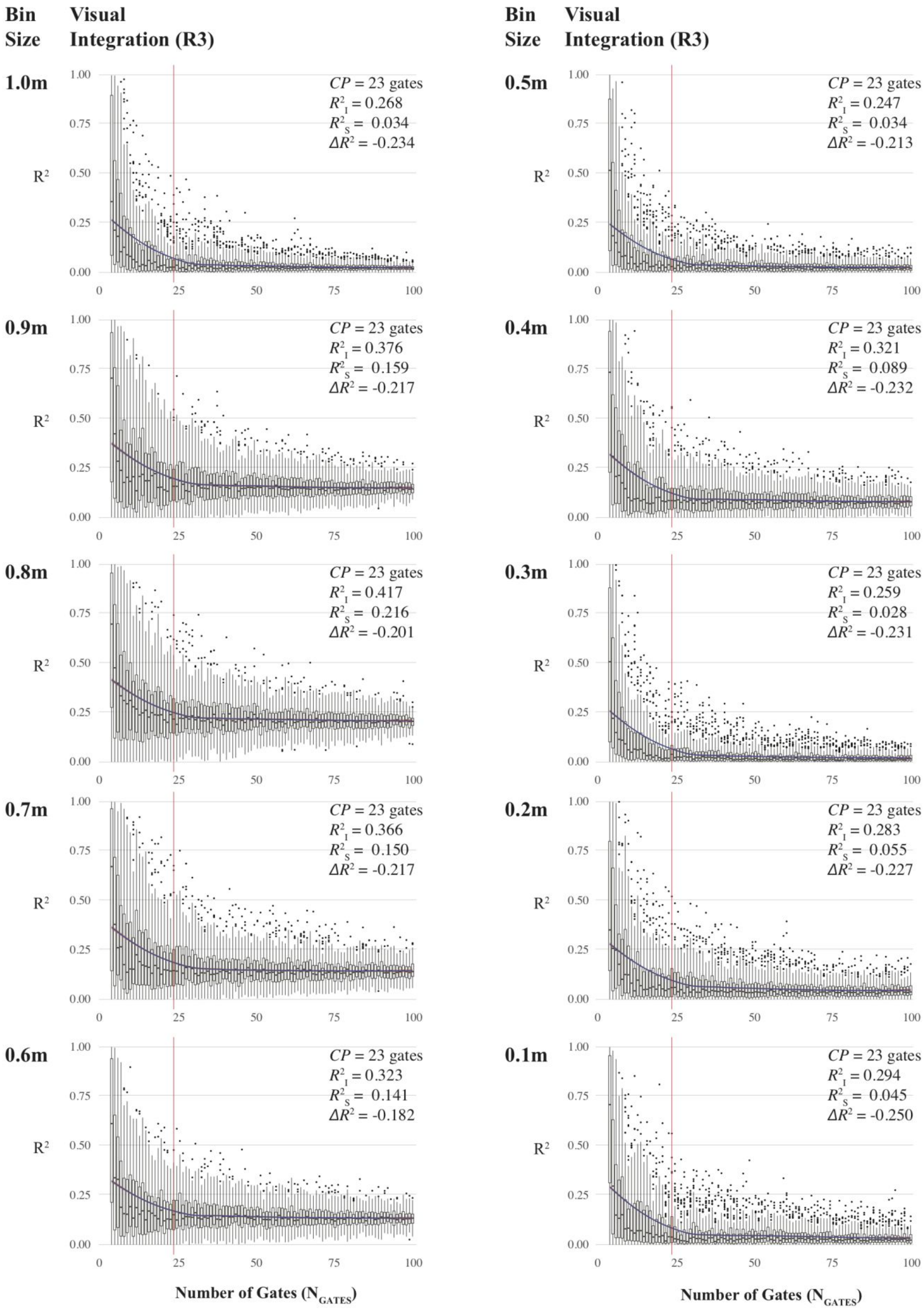

Figure 4. Simulations: Visual Integration (R3). *x-axes* (N_{GATES}): number of randomly sampled gate locations; 100 replications were simulated for each value of N_{GATES}. *y-axes*: R^2 values between the 100 simulations and the walking data. The mean and range are displayed for each value of N_{GATES}. *Boxplot whiskers*: min/max of 1.5x interquartile range. *Black dots*: outliers. *Trend lines*: best fit line for local polynomial regression (LPR) fit. *CP (red vertical line)*: first estimated change point value in LPR fitted R^2 values. R^2_I: initial R^2 value of LPR fit line (at $N_{GATES} = 3$). R^2_S: stabilized R^2 value, estimated by obtaining the mean of all LPR fitted R^2 from N_{GATES} = CP gates to $N_{GATES} = 100$. $\Delta R^2 = R^2_I - R^2_S$. For each value of N_{GATES}, 100 simulation runs were performed. *Boxplot hinges*: 25th and 75th percentiles.

SUPPLEMENTAL MATERIAL

Methods

Data Analysis

Data were analyzed using R (Version 3.3.1; R Core Team, 2016), R Studio (Version 1.0.136), Python (van Rossum and Drake, 2001), and depthmapX (Version 0.5b; Varoudis, 2015b). Points (x,y) defining the maze boundaries were imported into Adobe Illustrator's graph tool, converted to lines, and then exported to .dxf format. The maze and human data were imported into depthmapX (Figure 1B), and then binned at a variety of spatial resolutions.

Spatial binning. Because the walls of the maze corridors in the present study were 1.25m wide, employing bin sizes larger than 1.0m did not produce sensible results (e.g., a bin size of 1.1m produced cells that straddled both maze corridors and the inaccessible spaces between corridors). Due to the combinatorial explosion associated with computing space syntax measures at very high spatial resolution, it was impractical to compute space syntax measures for bin sizes smaller than .01 meters (3.94 inches).

Agent analysis. Default depthmapX settings for agent analysis were used, with the exception of a few parameters that were set to approximate those of the experimental design (for a detailed discussion of setting parameters for agent-based analysis using visibility graphs, see Turner 2003). First, the "Analysis length (timesteps)" parameter was set to 43,200 in order to approximate the parameters of the exploration phase (12 mins of exploration x 60 seconds per min x 60 Hz walking data sampling rate). Agents were released from a position at the center of the maze roughly corresponding to the location at which participants began the experiment. Finally, "Record trails for N" agents was set to N = 36, corresponding to the number of participants in the experiment. As Turner (2003) notes, these agent based analyses "approximate a Markov chain operating through locations on the visibility graph."

