The Comparative Gamma Map Method:

A Topo-Configurational Sketch Map Coding and Analysis Method for Survey View Building Sketch Maps

by

Timothy Bruce

September 2022

Supervisor: Professor Alan Penn

A Dissertation submitted in part fulfilment of the

Degree of Master of Science (MSc)

Space Syntax: Architecture and Cities

Bartlett School of Architecture

University College London

UCL FACULTY OF THE BUILT ENVIRONMENT
BARTLETT SCHOOL OF ARCHITECTURE

MSc SPACE SYNTAX: ARCHITECTURE AND CITIES
Coursework Submission Form

1.1.1. Student's Name (IN CAPS): TIMOTHY LAWRENCE BRUCE Word Count: 10'154

1.1.2. Dissertation Title: The Comparative Gamma Map Method: A Topo-Configurational Sketch Map Coding and Analysis Method for Survey View Building Sketch Maps

---

1.2. Deadline for Final Submission: **12.09.2022**

1.3. Actual Date of Submission (if different):

(If different please state whether this has been agreed with the module leader and the appropriate documentation provided)

1.4. Student Statement

- **I confirm I understand the meaning of plagiarism and that I may be penalised for submitting work that has been plagiarised.**
- **I confirm the work submitted is my own unaided work and that all sources used in its preparation and all quotations are clearly cited.**
- **I confirm that all work will also be submitted electronically and that this can be checked using the JISC detection service, Turnitin.**
- **I confirm that I understand that the work cannot be assessed unless I have handed in both hard copy and electronic versions.**

1.4.1. Student Signature: T. B

Your work will be returned unmarked if you fail to sign here

Abstract

Sketch mapping is a popular technique in spatial cognition research to collect information about an experimental participant's environmental knowledge. Reliably processing sketch maps into quantitative data for spatial cognition research is a widely recognised problem (Montello 2016). Addressing this, Bruce (2022) introduced the *comparative gamma map* method, a topo-configurational coding and analysis approach for survey view building sketch maps. So far, the method has been successfully applied to a sample of survey view sketch maps of simple buildings. However, survey view sketch maps of complex buildings undergo noticeable shifts in their graphical quality distribution and occurrence of labelled spaces which threaten methodological reproducibility and necessitate methodological adaptations. This study matures the comparative gamma map method into a building-scale-robust approach through three focus areas: 1) Development of new accuracy metrics and building graphs for complex buildings. 2) Methodological reproducibility benchmarking. 3) Experimentation with unsupervised classification algorithms. From this, three main outputs follow: 1) Alongside six novel sketch map accuracy metrics, the *hybrid-axial graph* is introduced as a graph that models the unique structure of complex buildings while meeting requirements of spatial cognition research. 2) A reproducibility benchmarking process is established in which initial findings suggest the comparative gamma map method behaves more consistently across building scales than conventional approaches. 3) Sketch maps are found to be classifiable into similar structural groups of building "ringiness" and size by clustering them by their sketch map measurements and metrics derived from the comparative gamma map method. Beyond addressing the problem of processing sketch maps, the comparative gamma map method presents an exciting prospect to overcome the longstanding divide between theory and coding approach in spatial cognition research. This promises to open the field to new research in which theory is testable at the level of the theoretical model.

Keywords: *sketch map coding, gamma map, comparative gamma map method, spatial cognition, space syntax*

Word Count: 10'154

Acknowledgements

I would like to express my gratitude towards my dissertation supervisor Professor Alan Penn. His guidance, encouragement, and expertise were indispensable to this dissertation. I would also like to thank Professor Kate Jeffery at the Department of Behavioural Psychology, UCL for providing me with sketch map data from her own cognitive mapping research and further guidance.

This dissertation is dedicated to my mum, Regula, my dad, Richard, my brother, Jonathan and all my friends who have unconditionally supported me throughout the past year. I would like to specifically thank the lucky individuals (Florine, Elisa, Pete, Konstantin and Luca) who braved listening to my long-winded explanations and thoughts on this dissertation – trust me, we can talk about other things again now.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Context

Humans and animals alike rely on mental models to navigate the world (Tolman 1948). These models are referred to as cognitive maps in spatial cognition research and the process through which they are formed from experienced and sensed external environments is referred to as cognitive mapping (Downs and Stea 1973). The construction, structuration, and use of cognitive maps by the mind have been studied through the collection and analysis of sketch maps as a technique to elucidate information about a subject's environmental knowledge (Golledge and Stimson 1997). However, working with sketch maps presents challenges which threaten to undermine the reproducibility of spatial cognition research. The crux of these challenges lies in converting sketch maps into quantitative information, as most maps are distorted, incomplete and devoid of an underlying metric system (Golledge and Stimson 1997). Qualitative approaches to quantifying sketch map accuracy, such as scoring their accuracy (in comparison to the reference environment) on a Likert scale, ranking from best to worst, or categorising by visually perceptible qualities, largely rely on subjective opinions and are confounded by the graphical abilities of the person producing the map. Furthermore, topology and configuration, which is the most reliable information underlying sketch maps, is disregarded (Golledge and Stimson 1997). Accordingly, there is a need for novel quantitative measures of sketch map accuracy.

In a previous research paper (Bruce 2022), the author introduced a novel method to quantify sketch map accuracy for survey-view sketch maps of small-scale buildings. This work is attributable to a general push within the field of spatial cognition research to incorporate analytical descriptions of space through the deployment of the space syntax toolkit (Dalton, Hölscher and Turner 2012). Inspired by Hillier and Hanson's (1984) gamma map, labelled topological graphs are abstracted from sketch and reference maps, called sketch and baseline graphs, and compared to each other through a set of accuracy metrics. The metrics are based on graph entity counts and graph entity matches. Through this approach a robust and reproducible assessment framework was developed, enabling more specific and novel cognitive mapping insights. This was evidenced through a proof-of-concept study on a pre-existing dataset of small-scale survey-view building sketch maps.

To date this method has only been tested on small-scale building systems. From these, labelled topological graphs can be abstracted and compared with minor interpretative biases. Survey-view sketch maps of more complex buildings undergo two noticeable shifts which necessitate further research and methodological developments. Firstly, they become increasingly unlabelled, as spaces within the building system either lose their spatial distinctiveness or the size of the system outgrows the capacity of the human mind to reference all spaces in the sketch map. Secondly, the sketch maps distort more and are increasingly incomplete as the differing graphical skills of the subjects are emphasised. The latter is particularly concerning, as it suggests that quasi qualitative approaches, such as scoring and ranking, may be increasingly biased by the aesthetic appearances of the sketch maps. This issue is not directly remediable through the application of the original comparative gamma map approach suggested by Bruce (2022), as it is bound to be more confounded by the increased ambiguity in abstracting sketch graphs. Furthermore, the original accuracy metrics can no longer rely on the labelled properties of the sketch graph, which calls for methods transcending entity counts and matches to account for the comparative "deep-structures" of the sketch and baseline graphs.

1.2. Aims

This dissertation seeks to mature the original comparative gamma map method into a building-scale-robust sketch map coding and analysis approach. Three interrelated research areas are identified and pursued through research questions.

The first area is concerned with developing novel graph comparison metrics for the partially labelled and unlabelled graphs of complex buildings. Accordingly, there is one research question:

- **RQ1:** How can the original comparative gamma map method proposed by Bruce (2022) be adapted to complex buildings with unlabelled or partially labelled graphs?

The second area is concerned with how to model complex buildings best as graphs. Importantly, an understanding of the reliability of the comparative gamma-map method for simple and complex buildings, in comparison to qualitatively based scoring and ranking approaches, should be developed. Accordingly, there are three research questions:

- **RQ2:** How can complex buildings best be represented as graphs?
- **RQ3:** How reproducible are qualitative sketch map scoring and ranking results compared to comparative gamma map results?
- **RQ4:** To what extent are qualitative scoring and ranking approaches and the comparative gamma-map method influenced by the perceptible shifts from simple to complex building sketch maps?

The last area is concerned with showcasing the potential of harnessing sketch graphs to arrive at novel insights through unsupervised learning algorithms. The aim is to explore possibilities to classify sketch maps according to structural properties that transcend immediate visual perception and classification. Accordingly, there is one last research question:

- **RQ5:** What kind of insights can be arrived at by applying an unsupervised classification algorithm to a sketch graph dataset?

1.3. Outline

Chapter 2 reviews the literature on analysing cognitive mapping data, with a specific focus on the issue of sketch map coding. In **Chapter 3** the adapted comparative gamma map method is presented, as well as how to best abstract a graph from complex buildings. This is followed by a reproducibility analysis of ranking approaches, scoring approaches and the comparative gamma map approach in **Chapter 4**. In **Chapter 5** the application of an unsupervised classification algorithm on sketch graph data is explored. This is followed by a discussion on embodying spatial cognition theory through sketch map coding in **Chapter 6** and a conclusion in **Chapter 7**.

2. Literature Review

2.1. Spatial Cognition, Cognitive Maps and Cognitive Mapping

Montello and Raubal (2012: 162) define spatial cognition as "the multi-disciplinary study of perception, thinking, reasoning, and communications that is fundamentally about spatial properties and relations […] in the environment, whether by humans, non-human animals, or computational entities such as robots." How humans and non-human animals alike structure spatial knowledge is a leading debate in spatial cognition research. The cognitive map theory is most widely supported and originates from Tolman's (1948) research on rats. He first conceived of the "cognitive map" as a map-like mental representation deployed by rats – and by extension humans – to purposefully navigate their surroundings. Spatial knowledge as structured in a metric Euclidian cognitive map has since been supported by breakthroughs in neuroscience, such as the discovery of the place cell (O' Keefe and Nadel 1978, Gallistel 1990) and grid cell (Bicanski and Burgess 2020). Whilst the notion of spatial information as mentally represented through a model-like entity is widely accepted, its structure, and therefore the term "cognitive map" itself (with its association to the Euclidian properties of everyday maps), is in contention (Peer *et al.* 2021). Opponents of the cognitive map theory have theorized spatial knowledge as more egocentrically represented through the structure of a topological graph (Kuipers, Tecuci and Stankiewicz 2003). Recent experimental evidence has opened the debate to a third alternative in which spatial knowledge lies between a Euclidian and topological structure (Ericson and Warren 2020). This structure is termed the "cognitive graph" and is conceived of as a graph "augmented by approximate local distance and angle information (Chrastil and Warren 2014)" (Ericson and Warren 2020: 1). Interestingly, space syntax research on the relationship of urban street network structures and pedestrian movement inadvertently supports this. Hillier and Iida (2005) showed that pedestrian movement in urban settings could best be summarised by modelling street networks as a graph of topological and angular relationships. The cognitive graph theory is also supported by evidence suggesting map and graph knowledge is learnable simultaneously for the same environment (Jacobs and Schenk 2003). However, whether maps and graphs are fundamentally the same thing, or two different spatial knowledge structures that lend themselves to navigation tasks in specific environments, remains unanswered (Peer *et al.* 2021).

Cognitive mapping, as the process through which a cognitive map, or an alternative spatial knowledge structure, is formed, may provide a focal point for research to fuel the debate with crucial evidence. Downs and Stea (1973:7) define cognitive mapping as "[...] a process composed of a series of psychological transformations by which an individual acquires, stores, recalls and decodes information about the relative locations and attributes of the phenomena in his everyday spatial environment". Accordingly, cognitive mapping is a precursor and, supposedly, determinant of the mental structuration of spatial information and studying it promises to reveal more about the structure of spatial knowledge in the mind.

2.2. Collecting and Analysing Cognitive Mapping Data

Lynch's (1960) pioneering work on uncovering how the city is mentally represented through the analysis of sketch maps is considered one of the first cognitive mapping studies. His work inspired a generation of behavioural geography and built environment research based on the theoretical assumptions of the cognitive map theory (Golledge and Stimson 1977). Since cognitive mapping research's early beginnings in the 1960s, two different data collection and analysis approaches have been developed: Behavioural methods and explicit reports (Montello 2016).

Behavioural methods record and study people's behaviour. This includes analysing where people move, what they look and point at and what they say and write. Accordingly, behavioural observations are categorizable into non-verbal and verbal behaviour (Montello 2016). The difficulty with such studies is that the records first need to be converted into data through a coding process. Often records are segmented into relevant units that subsequently are grouped into classes (Boehm and Weinberg 1997). As the coding process is hard, time consuming, and difficult to conduct in a reproducible manner, behavioural studies are marked by significant drawbacks (Montello 2016). However, space syntax, which could be classed as a non-verbal 'behavioural method', has proven itself particularly successful at overcoming coding impediments of the research environment. This is demonstrable by findings such as that the relationship between predictability of urban pedestrian movement flows from spatial network metrics appear to be related to the degree of intelligibility of the network (i.e. the local; global correlation of the spatial network metrics) (Hillier 1989).

Explicit reports make use of tests and surveys to measure the beliefs people express about things. They differ from behavioural methods because of the explicit recognition involved on behalf of the people being studied. Therefore, they are less prone to interpretative biases about people's motivations, internal mental states, and spatial understandings. However, the validity of findings may still be adversely affected by differences in the respondents' memories and truthfulness (Montello 2016). Explicit reports are broadly categorizable into uni-dimensional and two-dimensional tests (Kitchin 2000). One-dimensional cognitive map knowledge is evaluated through uni-dimensional tests. These often consist of different scaling and pointing tasks. (Montello 2016). Two-dimensional tests consist of graphic tasks, completion tasks and recognition tasks that all result in data generated on a single plain, such as a map. Sketch mapping, as a graphic task, is perhaps one of the most popular data collection methods in cognitive mapping research. However, as will be discussed in the following section, sketch maps are notoriously difficult to process. As Kitchin (2000: 12) writes: "[…] [C]ritics sugges[t] that sketch maps are difficult to subjectively score and code [as they] are dependent upon drawing abilities and familiarity with cartography […]".