Data alignment and correlation method. Due to minor inconsistencies between depthmapX's coordinate system and the coordinate system employed in R (which ranged from -10 to 10), a custom R script was used to align the configurational bins and binned pedestrian count data. Matrices containing binned walking data (Figure 1C) and syntactic measures computed with the same spatial resolution (Figure 1D) were superimposed and then systematically shifted (up, down, left, and right) until an optimal overlap was found using a least-squares criterion (maximum Pearson's product moment R^2 value).

Data Transformation Analysis: Examining whether (Q1) correlations are sensitive to data transformations. Silva (2013) explicitly recommends log transforming pedestrian movement data to ensure that both movement and syntactic data follow normal distributions, enabling statistical comparisons between them. Exploratory data analysis suggested a variety of possible related data transformations beyond those recommended by Silva (2013), so we decided to systematically examine the impact of additional data transformations on correlation strength. First, raw walking data (W) values were correlated with raw syntactic (S) values (W vs S). Second, because it was possible to obtain syntactic or walking data values of zero, the analysis was restricted to values that were greater than 0 for both the walking data and space syntax data (*W>0 vs S>0*). Third, the natural logarithm of values produced using the previous method was also examined. Finally, because $\log_{10}(0)$ and ln(0) are undefined, and because the log+1 transformation is commonly used to correct for departures from normality, log+1 transformed walking data was also compared to raw syntactic values [e.g., *log(W+1) vs. S*, and *ln(W+1) vs. S*].

Regression Analysis: Examining whether (Q2) correlations depend on the spatial resolution of the sampling grid. We wished to be conservative in testing (Q2) whether correlations would decrease with increased spatial resolution. This required identifying data transformations that would be most charitable (i.e., that would allow space syntax measures the greatest chance to remain high as we increased the spatial resolution of the underlying sampling grid) toward a wide variety of measures and spatial resolutions (bin sizes). To accomplish this, two complementary approaches were taken. First (Part 1), we identified and opportunistically applied whichever data transformation (of the 11 transformations examined) produced the highest correlation for a given measure-bin pair. Second (Part 2), we identified a single (best overall) data transformation that produced maximal correlations for the largest percentage of measure-bin size pairs (see Table 2).