2.3. The Issue of Coding Sketch Maps into Data

Sketch mapping is a popular data collection technique in cognitive mapping research (see Golledge *et al.* 1985, Moeser 1988, Jeffery *et al.* 2021). Its introduction to cognitive mapping research, as a technique to elucidate information about a subject's environmental knowledge, is widely accredited to Lynch (1960). Despite its popularity, its methodological validity has been criticized. Risks of confounding findings with the graphical skills of the participants are often highlighted. Sketch mapping necessitates an unnatural transformation from the subject's egocentric, three-dimensional perspective of the environment to a two-dimensional, allocentric perspective. Furthermore, the participants' ability to consistently produce identical sketch maps, and its implications on the validity of research has been questioned (see Golledge 1987, Siegel and Cousins 1985, Day 1976, Boyle and Robinson 1979, Downs 1985 or Saarinen 1988). However, Blades (1990) proved through a time-series experiment that sketch maps could be used as consistent sources of environmental data, and in a comparative study Newcombe (1985) found sketch maps to be no less reliable than other techniques in the field of spatial cognition research.

Validity concerns as a data collection method aside, the issue of how to code sketch maps into valid data looms large (Wood and Beck 1976). There is particular concern about retrieving data that transcends the mere ability of the mapper to cope with the task (Wood and Beck 1976). Montello (2016: 174) writes on sketch mapping: "[…] the ease and simplicity of collecting records is not matched by the ease and simplicity of coding and analysing records. […]. Analysing sketch maps is something of a notorious problem in research. One piece of good advice is that you should figure out what kind of information you want to get from the sketch maps, based on what research questions you want to address. There is no omni-relevant way to analyse them." This is reflected in Kuiper's (1983) five-fold categorization of information types contained in cognitive maps: Topological, route descriptions, fixed features, metric and sensory images. Depending on the research, different types of information are of interest, and, accordingly, sketch maps need to be coded case-specifically to arrive at relevant data.

A plethora of qualitative approaches to coding sketch maps have been taken. Some researchers have scored sketch maps by subjectively evaluating map features (Billinghurst and Weghorst 1995). Others tasked a panel of judges to subjectively rate "map goodness" or count the frequency of labelled landmarks and calculate landmark position scores by subjectively gauging how configurationally correct they are to other landmarks (Carassa *et al.* 2002, Coluccia *et al.* 2007, Zanbaka *et al.* 2005). In some cases, maps were evaluated through designed assessment rubrics (Brunyé and Taylor 2008a,b) or by using Likert scales (Woollett and Maguire 2010). In all cases the coding process is tainted by the marker's subjectivity, and the lack of applying a standardized, yet case-specifically flexible, coding approach has resulted in an incongruent and incomparable landscape of publications. As Gardony, Brunyé and Taylor (2015: 152) assess: "[W]hen it comes to sketch map analysis, "reinventing the wheel" appears to be the standard, and this variation in approach makes it difficult to interpret the literature as a whole".

A further issue with the current coding approaches is their disregard (or small regard) of topology and configuration. As sketch maps often are distorted, incomplete and devoid of an underlying metric system, topology and configuration constitute the most reliable types of information that can be gleaned about a subject's environmental knowledge (Golledge and Stimson 1997). This was first acknowledged by Lynch (1960), who was particularly interested in measuring subjects' local topological knowledge. However, the few efforts that have been made to code sketch maps by their topology (see Billinghurst and Weghorst 1995 or Gardony, Taylor, Brunyé 2015) have remained unconvincing, as they lack the methods to engage with more complex configurational aspects of the environment beyond simple, often canonical, local topological relationships between features. Peponis, Zimring and Kyung's (1990) critique of spatial cognition and navigation research identifies this issue on a more general level, asserting that the historical focus of spatial cognition research on cognitive processes has left it methodologically ill-equipped to describe research environments – and by extension sketch maps. Since space syntax's maturation into an established field of urban and architectural research, a general push to introduce the space syntax toolkit of environmental description to spatial cognition research has been made (Dalton, Hölscher and Turner 2012). Kitchin (2000: 18) asserts that the "[s]uccess of cognitive mapping research is dependent on overcoming issues of data validity and integrity", and, as will be discussed in the following section, space syntax may have an important contribution to make towards overcoming these issues.

2.4. Space Syntax and Sketch Maps

The interdisciplinary field of space syntax and spatial cognition has produced a handful of publications analysing the syntactical variables of sketch maps. These predominantly stem from a consensus that wayfinding performance is predictable from topological variables of the environment (Haq and Girotto 2003). By investigating the effects of environmental intelligibility on the configurational accuracy of cognitive representations, Kim (2001) found that inhabitants of more intelligible areas produced better sketch maps. Similarly, Kim and Penn (2004) investigated the linkages between syntactical properties of urban environments and their respective sketch maps. Haq and Girotto (2003) focused on the building scale by analysing the local topological correctness and intelligibility of hospital sketch maps. All these publications pioneered different adaptations of the space syntax methodology to sketch maps.

As the research environments were predominantly open air and of a larger scale, these methodological novelties were based on the axial line model. Axial line models are arrived at by drawing the fewest and longest lines of sight through a plan (Hillier 2007). The resulting model is restrictive in terms of topologically distinguishing between entities. Bruce (2022) identified this research gap and developed a topological graph comparison method based on Hillier and Hanson's (1984) gamma map – the comparative gamma map method - to produce seven different accuracy measures for survey view building sketch maps: Node Count Variance, Edge Count Variance, Node Accuracy, Edge Accuracy, Total Depth Variance, Global Convex Space Variance and Global Window Count Variance. The usefulness of this approach was evidenced through a proof-of-concept study on a pre-existing dataset of 156 small-scale survey-view building sketch maps from Jeffery *et al's* (2021) cognitive mapping research. Importantly Bruce (2022: 23) acknowledged the need to provide and develop the fundaments for a research-specifically adaptive sketch map coding system in cognitive mapping research, writing: "The general philosophy of the approach to harness the topological information embedded in sketch maps leaves potential to tailor and develop the accuracy measures to fit the specific needs of future spatial cognition research.” In line with this spirit, there is a clear research gap in extending the graph comparison method to survey-view building sketch maps of complex buildings. Labelled topological graphs can be abstracted from survey view building sketch maps of small-scale buildings with minor interpretative biases. However, sketch maps of complex buildings undergo two noticeable shifts that necessitate the further development of the original comparative gamma map method. Firstly, they become increasingly unlabelled, as spaces within the system either lose their spatial distinctiveness or the size of the system outgrows the capacity of the human mind to reference all spaces in the sketch map. Secondly, the sketch maps distort more and are increasingly incomplete as the differing graphical skills of the subjects are brought to the fore.

Consequently, the most suitable type of graph to be abstracted from these systems needs to be identified, as well as the impact of the increased ambiguity of the sketch maps on the reproducibility of the accuracy measures assessed. Furthermore, the original accuracy metrics cannot rely on the labelled properties of the sketch graph anymore, which calls for method extensions that transcend graph entity counts and matches to produce accuracy measures that reflect the comparative “deep-structures” of the sketch and baseline graphs.

3. Methodology

3.1. The Original Comparative Gamma Map Method for Survey View Sketch Maps of Simple Buildings

The comparative gamma map method for survey view sketch maps of simple buildings proposed in Bruce (2022) is based on Hillier and Hanson's (1984) gamma map. The gamma map is a topological graph in which the nodes (referred to as vertices in graph theory) represent discrete interior building spaces and the edges the permeability relationships between them. It was introduced as a technique with which to study the configurational properties of interior building spaces. In the gamma map nodes are laid out by topological depth from the building entrance (see **Figure 1**).

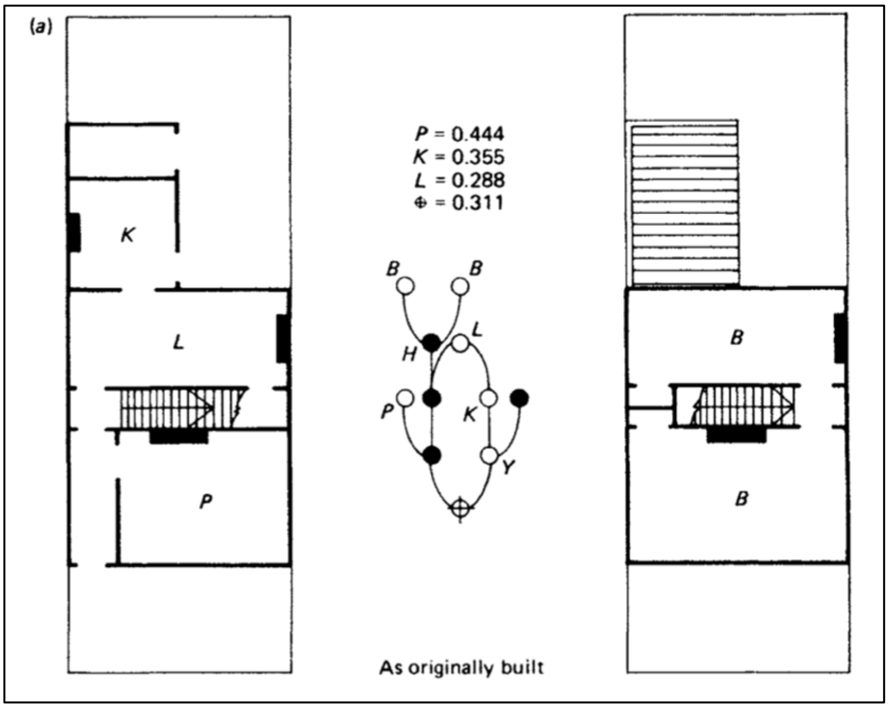


Figure 1: Hillier and Hanson's Gamma Map as abstracted from the example building. Source: Hillier and Hanson (1984: 156)

Bruce (2022) developed a set of sketch map accuracy metrics (see **Table 1**) by abstracting gamma-maps from both the reference floorplan, referred to as the ***baseline graph***, and the sketch map, referred to as the ***sketch graph,*** and comparing them to each other (see **Figure 2**). Accordingly, the most reliable information contained in sketch maps, topology and configuration, could be processed (Golledge and Stimson 1997), and the risk of confounding results by the participant's graphical skills largely averted. The efficacy of this method was proven through a proof-of-concept-study (see **chapter 5** in Bruce 2022).

Table 1: *Sketch Map Accuracy Metrics for Simple Buildings*

Metric	Calculation	Description
Node Count Variance	$\sum V_S - \sum V_B$ Where: V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex	Quantifies the deviation of number of spaces in the sketch map from the baseline map
Edge Count Variance	$\sum E_S - \sum E_B$ Where: E_S is Sketch Graph Edge E_B is Baseline Graph Edge	Quantifies the deviation of number of permeability relationships in the sketch map from the baseline map
Node Accuracy	$\sum (V_S \in V_B) / \sum V_B$ Where: V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex	A percentage score quantifying the share of original spaces in the baseline map contained within the sketch map. If all spaces are included the score is 100%
Edge Accuracy	$\sum (E_S \in E_B) / \sum E_B$ Where: E_S is Sketch Graph Edge E_B is Baseline Graph Edge	A percentage score quantifying the share of original permeability relationships in the baseline map contained within the sketch map. If all spaces are included the score is 100%
Total Depth Variance	$\sum_j \sum_i d(V_{Sj}, V_{Si}) - \sum_j \sum_i d(V_{Bj}, V_{Bi})$ Where: d is distance V_{Sj} is Sketch Graph Origin Vertex V_{Si} is Sketch Graph Destination Vertex V_{Bj} is Baseline Graph Origin Vertex V_{Bi} is Baseline Graph Destination Vertex	Quantifies the extent to whicht the configurational characteristics of the spatial system are changed by changes in the sketch map. This is the closest to measuring a form of allocentric spatial understanding

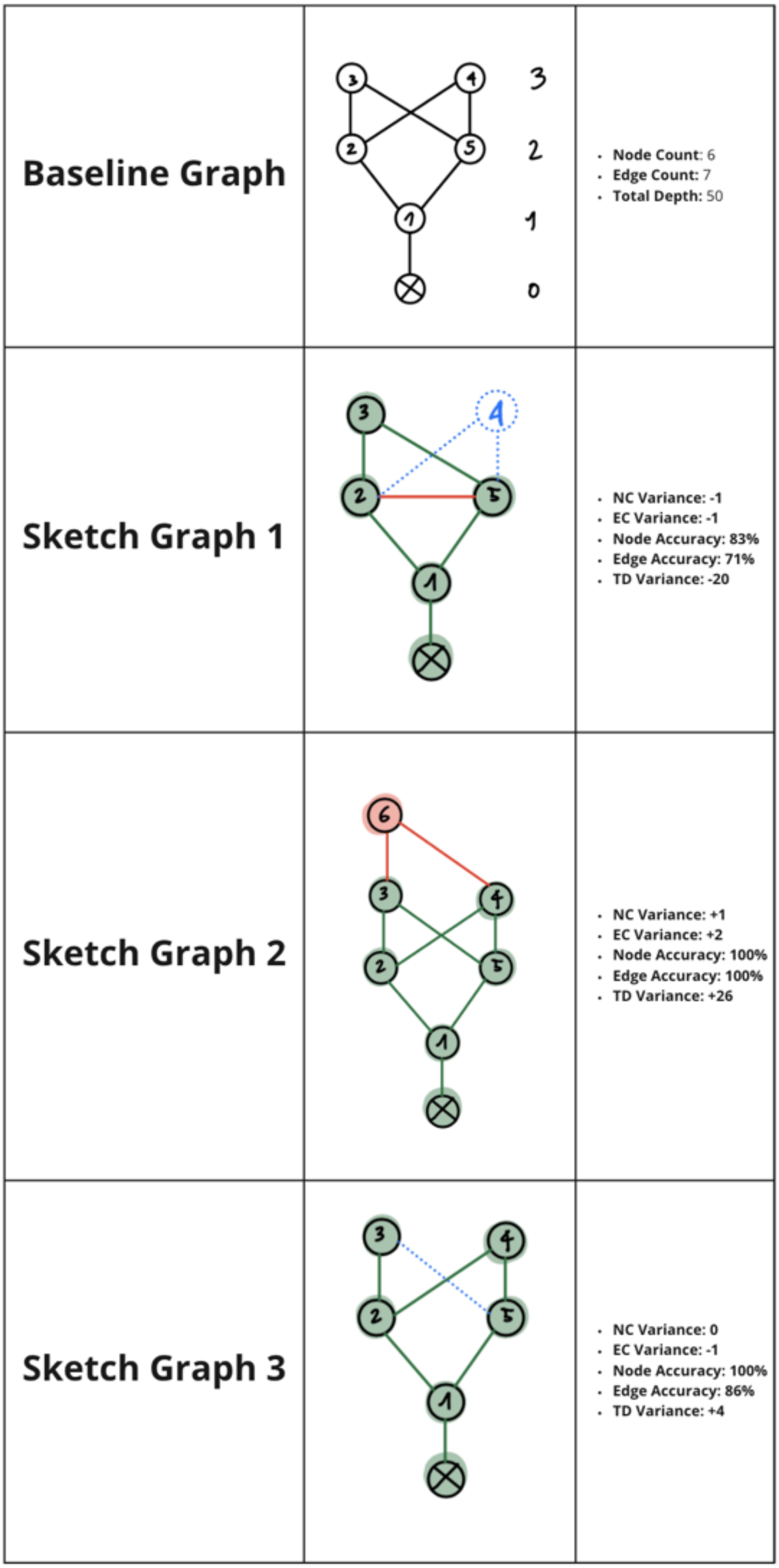


Figure 2*: Comparing sketch graphs and the baseline graph of a simple building through the accuracy metrics proposed by Bruce (2022).*