Simulations: Examining whether (Q3) a small sample of spatial locations would yield spuriously high correlations. Past research has generally sampled pedestrian data at subsets of locations ("gates" or grid cells) within the overall VGA sampling grid, rather than sampling pedestrian data at all possible sampling grid locations. Historically, the number of "gate" locations (N_{GATES}) has been limited due to data collection constraints (e.g., needing large numbers of researchers to collect data, or relative ease of counting pedestrians passing through doorways), and because gates are often positioned at locations convenient for researchers, it is possible that high correlations obtained in previous studies may be due to selection bias. In contrast, the motion tracking system used in the present study recorded *all* possible locations within the sampling grid, providing a more comprehensive assessment of syntactic predictions. In addition, we examined how correlations vary as increasingly large *subsets* of grid cells are randomly selected, simulating how stationing an increasing number of randomly located “experimenters” (N_{GATES}) to record gate counts impacts correlations. We evaluated whether using small subsets (n = 3) of grid cells to compute R^2 values would yield spuriously high correlations, and whether sampling from an increasing number of locations (up to n = 100) would yield more reliable or “stable” correlations (denoted R^2_S).

Several approaches were used to quantify how syntactic-behavioral correlations (R^2 values) vary as the number of “gates” (N_{GATES}) is increased. First, 100 replications (R^2 values) were computed for each simulated value of N_{GATES} (this corresponds to randomly distributing 100 distinct sets of N “experimenters” to count pedestrian flows for each N_{GATES} value, where N = N_{GATES}), yielding a total of 100,000 simulated R^2 values [(100 gates) * (100 replications/gate) * (10 bin sizes)] for each syntactic measure. Exploratory data analysis suggested that (a) R^2 values were highest when the number of sampling grid locations (N_{GATES}) was relatively low, and that (b) the mean R^2 value appeared to decline exponentially, before stabilizing above a critical value of N_{GATES}. Therefore, a change point approach was used quantify the presence of inflections or “change points” (CP, the gate count at which correlations tended to stabilize) in simulation data; thus, CP is the critical measure used to assess the minimum number of sampling grid locations (N_{GATES}) required to obtain reliable estimates of correlation strength at a given grid resolution.

Change points were detected by first computing local polynomial regression (LPR) fits for simulated R^2 values as a function of N_{GATES} (using the “loess” function from R’s “stats” package), and then obtaining the first detected change point in these regression fits (using R’s “cpm” package) (Ross, 2015). The initial LPR fitted the R^2 value at the minimum number of gates examined (N_{GATES} = 3), indicated by R^2_I. The arithmetic mean of the LPR fitted R^2 values between N_{GATES} = CP and N_{GATES} = 100 was used to estimate the point at which R^2 stabilized at a relatively constant value (R^2_S). Finally, the difference between the two R^2 values ($\Delta R^2 = R^2_I - R^2_S$) was computed to examine how correlations vary as the number of sampling grid locations (N_{GATES}) was increased.

Results

Q2: Do correlations depend on the spatial resolution of the sampling grid?

Heatmaps for Leading Syntactic Measures

This section examines whether (Q2) correlations depend on the spatial resolution of the sampling grid by discussing data for a syntactic measure that generally performs well in the space syntax literature [Visual Integration (R3)], and for the syntactic measure that yielded the highest correlation obtained in the present study [Metric Node Count (R1)]. Heatmaps in Figures 5 and 6 plot syntactic values and binned walking data at corresponding grid resolutions for Visual Integration (R3) and Metric Node Count (R1) respectively. Heatmap values and correlations were computed at each bin size after applying a $Log_{10}(W)>0$ *vs. S* data transformation; color scales indicate data ranges and color mappings.

Heatmaps and Correlations for Visual Integration (R3). As previously noted, Integration is a commonly reported measure in the space syntax literature. Visualizations of Visual Integration (R3) and binned walking data for all 10 bin sizes appear in Figure 5.

[Insert Figure 5]

For this measure, R^2 values increased from R^2 = .17 at the lowest spatial resolution (1.0m) to a maximum of R^2 = .37 at an intermediate resolution (0.7m), and decreased to R^2 = .13 at the highest spatial resolution (0.1m) examined. Thus, we found that (Q2) correlation strength decreased (by .05) as spatial resolution was increased. It is worth nothing that R^2 values reached a peak at $R^2 = .37$, which may provide support for the notion that there is an ideal resolution for syntactic measures (Al Sayed et al., 2014; Turner et al., 2001).

Heatmaps and Correlations for Metric Node Count (R1). Visualizations of Metric Node Count (R1) and binned walking data for all 10 bin sizes appear in Figure 6.

[Insert Figure 6]

Metric Node Count (R1) yielded the highest correlation value found in the present study. For this measure, R^2 values increased from $R^2 = .18$ at the lowest spatial resolution (1.0m) to a maximum of $R^2 = .41$ at an intermediate resolution (0.6m), and decreased to $R^2 = .19$ at the highest spatial resolution (0.1m) examined. With respect to Q2, this syntactic measure exhibited a more complex pattern of results than we predicted, with correlations peaking at intermediate bin sizes (see Figures 2 and 3), which may be consistent with the claim that there is an ideal spatial scale for computing syntactic-behavioral correlations (Al Sayed et al., 2014; Turner et al., 2001).

Q3: Does a small sample of spatial locations yield spuriously high correlations?

Each boxplot (Figures 4 and 7) shows simulated R^2 values (y-axis) against the number of randomly sampled gate locations (x-axis N_{GATES}) for a given bin size. Boxes and whiskers summarize the distribution of the results from all 100 replications at each value of N_{GATES}; whiskers extend to the minimum and maximum simulated R^2 values, and extend no further than 1.5 times the interquartile range (IQR); box hinges indicate the 25th and 75th percentiles of the simulated R^2 values; outlying points are indicated as black dots.

Metric Node Count (R1). At all ten spatial resolutions examined, correlations between Metric Node Count (R1) and walking data decreased (ΔR^2; $M = -.18$, $SD = .042$) as N_{GATES} increased. The first value of N_{GATES} (Figure 7, x-axis) at which a significant change (Ross, 2015) in LPR fitted R^2 values (Figure 7, y-axis, blue best fit line) was detected was $N_{GATES} = 23$, just as we found for Visual Integration (R3).

[Insert Figure 7]

This value was consistent across all ten of the spatial resolutions examined. Beyond 23 gates, correlations tended to stabilize (R^2_S) at a low but relatively constant value (*mean* $R^2_S = .258$; $SD = .151$). With respect to Q3, when fewer than 23 gates were used to compute correlations, perfect positive correlations ($R^2 = 1$) between random noise and walking data were obtained, strongly suggesting that using a small number of sampling grid locations can inflate correlations.

Comparisons to random noise. To assess whether this measure correlated with walking data above chance levels, random noise was substituted for syntactic data, and correlated with walking data. Initial correlations (R^2_I) between syntactic data and walking data (*mean* $R^2_I = .43$, $SD = .13$) were 72% higher than correlations between random noise and walking data (*mean* $R^2_I = .25$, $SD = .01$), $t(9) = 4.37$, $p < .01$. Stabilized correlations (R^2_S; beyond $N_{GATES} = 23$) with walking data were also higher for syntactic data ($M = .257$, $SD = .15$) than random noise data ($M = .02$, $SD = .001$), $t(9) = 4.98$, $p < .001$. Thus, syntactic measures performed better than chance. However, with respect to Q3, when fewer than 23 gates were used to compute correlations, perfect positive correlations ($R^2 = 1$) between random noise and walking data were obtained, strongly suggesting that using a small number of sampling grid locations can inflate correlations.