3.2. Adapting the Comparative Gamma Map Method to Complex Buildings

Sketch graphs abstracted from survey view building sketch maps of complex buildings, opposed to simple building, are unlabelled or partially labelled. Accordingly, global metrics relying on graph entity matches, such as the node accuracy and edge accuracy metrics, are no longer computable, and so are their local equivalents, such as node and edge error omission distributions (see **section 5.5** in Bruce 2022**)**. Furthermore, complex buildings are often composites of multiple sub-complexes connected through circulation systems and, therefore, are not just different in magnitude (e.g., area covered, number of spaces etc.) but also in topo-configurational (how spaces, as topological entities, relate configurationally to each other) properties from simple buildings. Therefore, new metrics measuring aspects of sketch map accuracy for complex buildings are introduced (addressing **RQ1**). These metrics are differentiated into global metrics describing overall topo-configurational differences and a local metric describing the topological neighbourhood similarity of labelled nodes in the sketch and baseline graph.

3.2.1. Global Metrics

The proposed global metrics are summarised in **table 2**. Three metrics, the *node count, edge count and total depth variance*, are applicable to simple and complex buildings. However, as will become apparent from **section 3.3**, they no longer purely represent discrete spaces and their permeability relationships, but a mixture of discrete spaces and discretised hallways and their permeability relationships. Therefore, there is some definitional ambiguity surrounding what constitutes a complex building space in these metrics.

The *node count variance* describes whether a participant could remember the number of spaces present in a building. The location of specific spaces is irrelevant to the metric. The *edge count variance* closely relates to the *node count variance,* as the two influence each other, and expresses whether the participant could remember the number of permeability relationships present in a building. Again, the recall of specific permeability relationships is irrelevant to the metric.

The *total depth variance* measures how configurationally similar the spatial layout in the sketch map is to the baseline map. This metric is inspired by Hillier and Hanson's (1984) research, in which it figures as a variable defining the configurational properties of buildings.

Small differences in total depth values are assumed to signify configurational similarity between sketch and baseline map layout.

The *average degree variance* describes the extent to which the participant has understood the movement potential to elsewhere embodied in an average space of the building.

The *cycle count variance* measures how well the participant has registered ring structures in the building. Especially in complex buildings, sequences of spaces often link up in cycles through which flows of people, goods, ideas etc circulate. Hillier and Hanson (1984) termed "ringiness" as a property of buildings with an abundance of such cycles.

The *A, B, C and D-space share variances* quantify how well the participant has understood the distribution of topologically different types of spaces in the building. The ABCD-space classification of gamma map nodes was proposed by Hillier (2007), who argued that the type of topological embeddedness of a space brings with it different potentials for occupation and movement. A-spaces are dead-end spaces with a single link to other spaces. Movement is only possible to and from them. They are characterised by occupational functions. B-spaces are in tree formed subgraphs of the gamma-map. They are always thoroughfare spaces on-route to an A-space. In tree subgraphs only one movement route exists to and from spaces, which endows B-spaces with great control over accessibility to other spaces. C-spaces lie on exactly one topological cycle of the gamma map. Accordingly, C-spaces are like B-spaces, but are accessible via two neighbouring spaces in two movement directions. Lastly, D-spaces lie on multiple topological cycles of the gamma map and serve as hubs of movement dispersion and confluence (Hillier 2007) (see **Figure 3**).

Table 2: *Global Sketch Map Accuracy Metrics for Complex Buildings*

Metric	Calculation	Description
Node Count Variance	$\sum V_S - \sum V_B$ **Where:** V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex	Quantifies the deviation of number of spaces in the sketch map from the baseline map
Edge Count Variance	$\sum E_S - \sum E_B$ **Where:** E_S is Sketch Graph Edge E_B is Baseline Graph Edge	Quantifies the deviation of number of permeability relationships in the sketch map from the baseline map
Total Depth Variance	$\sum_j \sum_i d(V_{Sj}, V_{Si}) - \sum_j \sum_i d(V_{Bj}, V_{Bi})$ **Where:** d is distance V_{Sj} is Sketch Graph Origin Vertex V_{Si} is Sketch Graph Destination Vertex V_{Bj} is Baseline Graph Origin Vertex V_{Bi} is Baseline Graph Destination Vertex	Quantifies the extent to which the configurational characteristics of the spatial system are changed by changes in the sketch map. This is the closest to measuring a form of allocentric spatial understanding
Average Degree Variance	$\frac{1}{n_S}\sum D(V_S) - \frac{1}{n_B}\sum D(V_B)$ **Where:** n_S is number of vertices in Sketch Graph n_B is number of vertices in Baseline Graph $D(V_S)$ is degree of Sketch Vertex $D(V_B)$ is degree of Baseline Vertex	Quantifies the extent to which the average movement choices between spaces in the sketch map deviate from those in the baseline map. The degree of a vertice V, formulated as D(V), expresses the number of connected, neighbouring vertices V has (see Clark and Holton (1991)).
Cycle Count Variance	$\sum C_S - \sum C_B$ **Where:** C_S is Cycle in Sketch Graph C_B is Cycle in Baseline Graph **Sidenote:** A cycle is a closed sequence of discrete edges in which the start and end vertex are the same (see Clark and Holton (1991: 29) for an exact definition)	Quantifies the extent to which the sketch map reflects the ring structure (or ringiness) of the building complex represented by the baseline map.

ABCD-Space Share Variance A-Space Share Variance B-Space Share Variance C-Space Share Variance D-Space Share Variance	$A\Delta = \frac{1}{n_S}\sum A(V_S) - \frac{1}{n_B}\sum A(V_B)$ **Where:** n_S is number of vertices in Sketch Graph n_B is number of vertices in Baseline Graph $A(V_S)$ is a classification function that returns a 1 if the Sketch Vertex is an A-space and 0 otherwise $A(V_B)$ is a classification function that returns a 1 if the Baseline Vertex is an A-space and 0 otherwise **Sidenote:** The same formula is applied for the B, C and D-Space Share Variance calcualtion	A set of four variables quantifying the extent to which the sketch map reflects the distribution of topologically different types of spaces in the baseline map. A, B, C and D spaces differ by their topological embeddedness in the building and accordingly reflect different potentials for occupation and movement (Hillier 2007).
Cluster Count Variance	$\sum C_S - \sum C_B$ **Where:** C_S is a Sketch Graph cluster C_B is a Baseline Graph cluster **Sidenote:** Clusters are computed using the Louvian method, which is a modularity maximising algorithm (Traag, Waltman and Van Eck 2019)	Often complex building systems are a composite of sub-complexes. This metric attempts to detect these through topological community analysis of the nodes in the sketch and baseline graphs. If the Baseline and sketch map have the same cluster count the baseline map's sub-complex composition may have been correctly understood.
Isomorphism Test	Using the Weisfeiler-Lehmann graph hashing technique (also referred to as graph colouring) whether two graphs are topologically isomorphic is tested (Shervashidze *et al.* 2011)	A binary variable expressing whether the baseline and sketch map layout are definitely different or likely the same.

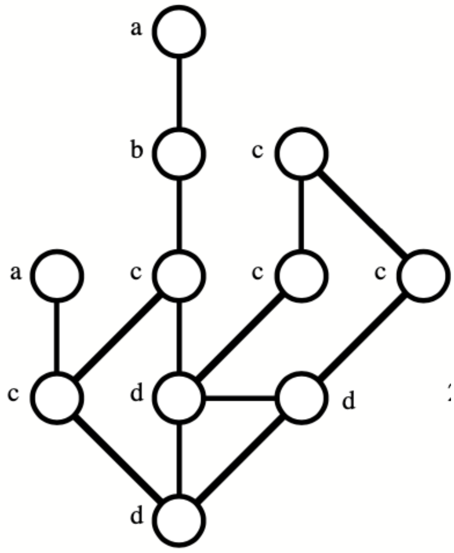


Figure 3*: ABCD-space classification of a gamma map. Source Hillier (2007: 249).*

The *Cluster Count Variance* intends to capture how well the participant has understood the sub-complex composition of the building by comparing the number of node clusters detected in the sketch graph to that of the baseline graph. Clusters are detected using the Louvian algorithm, which is an unsupervised community detection algorithm for graph vertices. The algorithm partitions a graph into sub-groups of vertices to maximise graph modularity. Modularity measures the density of connected vertices within sub-groups by comparing them to the expected density in a random graph. Maximising the modularity measure leads to denser connections of vertices within the sub-groups than between sub-groups (for a detailed description see Trag, Waltman and van Eck 2019). This measure should be used carefully, as there is no guarantee that clusters detected in the baseline graph correspond with the actual sub-complex structure of the building.

Lastly, the *Isomorphism Test* is a binary variable expressing whether the layout of spaces in the sketch and baseline map are different or likely the same. Graphs are isomorphic if they have the same number of vertices and an edge set of corresponding relations. The Weisfeiler-Lehman algorithm is used to identify isomorphic graphs (see Shervashidze *et al.* 2011). In a first step it assigns each vertex an initial colour. Then a signature string is created for each vertex by concatenating its colour with the colours of its neighbouring vertices. In a second step all graph vertices are recoloured by their new signatures. These steps are repeated until convergence. The result is a list of vertex signatures (a graph-hash to be exact) describing the exact topological embeddedness of each vertex in the graph (see **Figure 4**). If they are identical, the graphs are likely isomorphic, otherwise they are not. This metric must be used cautiously, as two functionally different buildings can have isomorphic gamma maps. As Dalton and Kirsan (2007: 814) write: "[…] [T]he use of labelled graphs does seem crucial from an architectural standpoint. For example, it would be possible to have two buildings which were identical from the point of view of their graphs, but fundamentally different from a functional perspective". However, the aim is to compare sketch maps to a baseline map that, presumably, references the same building. This reduces that risk. **Figure 5** visualises how the accuracy measures are applied in practice on a baseline graph which has undergone several transformation scenarios that represent sketch graphs.

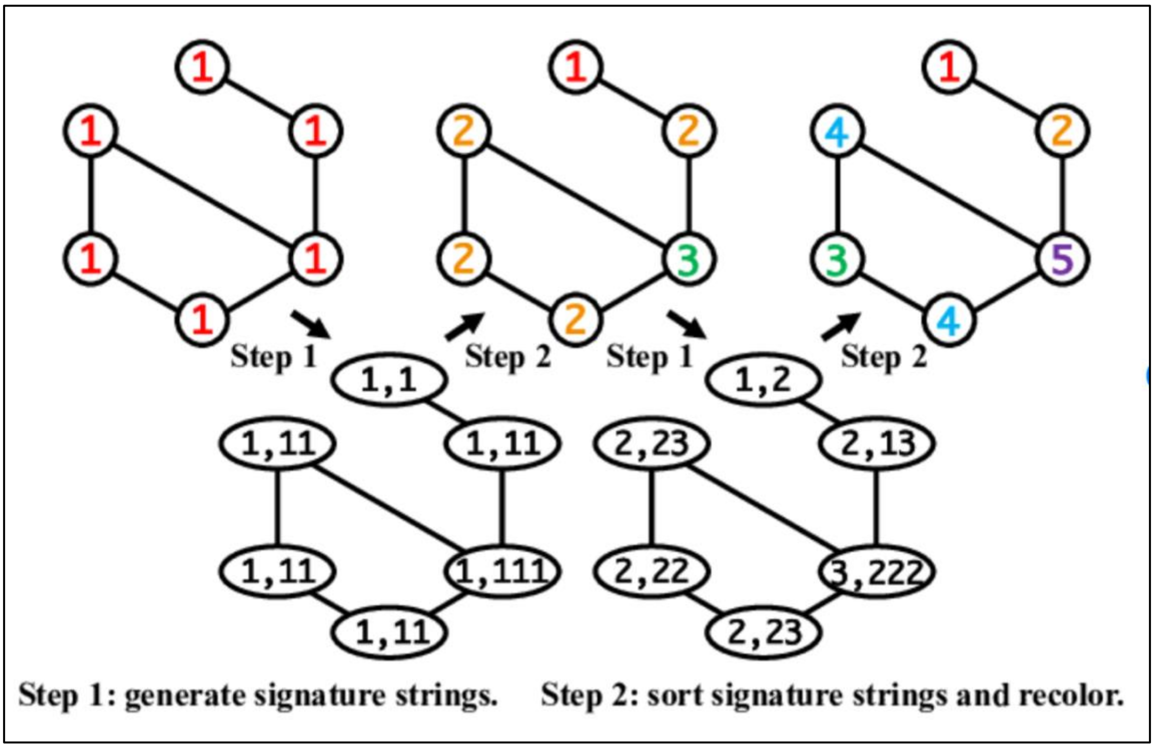


***Figure 4**: Illustration of two iterations of the Weisfeiler-Lehman algorithm. Source: Zhang and Chen (2017: 577).*

Figure 5: *Comparing sketch graphs and the baseline graph through the accuracy metrics.*

3.2.2. Proposing a Local Metric

The global metrics rely on comparing unlabelled structural aspects of the sketch and baseline graph. However, sometimes sketch graph nodes are labelled, especially when the participant is referring in the sketch map to spaces of significant importance or differentiation within the building complex. The purpose of this section is to introduce, explore and adapt the *signature similarity metric between nodes* introduced by Milenković and Pržulj (2008) as a local metric to compare topological neighbourhood similarities between corresponding labelled nodes in the sketch and baseline graph. The aim is to propose a measure that quantifies how well the participant understood the spatial context of labelled spaces.