Discussion

Local maxima. Some measures exhibited a small "hump" or local maximum in correlation strength (R^2) at an intermediate spatial resolution near 0.7m (see Figures 2 and 3). In an effort to adopt a scale of analysis commensurate with typical human walking behavior, Turner et al. (2001) employed a 1m grid spacing, and Al Sayed et al. (2014; *depthmapX* handbook) recommends that researchers depthmapX users select “a sensible grid spacing values that match the human scale (0.6 - 0.7 meters).” Thus, our results could be interpreted as supporting the claim that there is an optimal spatial scale for correlating space syntax measures with pedestrian behavior; the large number of syntactic measures examined in the present seems to support the recommendations made by other researchers. However, we urge caution with respect to this interpretation of our results. While several previous studies (Emo et al., 2012; Ferguson et al., 2012; Turner, 2003) cite Gibson’s (1950, 1986) ecological approach to visual perception as the theoretical basis for positing a causal relationship between syntactic variables and pedestrian behavior, they do not clearly articulate why syntactic-behavioral correlations should be maximal at human scale. Moreover, operational definitions of “human scale” have been extensively debated, and remain controversial (see Ewing & Handy, 2009 for a review).

Tables

[Insert Table 2]

Table 2
Summary of syntactic-behavioral correlations found in selected previous studies

Study	Mode	Environment	Syntactic Measure(s)	Data Transformations	Correlations
Hillier et al. (1996)	Walking	Museum	Integration	ln (movement rates)	$.37 < R^2 < .86$
de Arruda Campos (1997)	Walking	Urban area	Integration (R3) Integration (RN)	(Unknown)	$.81 < R^2 < .88$ $.80 < R^2 < .80$
Penn, Hillier, Banister, and Xu (1998)	Vehicle	Urban area	Integration (R3, R5, R7, R9)	Fourth root of flow rates	$.34 < R^2 < .83$
	Walking	Urban area	Mean integration (R3) and development density	Net capacity	$R^2 = .98$
Turner & Penn (1999)	Walking	Museum	Isovist Integration	Log of mean occupancy levels	$R^2 = .585$
		Store	Isovist Area		$.324 < R^2 = .578$
Desyllas & Duxbury (2001)	Walking (5m and 3m)	Urban area	Axial Map Analysis	ln (mean visibility) and ln (mean pedestrian movement data)	$R^2 = .456$ (5m) $R^2 = .625$ (3m)
Turner (2003)	Walking (3m)	Urban area	Various	Log transformed agent simulation data	$.29 < R^2 < .73$
Turner (2007)	Walking	Museum	Through vision (agent simulation)	ln (movement rates)	$.68 < R^2 <.74$
Mora, Astudillo, and Bravo (2014)	Walking	Urban area	Gate counts ($N_{GATES} = 203$)	Mean gate counts over six consecutive workdays	$.142 < R^2 < .271$
Okamoto et al. (2013)	Walking	Commuter rail mall	Gate counts ($N_{GATES} = 50$)	Connectivity, visual step depth, shortest distance, integration	$.2 < R^2 < .598$

Note: Mode column indicates whether pedestrian (walking) data or vehicular data were correlated with syntactic measure(s). Correlations column includes minimum and maximum syntactic-behavioral correlations found in the study. N_{GATES} = the reported number of gate locations at which pedestrian flows were counted.