3.2.2.1. Graphlets, Graphlet Degree Vectors and Node Similarity Signatures

The *signature similarity metric between nodes* was invented by Milenković and Pržulj (2008) as a bioinformatical method to quantify the local topological similarity of proteins in protein-protein interaction networks. In Milenković and Pržulj's (2008: 258) own words: "[…] [The] node similarity generalizes the degree of a node […] into the vector of *graphlet degrees*, counting the number of graphlets that the node touches; *graphlets* are small connected non-isomorphic induced subgraphs of a large network (Pržulj *et al.* 2004)". The graphlet degree vector then summarizes the frequency with which each graphlet orbit is touched upon by a node for all 2-5 node graphlets depicted in **figure 6**. Larger graphlet counts are not included due to increasing computational complexity. As often graphlet nodes are positionally symmetrical to each other, orbits are defined as the uniquely distinguishable positions within the graphlet that can be touched upon by the node. For all 2-5 node graphlets there are 73 orbits, resulting in a graphlet degree vector of length 73 for each node.

To calculate the *signature similarity* $S(u, v)$ between nodes u and v, which expresses how similar the graphlet degree vectors of u and v are to each other, five computations are necessary (Milenković and Pržulj 2008):

1) The graphlet degree vector V_n for each node n is computed.
2) Orbit weights are calculated to account for the dependency of some orbits on others. Higher weights are assigned to "important" orbits less affected by others. Low weights are assigned to "redundant" orbits dependent on many others. The weight w_i for orbit i is calculated followingly:

$$w_i = 1 - \frac{\log(o_i)}{\log(73)}$$

Where:

o_i is the count of orbits affecting i.

3) After obtaining the orbit weights w_i the distance $D_i(u,v)$ between the ith orbits of nodes u and v is calculated:

$$D_i(u,v) = w_i \times \frac{|\log(u_i + 1) - \log(v_i + 1)|}{\log(\max\{u_i, v_i\} + 2)}$$

4) The total distance $D(u,v)$ between nodes u and v is then calculated:

$$D(u,v) = \frac{\sum_{i=0}^{72} D_i}{\sum_{i=0}^{72} w_i}$$

5) Lastly, the signature similarity $S(u,v)$ between nodes u and v is calculated:

$$S(u,v) = 1 - D(u,v).$$

This results in a measure of signature similarity between nodes that runs from 0 to 1, where nodes with an identical local topological neighbourhood take on a value of 1 and nodes with completely dissimilar neighbourhoods a value of 0.

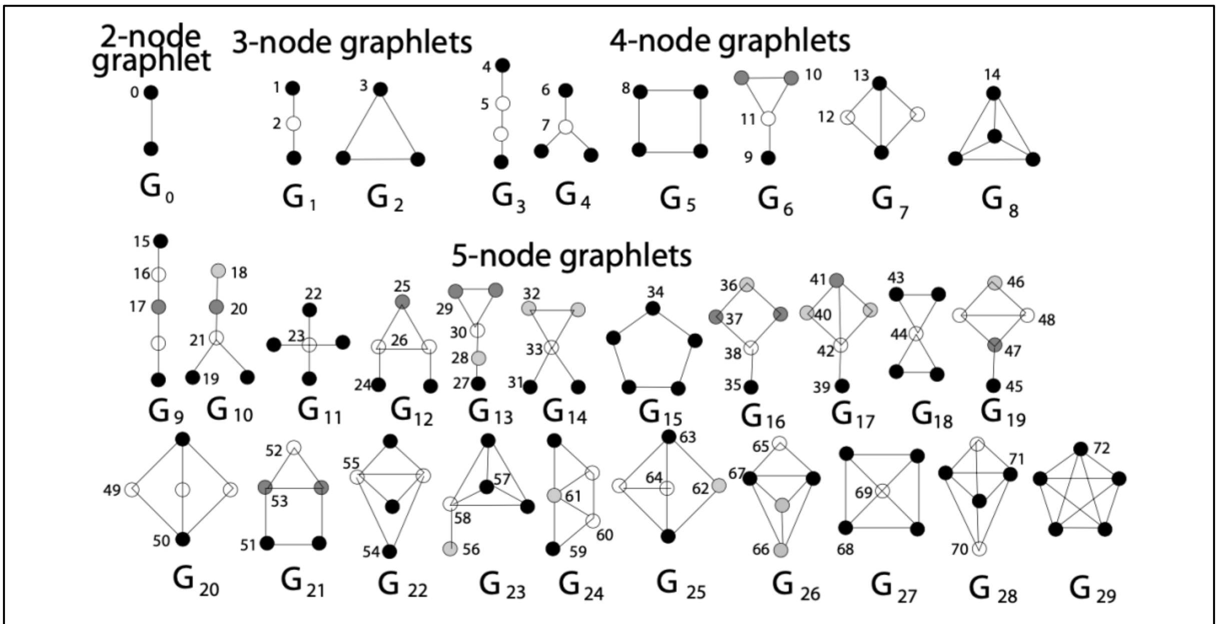


Figure 1. The thirty 2-, 3-, 4-, and 5-node graphlets $G_0, G_1, \ldots, G_{29}$ and their automorphism orbits 0, 1, 2, ..., 72. In a graphlet G_i, $i \in \{0, 1, ..., 29\}$, nodes belonging to the same orbit are of the same shade (Pržulj, 2006).

Figure 6: *All 2-5 graphlet node orbits. Source: Milenković and Pržulj (2008: 259).*

3.2.2.2. Exploring and Adapting the Node Signature Similarity Measure to a Mirror Symmetrical Building

The node signature similarity measure can be used to compare the topological neighbourhood similarity of labelled nodes in the sketch graph to those in the baseline graph. To understand how the signature similarity measure works for complex buildings, the pairwise similarity of all building spaces, as graph nodes, to a specific building space can be computed (node signature self-similarity), plotted, and examined.

In **figure 7** node signature self-similarities for a mirror symmetrical building also used in **chapter 5** are plotted (see **B1** for all self-similarity value tables). The similarities are computed for a *hybrid-axial graph* model of the building – more detail on this graph is provided in **section 3.3**. In the first plot, node 21 is set as the reference node for the pairwise comparisons and the resulting signature similarities are plotted. Node 21 represents an enclosed dead-end space arranged around one of the two building courtyards. Darker node colours denote more similarity to node 21 and lighter ones less. The first, and all subsequent node-comparison plots in **Figure 7**, show that the signature similarity measure is categorising the nodes into five groups with similar topological neighbourhoods (especially visible in the pairwise comparison plot to node 10) – these groups will be considered in more detail in **section 3.2.2.3**. This highlights that many spaces within the spatial configuration of this building layout are topologically equivalent to each other. For example, all dead-end spaces around the courtyards have the same topological neighbourhood, and accordingly are considered the same by the signature similarity measure. For spatial cognition research this level of differentiation between spaces may be too spatially reductive to assess how well labelled spaces are understood in their spatial context. After all, more complex spatial relationships such as adjacency or ego- and allocentric left and right-sidedness of spatial feature relations are disregarded for pure topological permeability relationships between spaces. The following paragraphs demonstrate how more spatial information can be incorporated into the graph through the addition of edges and labels.

Figure 8 shows the self-similarity measures for an *adjacency graph* model of the building. In the *adjacency graph* edges are added to the *hybrid-axial graph* that represent adjacency relationships between spaces. The result is a more fine-grained categorisation of self-similarities in which nodes are differentiated by their lateral “distance” from the local bi-lateral symmetrical axes of the courtyards.

Global and local left and right-sided relations of the spaces can be introduced to the *adjacency graph* to render each node topologically unique (see **figure 9**). This is done by introducing one large and two small *L-nodes* which indicate spaces left of the global bi-lateral axis of the building layout and/or left of the local bi-lateral axis of the courtyards. The resulting graph is an *L-node graph*. Of course, this graph caters to this building's specific symmetrical properties and relative explicitness of what is left and right sided and is less easily reproducible on other buildings. Furthermore, dead-end spaces across from each other in the right courtyard are deemed more similar by the measure than those in the left courtyard (presumably as the nodes there are differentially touching on less graphlets). However, for now, this invention solely serves to illustrate how the graph can be augmented to incorporate further spatial information into the node signature similarity measure, and the extent to which unwanted noise is introduced necessitates further research.

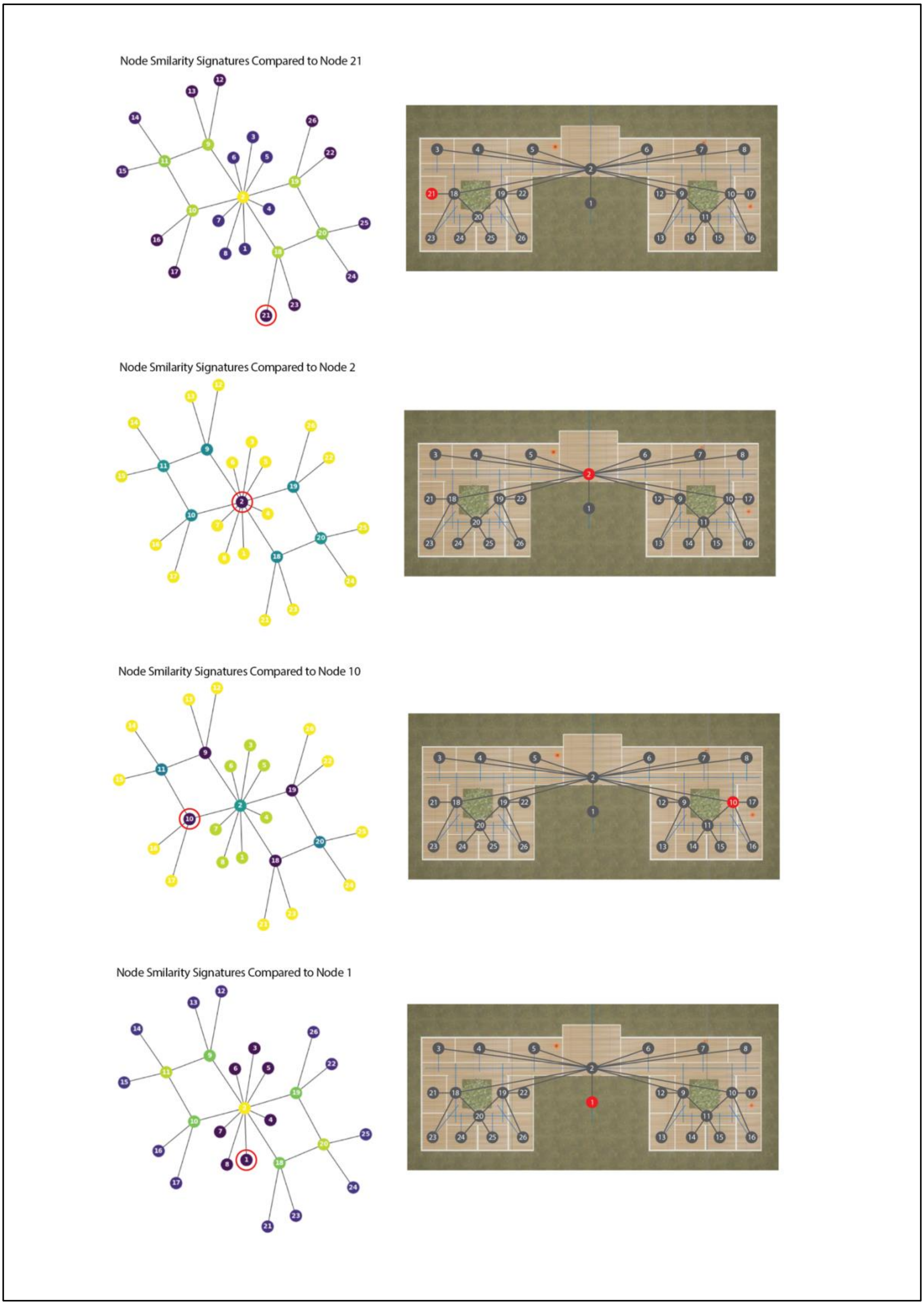


Figure 7: *Node signature self-similarity measures for a hybrid-axial graph modelled mirror symmetrical building.*

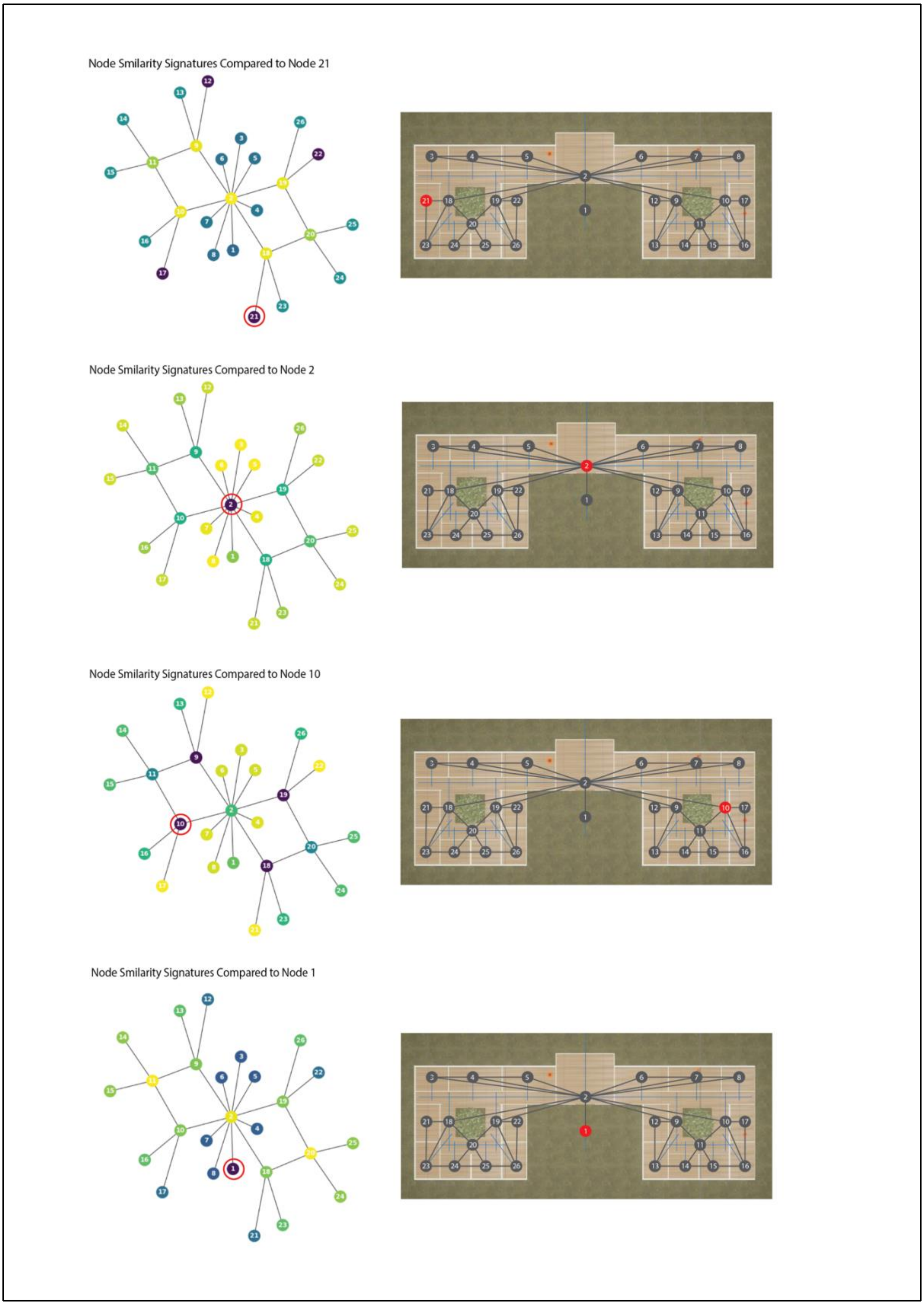


Figure 8: *Node signature self-similarity measures for an adjacency graph modelled mirror symmetrical building.*

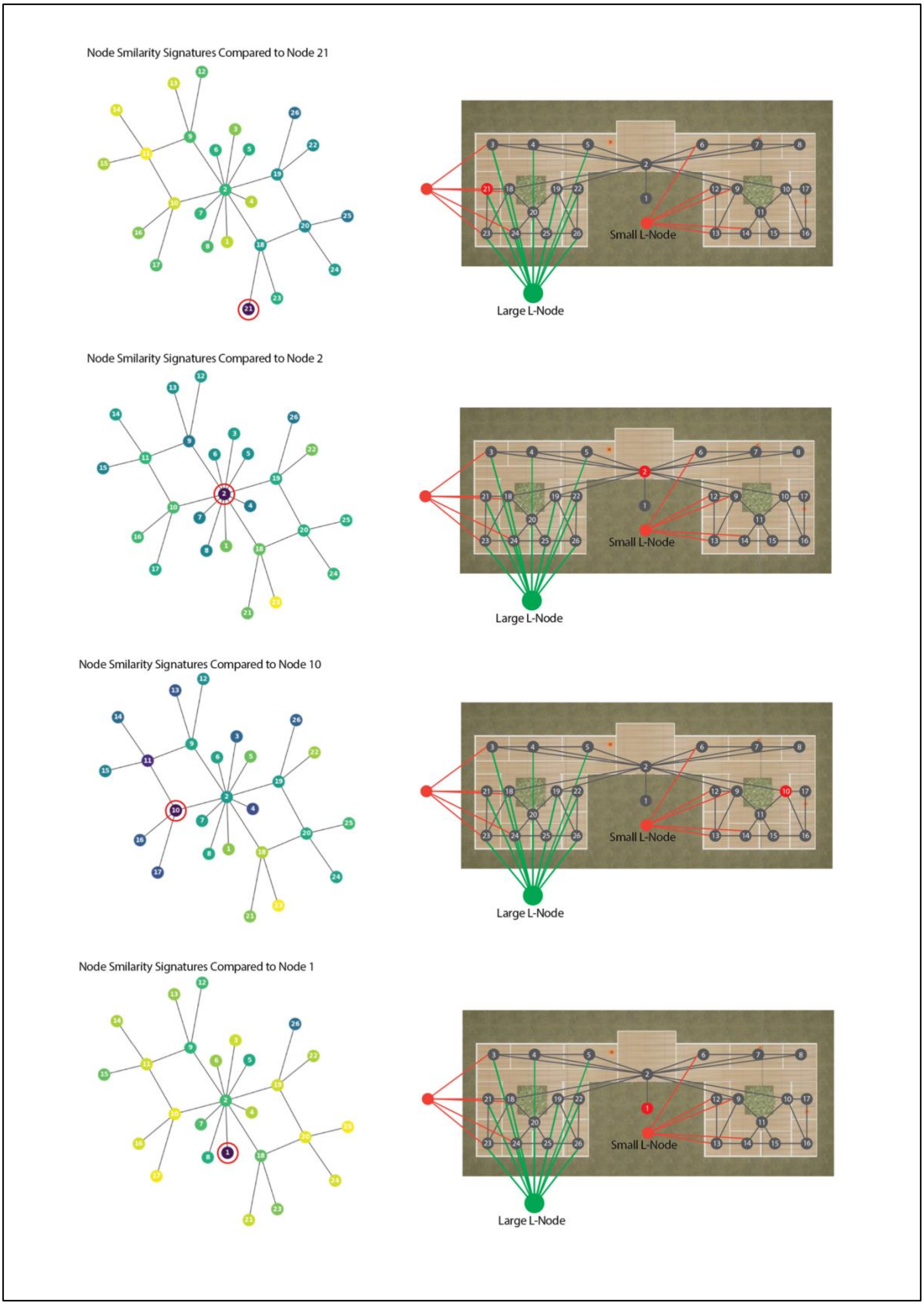


Figure 9: *Node signature self-similarity measures for an L-Node graph modelled mirror symmetrical building.*

3.2.2.3. Conceptually Extending the ABCD-space Classification

One of the striking things about how the signature similarity measure categorises the building spaces into five groups in **figure 7** (especially visible in the pairwise comparison plot to node 10), is that the groups correspond to the "distributedness" of the spaces, that is, whether they are destination spaces or movement spaces lying on cycles. Therefore, the signature similarities correspond closely with Hillier's (2007) ABCD-space categorisation introduced in **section 3.2.1**. However, the measure is more discerning of different types of destination and circulation spaces by considering their extended topological embeddedness. This is demonstrated in **figure 10**, where the five-fold categorisation of the spaces in **figure 7** is reproduced by applying one iteration of the Weisfeiler-Lehman algorithm to the ABCD-space categorised *hybrid-axial graph* of the building. Therefore, the measure does not just differentiate between A, B, C and D spaces but also by how they connect to other A, B, C and D spaces neighbouring them to form new categories such as AB, AC, AD spaces and so on. Accordingly, by applying the logic of the Weisfeiler-Lehman algorithm to the ABCD-space classification, it can be conceptually extended, which is what the signature similarity measure appears to be doing.

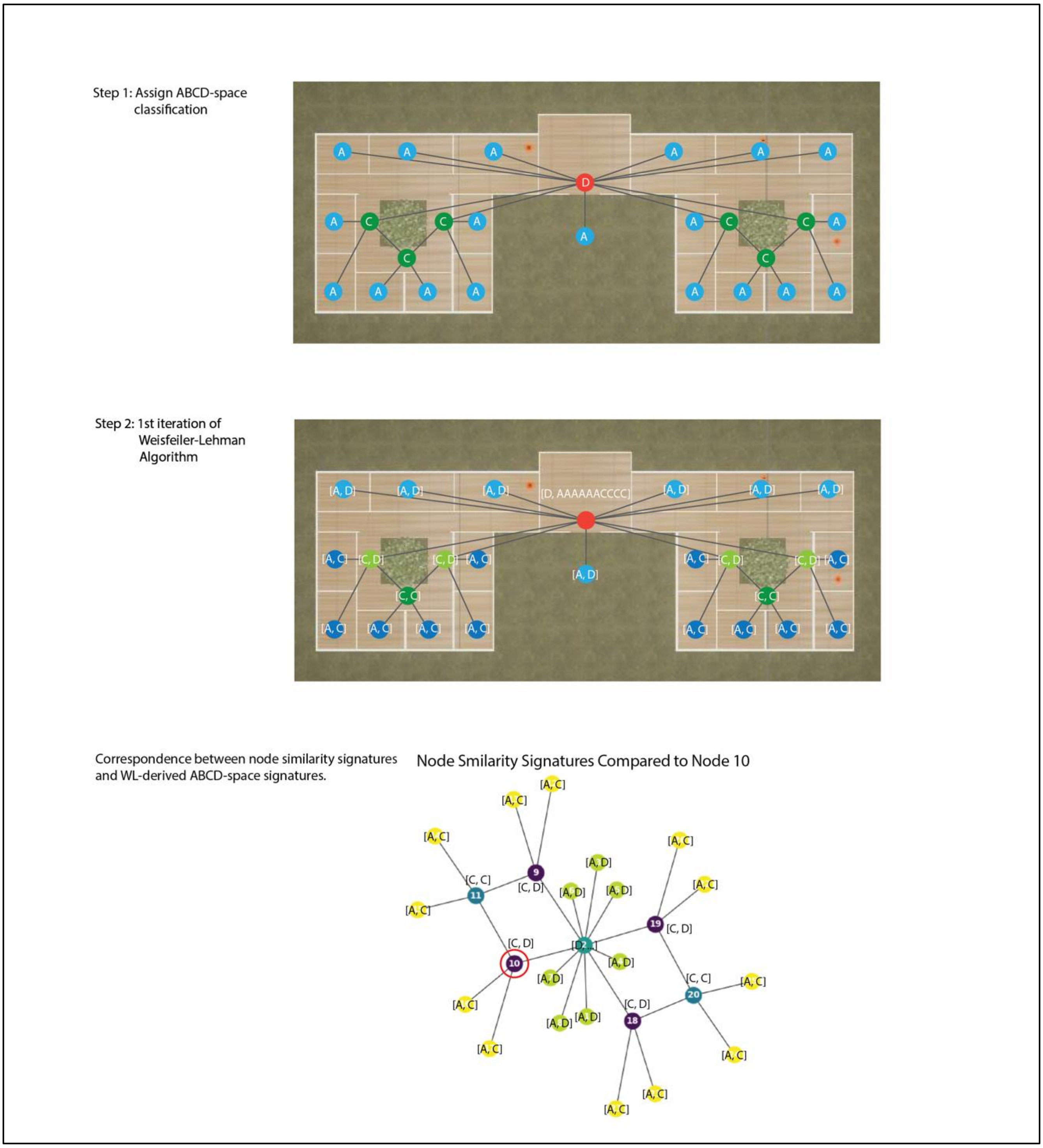


Figure 10: *Applying 1st iteration of the Weisfeiler-Lehman algorithm to an ABCD-space categorized gamma map to derive the same grouping as with the signature similarity measure.*

3.3. Introducing the Hybrid-Axial Graph

As the topo-configurational structure of complex buildings is different to simple buildings, a different graph than the gamma map is required to model them (addressing **RQ2**). In the gamma map each enclosed space is represented as a node and the permeability relationships between them as edges. This suffices to describe most simple buildings (and was deployed as such in Bruce (2022)'s study). However, this representation smooths over the structural aspects that differentiate complex from simple buildings; namely, the subcomplexes and circulation systems. Accordingly, the field of space syntax has applied axial models to study more complex buildings (e.g. see Hillier and Penn 1991). In an axial model the longest and fewest lines of sight are drawn through a plan (Hillier 2007) (see **figure 11**). From a graph perspective, the lines represent nodes and their intersections the edges. Accordingly, hallways are discretised into spaces by following a logic on how space is cognitively perceived (Dalton 2005). However, the axial model, in its pure application, often leads to enclosed and open spaces being conflated into one line (or node from the graph perspective). This is not particularly useful for spatial cognition research, as the graph should ideally represent enclosed spaces separately as well as embody the structural particularity of complex buildings. Accordingly, a *hybrid-axial* graph representation is suggested in which circulation spaces are discretised into nodes axially and the enclosed spaces are represented as separate nodes in the graph, all of which are connected by edges according to their topological permeability relationships. See **figure 12** for an example on how to abstract the *hybrid axial graph.* As touched upon earlier, compared to

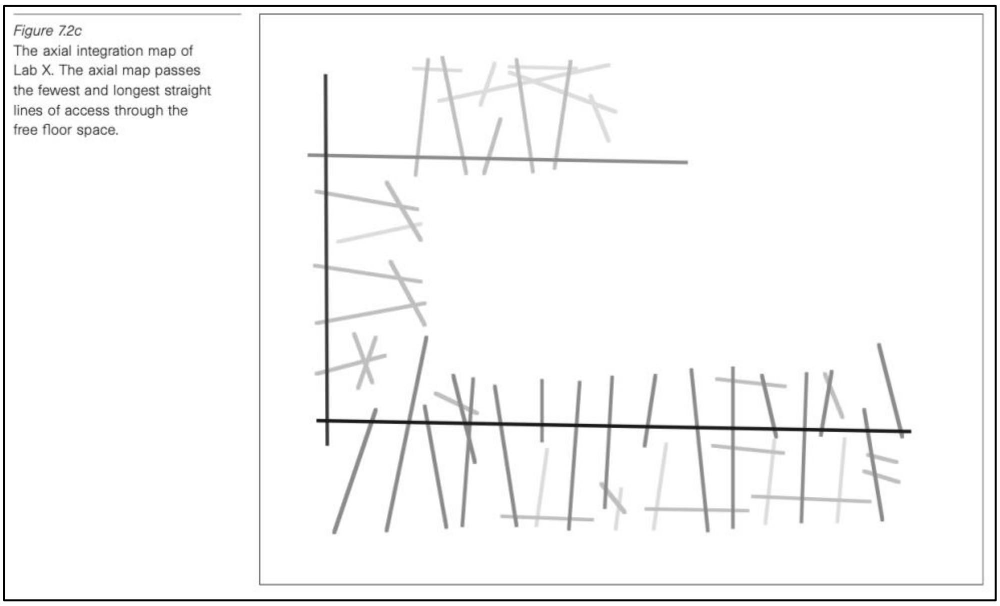


Figure 11: *Axial map of a laboratory. Source: Hillier (2007: 205).*

the gamma map, the resulting definition of a space as represented through the hybrid axial graph nodes is more ambiguous.