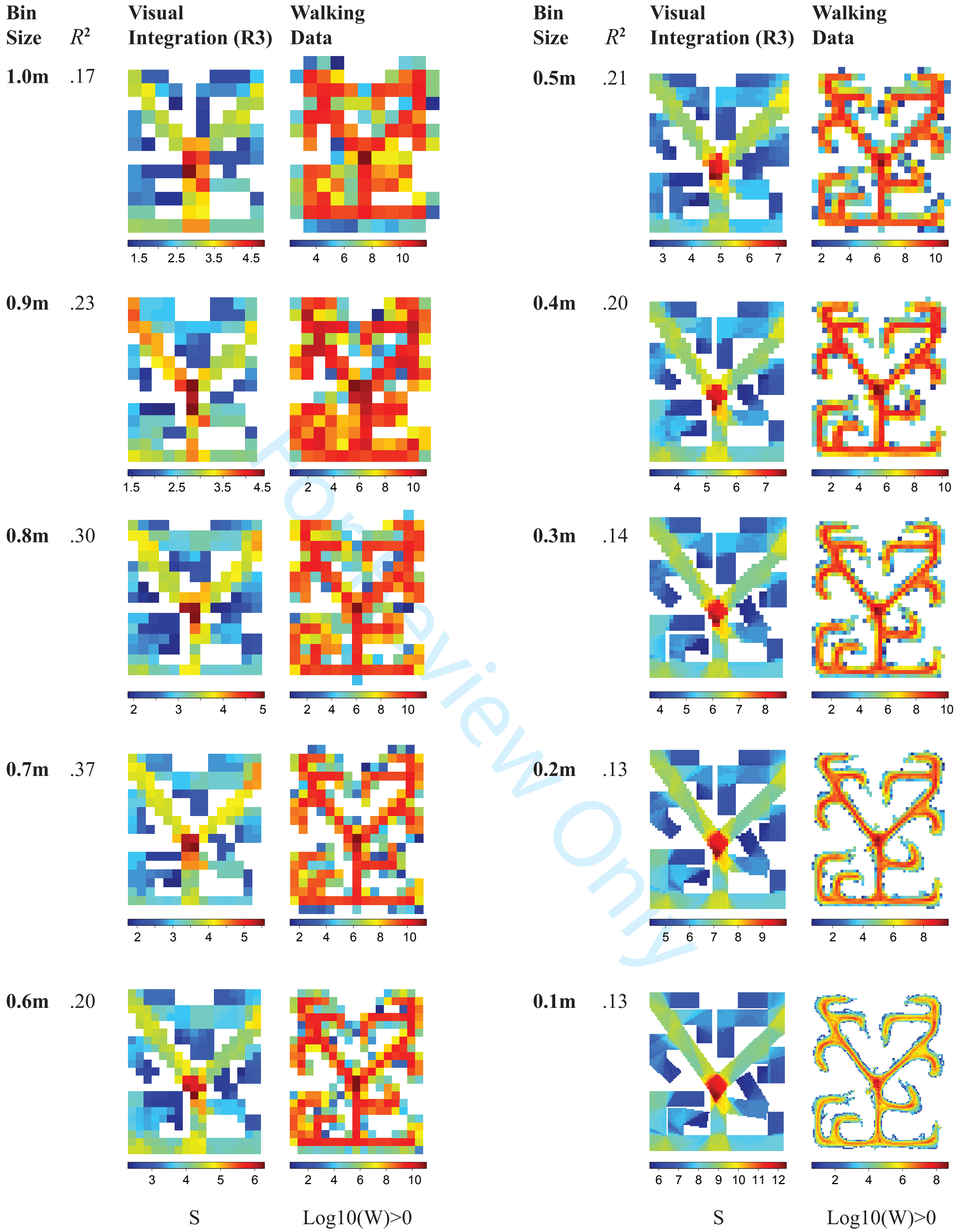

Figure 5. Heatmaps and Correlations for Visual Integration (R3). Heatmap values and correlations were computed at each bin size after applying a Log10(W)>0 vs. S data transformation. For Visual Integration (R3), colors indicate raw syntactic measure values (S). As previously noted, the Integration value for a grid cell is obtained by computing the average depth (i.e., topological distance) of that cell to neighboring cells within a specified topological distance (radius), effectively ranking cells "from the most integrated to the most segregated" (Klarqvist, 1993). For the walking data, colors indicate Log10(W)>0 transformed position data count values computed for each grid cell; the same walking data are plotted in both Figures 5 and 6.