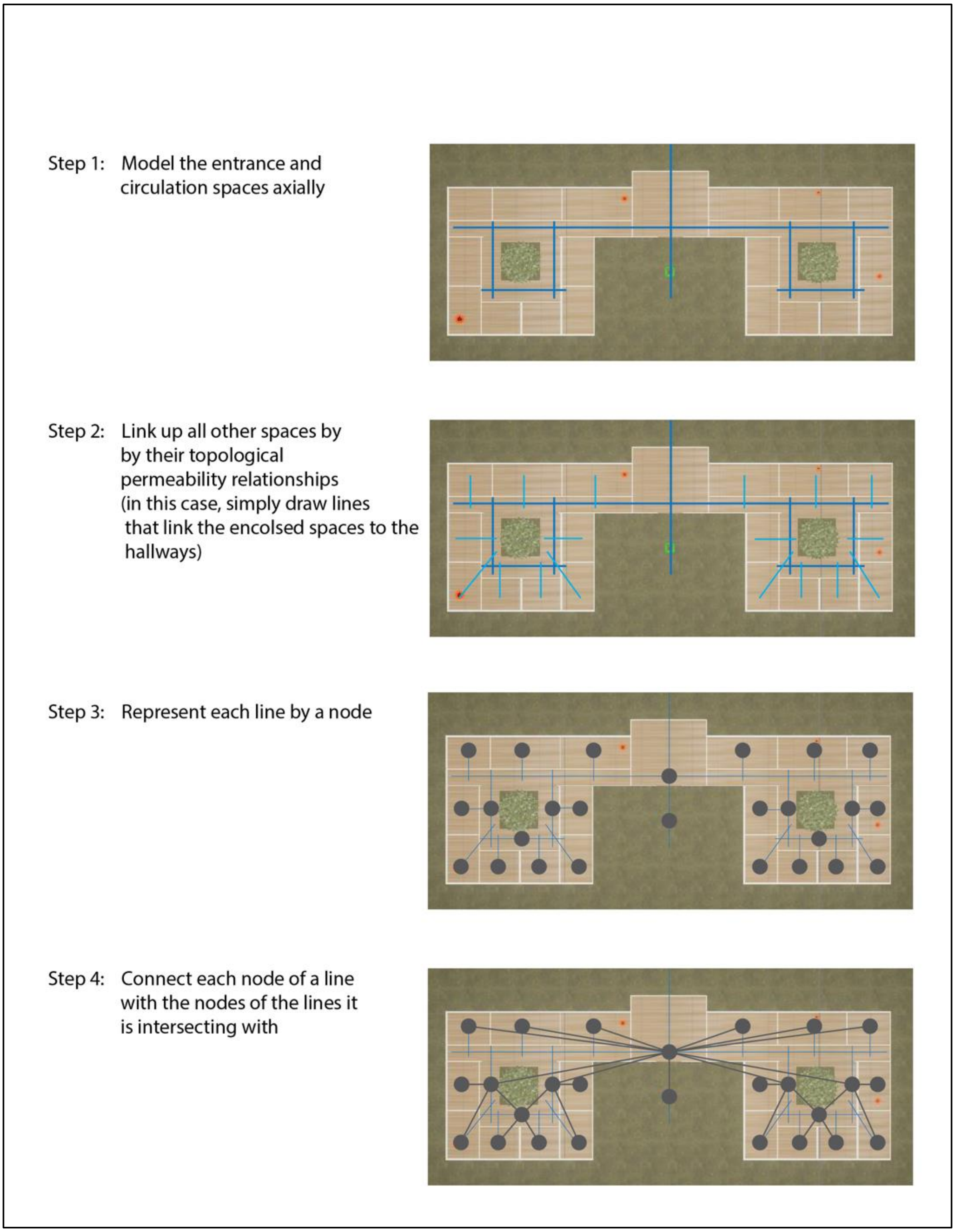


Figure 12: *How to create the hybrid-axial graph.*

3.4. Digitizing the Analysis

In the following chapters the methods are applied to a survey view sketch map data set of complex buildings. All global metrics were computed with a custom graph analyser written in python and the resulting outputs analysed in R. (see **B2**). NetworkX (NetworkX 2022), which is a python package for creating and manipulating graph objects, was an essential component of the analyser. Furthermore, a node-signature self-similarity analyser was programmed in python for the analysis in **section 3.2.2.2** which takes a matrix of all graphlet degree vectors of a building as an input (see **B2**). Graphlet degree vectors were computed in R. (see **B2**) with the orca package (Hočevar and Demšar 2016).

4. Reproducibility

This chapter establishes initial benchmarks of measurement integrity and reliability for different sketch-map analysis methods. As subjective scoring and ranking approaches are widely used in the field, their reliability, as measures of sketch map accuracy, is compared to the comparative gamma map method proposed in this dissertation. Two investigative strands are followed. The first concerns the reliability of scoring and ranking measures (addressing **RQ3**) and if, as proposed in the introduction, they are biased by the perceptible shifts from simple to complex building sketch maps (addressing **RQ4**). The second concerns the strength of consensus in abstracting sketch graphs (addressing **RQ3**) and the extent to which this is affected by the perceptible shifts from simple to complex building sketch maps (addressing **RQ4**). A simple and complex building sketch map sample were used in the benchmarking, both of which were collected by Professor Kate Jeffrey at the Department of Behavioural Psychology, UCL.

4.1. Survey

The insights in the following sections are based on data collected through a survey circulated among 29 built environment students and professionals, the majority of which have a background in architecture (**see B3**). This demographic was specifically selected for its working familiarity with floorplans. Participants were tasked with ranking and scoring a sample of 12 randomly selected simple and complex building sketch maps by their accuracy compared to the baseline map. Ranks ranged from 1 to 12, with 1 being the most accurate sketch map and 12 the least accurate. Scores ranged from 1 to 5, 1 being “no resemblance” and 5 being “practically identical”. The order of the sketch maps was randomised, and half the participants received the reversed, randomised order to avoid influencing results by the order of presentation. A subset of the participants with a background in space syntax was tasked to abstract hybrid-axial graphs from a sample of 9 randomly selected complex building sketch maps. This task was partially supervised, but guidance was kept to a minimum to assess both the accessibility of the method as well as issues impeding optimal graph production.

4.2. Ranking and Scoring Consensus

Sketch map ranks deviate on average by **2.3** for complex and **2.1** for simple buildings. Sketch map scores deviate on average by **1** for complex and **0.9** for simple buildings. Therefore, ranking and scoring consensus is similar for both complex and simple buildings, and the difference statistically insignificant at a confidence level of 95% (two-sided Welch t-test, p = 0.3 (rank comparison) and p = 0.4 (score comparison)). These results show that both ranking and scoring is quite subjective, as there is some disagreement.

Figure 13 depicts the scattergrams and Pearson correlation statistics for scores against ranks of the simple and complex buildings at the data collection level and aggregated. At the aggregated level there are twelve data points representing average ranks and scores received by each sketch map. The statistically significant correlations (at a 95% confidence level) at both levels reveal that the score and rank measure are capturing a similar concept of measurement accuracy (level of the data: ***R*=-0.76,** ***R*=-0.77**, aggregate level: ***R*=-0.96,** ***R*=-0.97**).

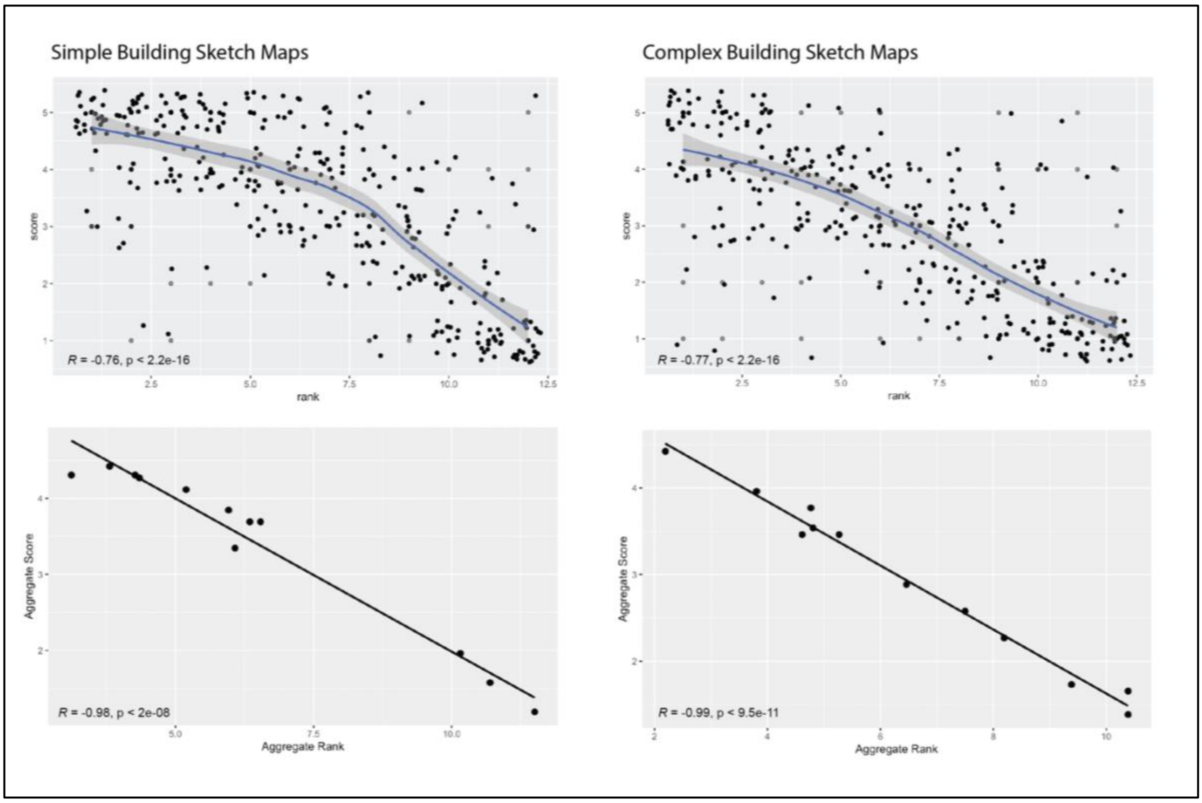


Figure 13: *Investigating the relationship between scores and ranks for simple and complex building sketch maps.*

Importantly, the aggregate distributions reveal that simple buildings are characterised by a binary "good/bad" distinction in average scores and ranks compared to a graduated linear spread for the complex buildings. This could indicate that sketch maps of simple buildings differentiate themselves from complex buildings by either a binary distribution of map quality or judgement. Accordingly, the assumption in the introduction, that the graphical abilities of the sketch map producing participants is more emphasised in complex buildings, may be evident, leading to a more graduate distinction in score and ranks from "bad" to "good". This poses questions about the suitability (weaknesses) of subjective scores for different building scales. For one it could be that the "good/bad" distinction in simple buildings leads to over generalisations. Whereas the heightened graphical confoundedness in complex buildings may lead to an overemphasis of differences.

By plotting the standard deviation against the average rank and score received, **figure 14** reveals that simple and complex building sketch maps also differ in where there is consensus on map quality. Simple building score and ranking consensus is highest for "bad" maps and lowest for "good" maps. In contrast, the complex building scattergrams seem to form an inverse hyperbole, indicating consensus is highest at both extremes of map quality and lowest in the middle ground. However, a larger sample size is needed to confirm this trend.

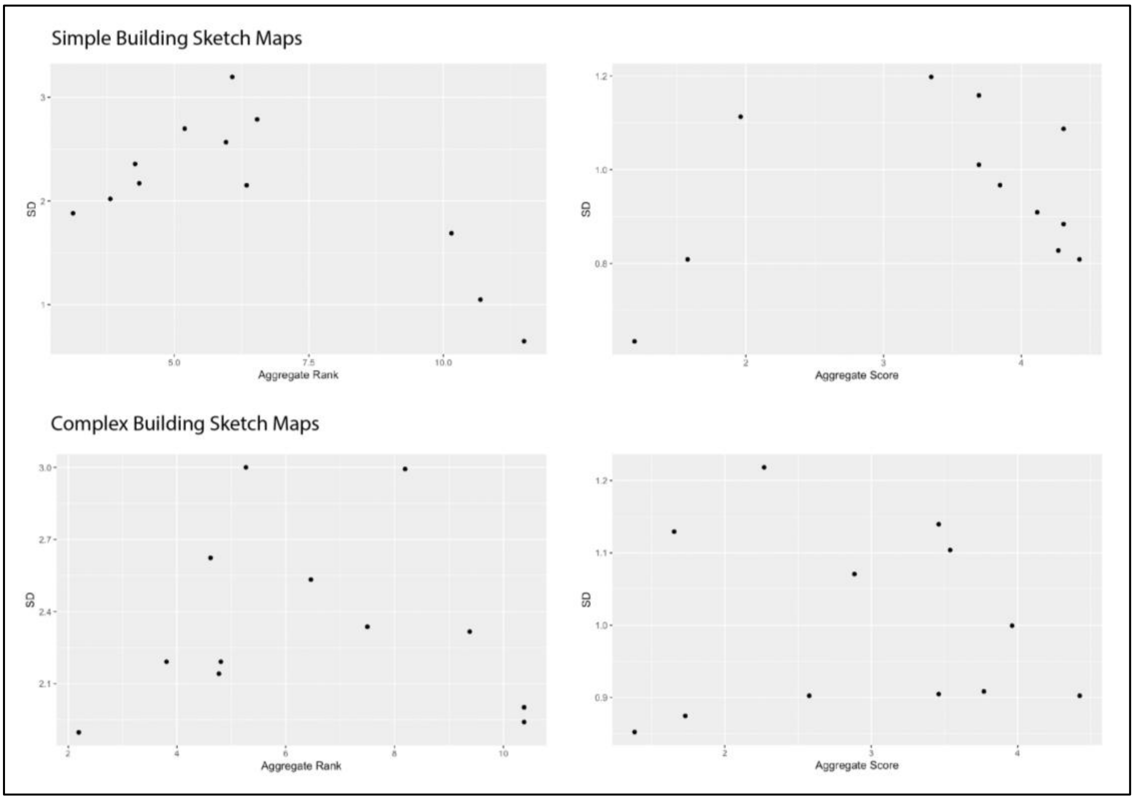


Figure 14: *Investigating score and rank consensus for simple and complex building sketch maps.*

Taken together, these findings suggest that when using ranking and scoring approaches there is some disagreement between practitioners on map quality. Furthermore, the identified shifts from simple to complex building sketch maps seem to lead to different map quality consensus outcomes for simple and complex buildings.

4.3. Hybrid-Axial Graph Consensus

22 from 29 participants completed the hybrid-axial graph abstraction task, out of which **63% correctly or partially correctly produced graphs**. The most common reason for not completing the task correctly was a lack of understanding on how to produce the graph, resulting in e.g. disconnected graphs. This highlights that the method is not immediately accessible to everyone, and, beyond an applied Space Syntax training, an introduction to the fundamental properties of graphs is necessary.

7 participants produced partially correct graphs. The most common errors were inclusions of impossible links (e.g. between spaces that are physically not accessible to each other) or the introduction of C-space paradoxes. A C-space paradox occurs when an A-space situated at the intersection of two hallways is topologically linked to both hallways and thereby introduces meaningless cycles into the graph (see **figure 15**). Underlying this paradox is the general issue of handling corner spaces where the allocation to a hallway becomes ambiguous. Such issues are solvable but require a combination of modelling and data collection rules (the likes of which exceed the bounds of this dissertation).

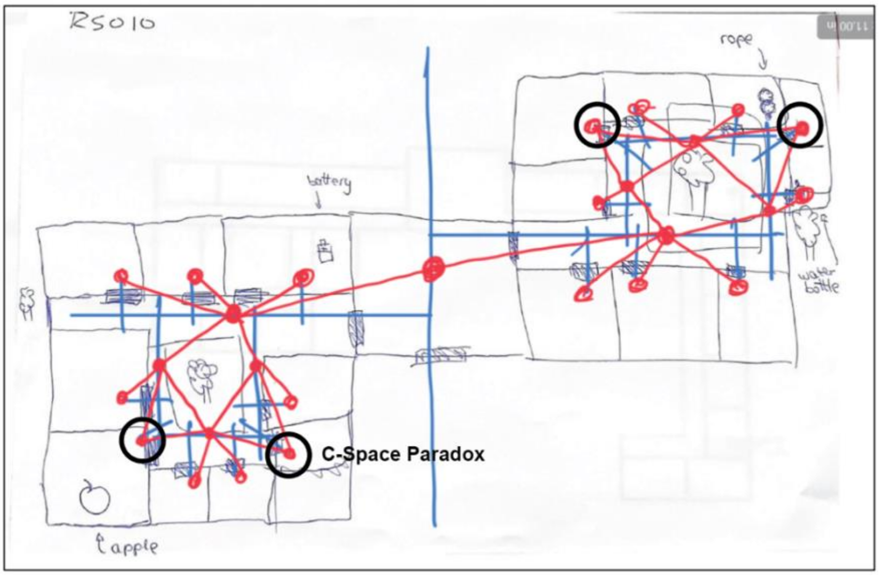


Figure 15: *The C-Space paradox.*

To assess graph abstraction consensus among those 7 participants that produced completely correct graphs, all unique pairwise comparisons of the hybrid axial graphs for each of the 9 sketch maps were computed through the global accuracy metrics developed in this dissertation (resulting in 21 unique pairwise comparisons per sketch map). **Table 3** reports the resulting average distances between all graphs for each sketch map. The results show that disagreement on how to abstract the graphs varies markedly between sketch maps.

Table 3: Average Graph Distances per Sketch Map (21 unique pairiwse comparisons each)

Sketch Map ID	Node Count Distance	Edge Count Distance	Total Depth Distance	Cluster Count Distance	Average Degree Distance	Cycle Count Distance	A-Space Share Distance	B-Space Share Distance	C-Space Share Distance	D-Space Share Distance	Isomorphism Share
RS010	0.571	0.571	289	0	0.002	0	0.017	0.013	0.016	0.012	0.143
AS003	1.238	1.571	193	0.571	0.031	0.333	0.027	0.059	0.053	0	0.048
MS001	4	4.238	708	0	0.034	0.333	0.042	0.042	0.044	0.017	0.048
AS008	2.619	1.714	403	0.619	0.039	0.476	0.032	0	0.053	0.036	0
MS003	1.238	1.524	254	0.571	0.02	0.286	0.027	0.048	0.023	0.01	0.048
MS007	5.81	6.048	1073	0.333	0.038	0.333	0.076	0.009	0.061	0.032	0.143
RS009	2.095	2.381	193	0.571	0.027	0.286	0.048	0	0.034	0.014	0.048
AS009	2.048	2.19	364	0.762	0.04	0.333	0.073	0.067	0.088	0.031	0

By normalizing the distance distribution of each sketch map metric between 0 and 1 accordingly:

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

and creating a compound z_i score for each sketch map by summing all its normalized metrics together, graph abstraction consensus can be explored further by relating it to the subjective survey ranks and scores of each sketch map. **Figure 16** depicts this relationship. The strong statistically significant correlations (at a confidence level of 95%) evidence that graph abstraction consensus is stronger for better ranked and scored sketch maps and thus influenced by perceived map quality.

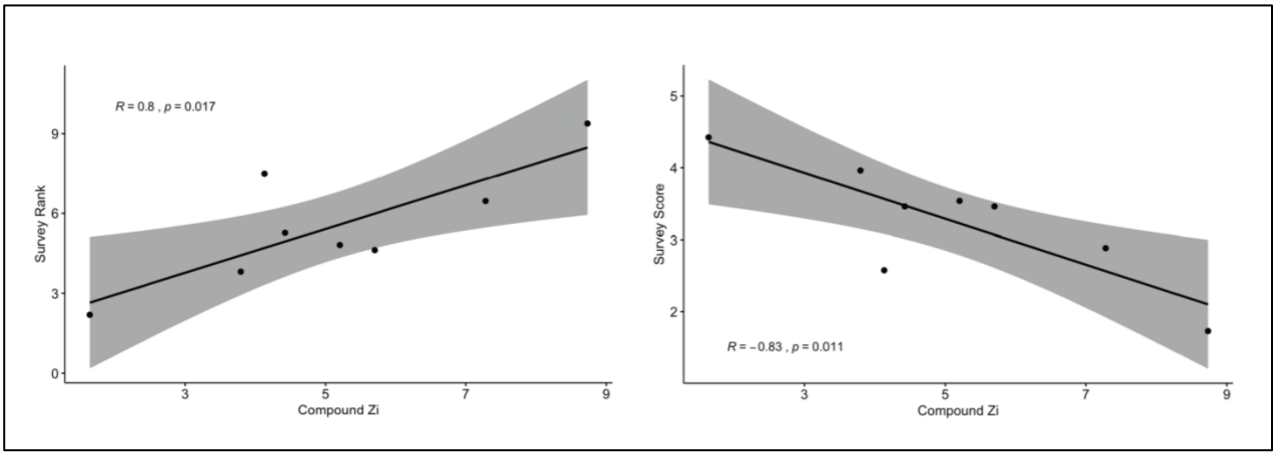


Figure 16: *The relationship between Compound Zi and Survey Rank/Score.*

4.4. Concluding Thoughts

The following findings have shown that, as posited in the introduction, there are two different types of sketch map quality distributions for simple and complex buildings: Binary and linearly graduated. These partially influence the types of consensus patterns observable across building scales. However, it is posited, that with further analysis, it could be proven that consensus behaves more consistently across building scales for the comparative gamma map method than the ranking and scoring approaches. In the ranking and scoring approaches "bad quality consensus" was observable for simple buildings and there seems to be "quality extremes consensus" for complex buildings. For the comparative gamma map method, however, consensus improved gradually by perceived map quality for complex buildings. The missing link in this argument is whether this is the same for simple buildings. From the author's previous experience with abstracting gamma maps from simple building sketch maps (Bruce 2022), it can be said, that particularly bad sketch maps were representationally ambiguous, whereas there was no ambiguity for the remaining maps. Accordingly, it is assumed, that there are not two different types of consensus patterns present across building scales for the comparative gamma map method, but rather two different types of manifestations of the same consensus, which reflect the underlying quality distributions of sketches for different building scales. If this is true, it suggests that, whilst there are reproducibility issues with the comparative gamma map method (especially for complex buildings), it behaves more consistently across building scales than ranking and scoring approaches.

5. Unsupervised Sketch Map Classification

In this chapter the application of an unsupervised classification algorithm on a sketch map data set is demonstrated (addressing **RQ5**). Using the sketch map measurements and metrics proposed in this dissertation, the sketch maps can be re-expressed as multidimensional, quantitative information processable by an algorithm. The aim is to explore whether some novel, meaningful classification outcomes can be achieved. The K-means algorithm is used, which is one of the simplest and most commonly used unsupervised learning algorithms. The goal of K-means clustering is to learn something about the data by grouping similar observations together into *k* pre-specified clusters. Accordingly, the user needs to define the target number of *k,* which is the number of cluster centroids needed. Data points are then allocated to the nearest cluster as such that the centroids are kept as small as possible. In other words, clusters are formed by allocating data points to them that minimize the in-cluster sum of squares (Hammerly and Elkan 2003). The resulting clusters are characterised by analysing the values of their centres as well as plotting them along the principal component analysis scales of the first two dimensions. Principal component analysis (PCA) is another unsupervised measurement method. The aim of PCA is to re-describe the data in a smaller number of uncorrelated variables. The principal components are the resulting uncorrelated variables that capture as much of the data's total variance as possible (Lauderdale 2021). To avoid the variance described in PCA being influenced predominantly by different units of measurement, the data is standardised.

5.1. Data

The data used in this chapter was collected by Professor Kate Jeffery at the Department of Behavioural Psychology, UCL, as part of her ongoing cognitive mapping research. Participants explored complex buildings virtually with headsets and afterwards sketched survey view maps of them. Sketch maps of three different complex building types, which distinguish themselves by their floorplan symmetry, are present in the sample (see **Figure 17**). The sample encompasses a total of 32 sketch maps, 11 of which represent the asymmetric building, 10 of which represent the mirror symmetric building and 11 of which represent the rotationally symmetric building.