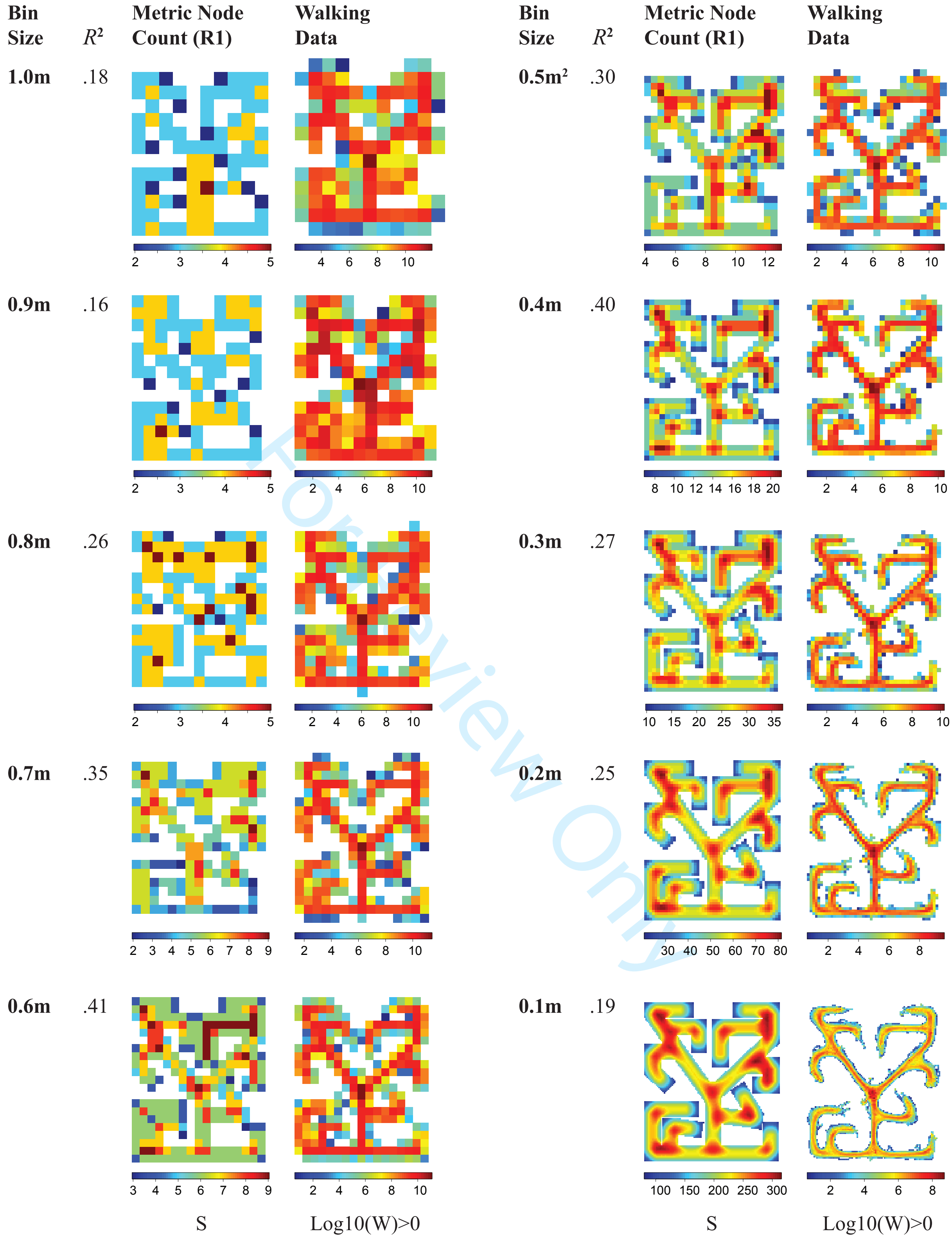

Figure 6. Heatmaps and Correlations for Metric Node Count (R1). Heatmap values and correlations were computed at each bin size after applying a Log10(W)>0 vs. S data transformation. For Metric Node Count (R1), colors indicate raw syntactic measure values (S). Metric Node Count (R1) is the number of neighboring nodes within a specified topological distance (radius) (Turner, 2004). For the walking data, colors indicate Log10(W)>0 transformed position data count values computed for each grid cell; the same walking data are plotted in both Figures 5 and 6.

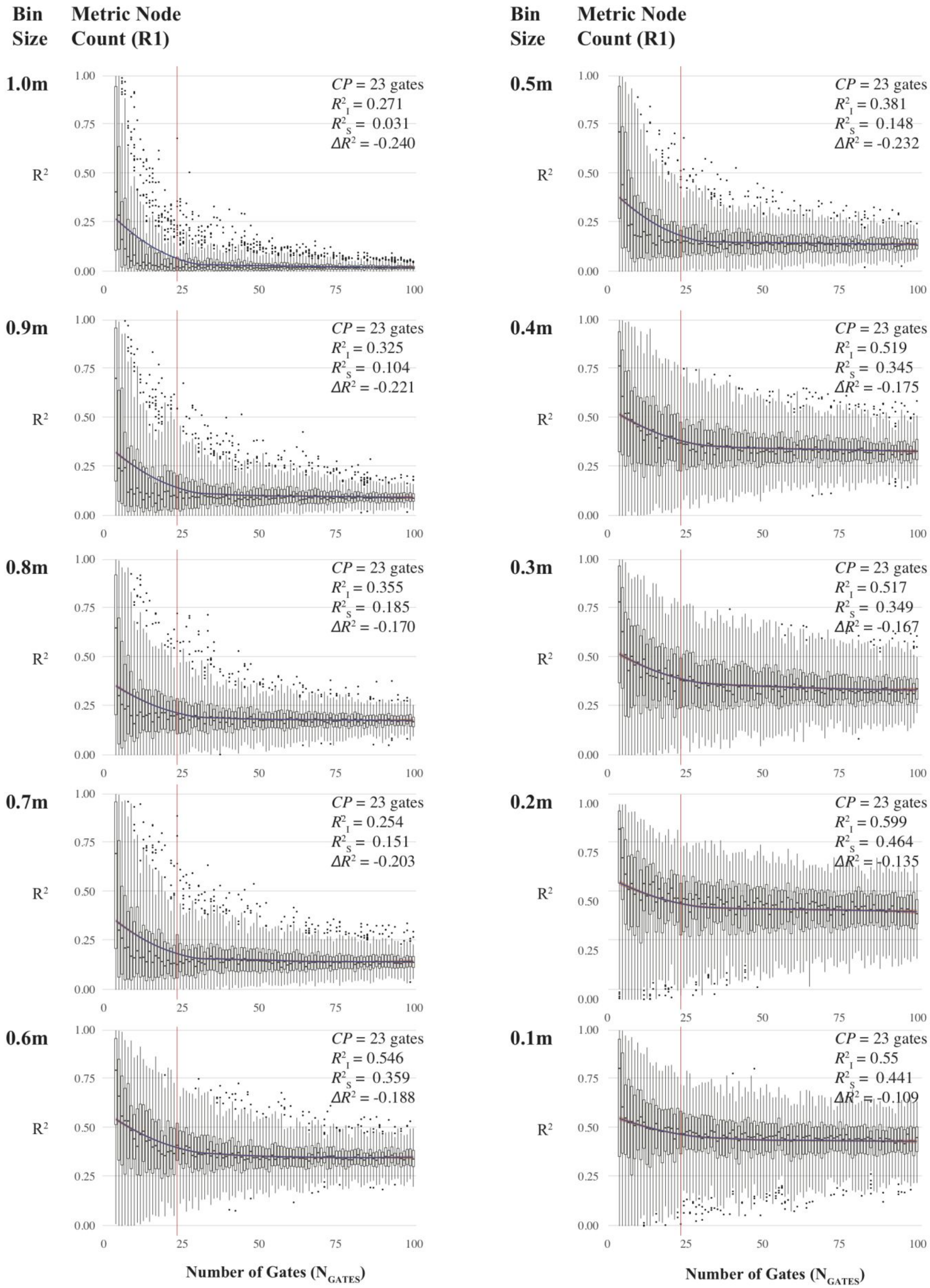

Figure 7. Simulations: Metric Node Count (R1). *x-axes* (N_{GATES}): number of randomly sampled gate locations; 100 replications were simulated for each value of N_{GATES}. *y-axes*: R^2 values between the 100 simulations and the walking data. The mean and range are displayed for each value of N_{GATES}. *Boxplot whiskers*: min/max of 1.5x interquartile range. *Black dots*: outliers. *Trend lines*: best fit line for local polynomial regression (LPR) fit. *CP (red vertical line)*: first estimated change point value in LPR fitted R^2 values. R^2_I: initial R^2 value of LPR fit line (at $N_{GATES} = 3$). R^2_S: stabilized R^2 value, estimated by obtaining the mean of all LPR fitted R^2 from N_{GATES} = CP gates to $N_{GATES} = 100$. $\Delta R^2 = R^2_I - R^2_S$. For each value of N_{GATES}, 100 simulation runs were performed. *Boxplot hinges*: 25th and 75th percentiles.