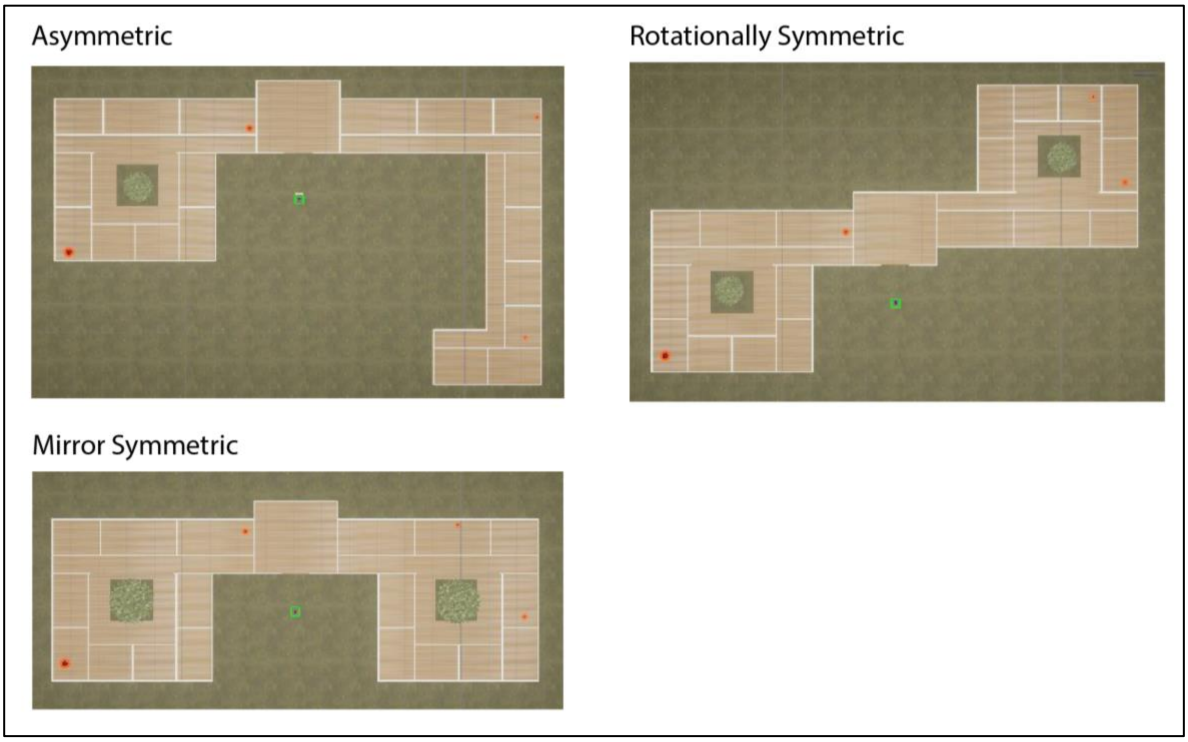


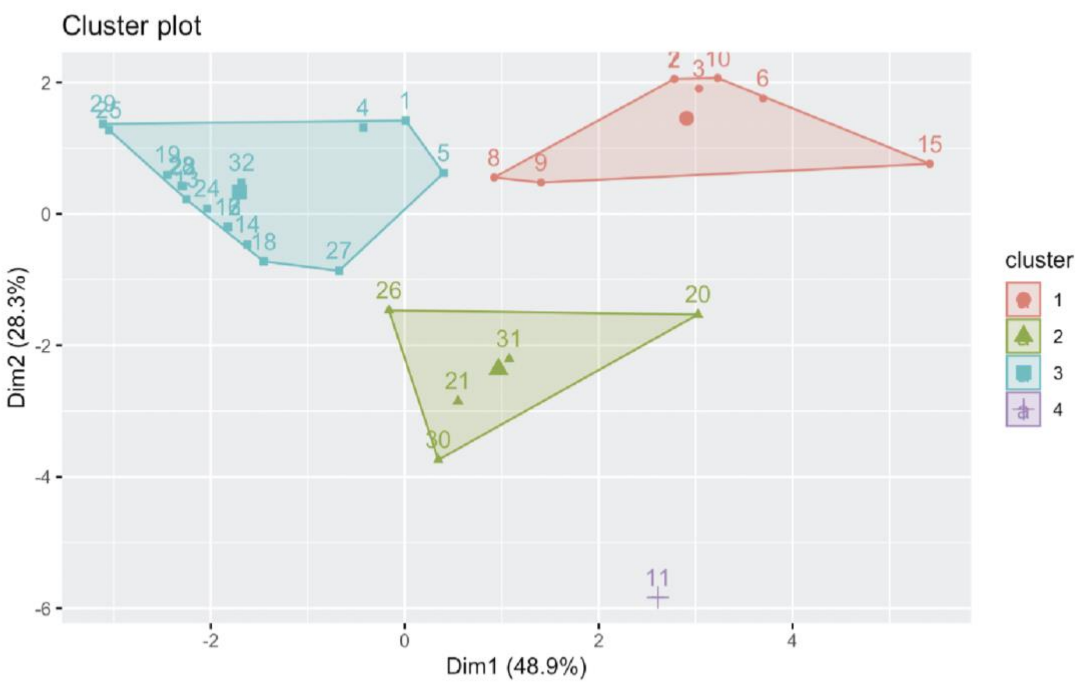
Figure 17: *The three different building types present in the data.*

5.2. Analysing K-means Clusters

The sketch maps are classified twice using the K-means clustering algorithm. The first classification uses the sketch graph measures and the second one the sketch graph accuracy metrics, which are derived by comparing the sketch graph measures to the baseline graph measures as proposed in **section 3.2.1**. Therefore, the first classification reflects sketch map properties, whereas the second one reflects sketch map accuracy.

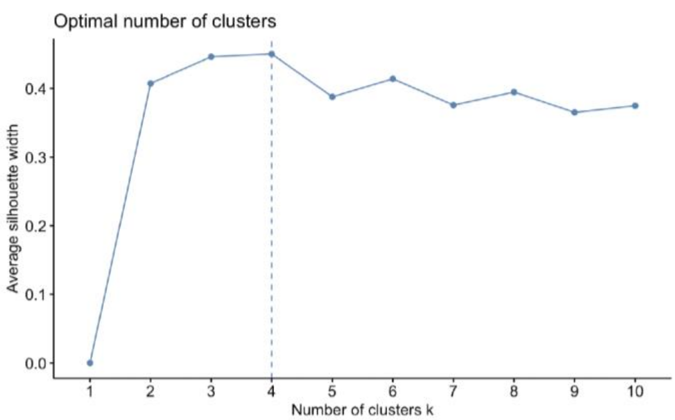
Figure 18 depicts the K-means classification of the sketch maps by sketch graph measures. A Screeplot is used to identify four as the optimal number of clusters to best explain the variance in the data. In the cluster plot the sketch maps are allocated into clusters and mapped out along the 1st and 2nd principal component dimensions, which combined explain about 80% of the total variance in the data. By analysing the cluster centres and the loadings of the two first principal components, the four clusters can be characterised. Accordingly, the sketch maps in cluster one can be classified as "rather node and edge rich sequential layouts", those in cluster two as "rather node and edge sparse ringy layouts", those in cluster three as "node and edge rich ringy layouts" and the map in cluster four as a "completely sequential layout". Evidently, the K-means clustering is classifying the sketch maps according to their system size and "ringiness".

Interestingly, the clusters do not really reflect the different types of buildings represented in the sample. This is visible in **Figure 19** depicting the building type share per cluster.



Cluster Centres: Sketch Graph Measures

Cluster	Node Count	Edge Count	Average Degree	Total Depth	Cycles	A	B	C	D	Cluster Count
1	21.6	20.8	1.92	1263	0.25	0.79	0.16	0.03	0	3.7
2	17.6	18.4	2.09	780	1.8	0.68	0	0.24	0.06	3.8
3	26.6	27.5	2.06	2226	1.83	0.72	0.02	0.23	0.02	4.8
4	15	15	2	840	1	0	0	1	0	4



Principal Component Loadings

	PC1	PC2
Node Count	-0.325	0.34
Edge Count	-0.364	0.276
Average Degree	-0.343	-0.29
Total Depth	-0.343	0.26
Cycle Count	-0.374	-0.236
A-Space Share	0.006	0.442
B-Space Share	0.312	0.334
C-Space Share	-0.114	-0.439
D-Space Share	-0.085	-0.275
Cluster Count	-0.357	0.037

Figure 18: *K-means classification of sketch maps by sketch graph measures.*

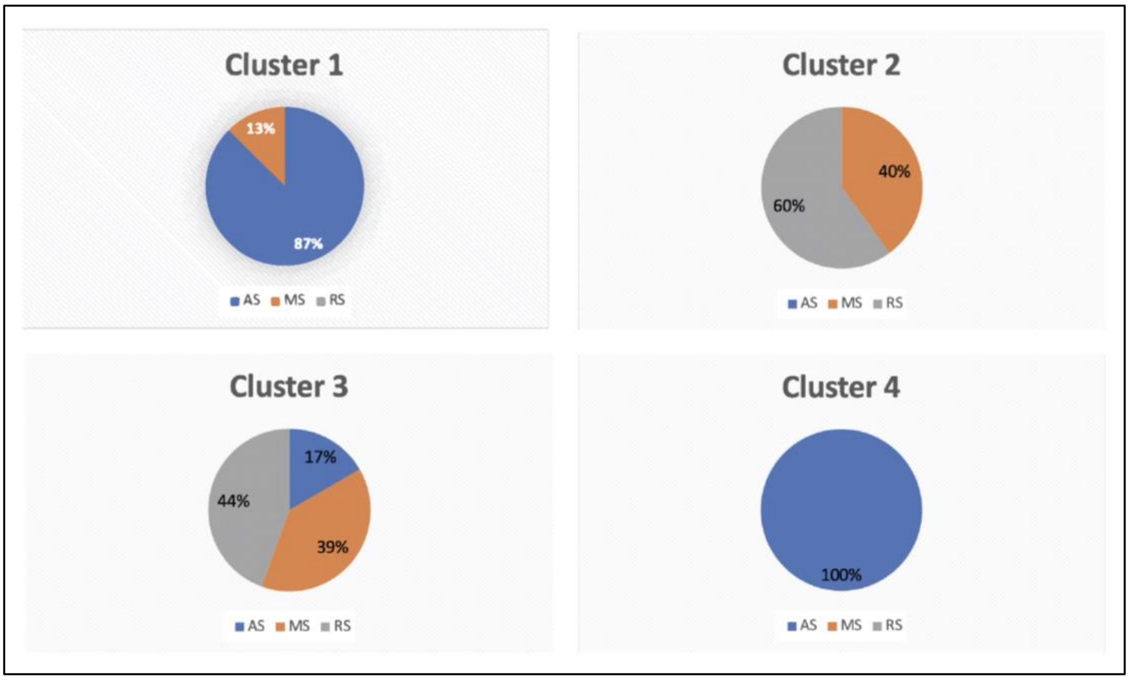
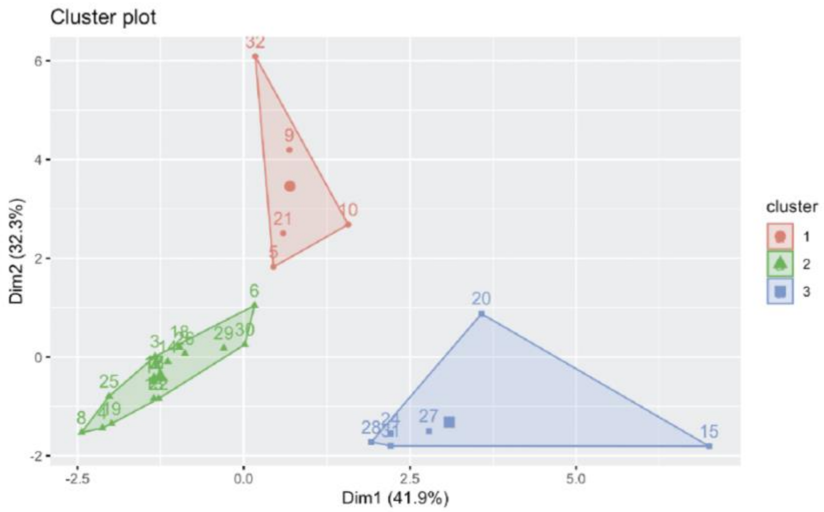


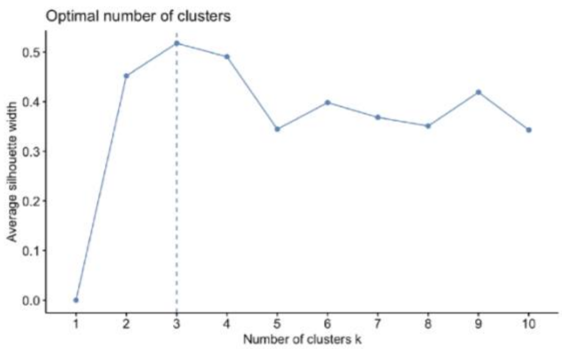
Figure 19: *Building type share in each measurement cluster.*

Figure 20 depicts the K-means classification of the sketch maps by the sketch graph accuracy metrics. Via the Screeplot, three clusters were identified as optimal. In the cluster plot the sketch maps are allocated into clusters and mapped out along the 1st and 2nd principal component dimensions, which combined explain about 73% of the total variance in the data. Again, by analysing the cluster centres and the loadings of the two first principal components, the three clusters can be characterised. Accordingly, the sketch maps in cluster one can be classified as “edge and node count inaccurate”, those in cluster two as “least inaccurate” and those in cluster three as “ringiness inaccurate”. Evidently, K-means clustering is classifying the maps into size and “ringiness” inaccurate ones, and those that are rather topologically accurate. Again, the clusters do not reflect the different types of buildings in the sample, as is shown in **figure 21** depicting the building type share per cluster. Lastly, **Figure 22** shows for those sketch maps that were subjectively ranked and scored in the survey (see **chapter 4**) in which accuracy metric cluster they are located. Whilst there is only ranking and scoring data for a subset of all the maps classified by the K-means algorithm, this still shows quite nicely that there is no correspondence between perceived map quality and the K-means classification. This could suggest that the K-means classification is not confounded by aesthetic appearances of the sketch maps and only classifies them by topo-configurational properties. Alternatively, this could be interpreted as the classification not resembling a real concept of map quality.



Cluster Centres: Accuracy Metrics

Cluster	Node Count Variance	Edge Count Variance	Total Depth Variance	Cluster Count Variance	Average Degree Variance	Cycle Count Variance	A-Space Variance	B-Space Variance	C-Space Variance	D-Space Variance
1	-9	-9	-1510	-1	0.03	0	-0.15	-0.03	0.13	0.05
2	0.1	0.1	-37	-0.2	0	0	0.01	0	-0.01	0
3	-4.7	-5.8	-678	-1.42	-0.1	-1.1	0.05	0.1	-0.14	-0.01



Principal Component Loadings

	PC1	PC2
Node Count Variance	-0.352	-0.369
Edge Count Variance	-0.384	-0.328
Total Depth Variance	-0.308	-0.35
Cluster Count Variance	-0.404	-0.062
Average Degree Variance	-0.372	0.298
Cycle Count Variance	-0.433	0.216
A-Space Share Variance	0.014	-0.399
B-Space Share Variance	0.349	-0.322
C-Space Share Variance	-0.119	0.398
D-Space Share Variance	-0.071	0.263

Figure 20: *Building type share in each metric cluster.*

Figure 22: *Building type share in each accuracy metric cluster.*

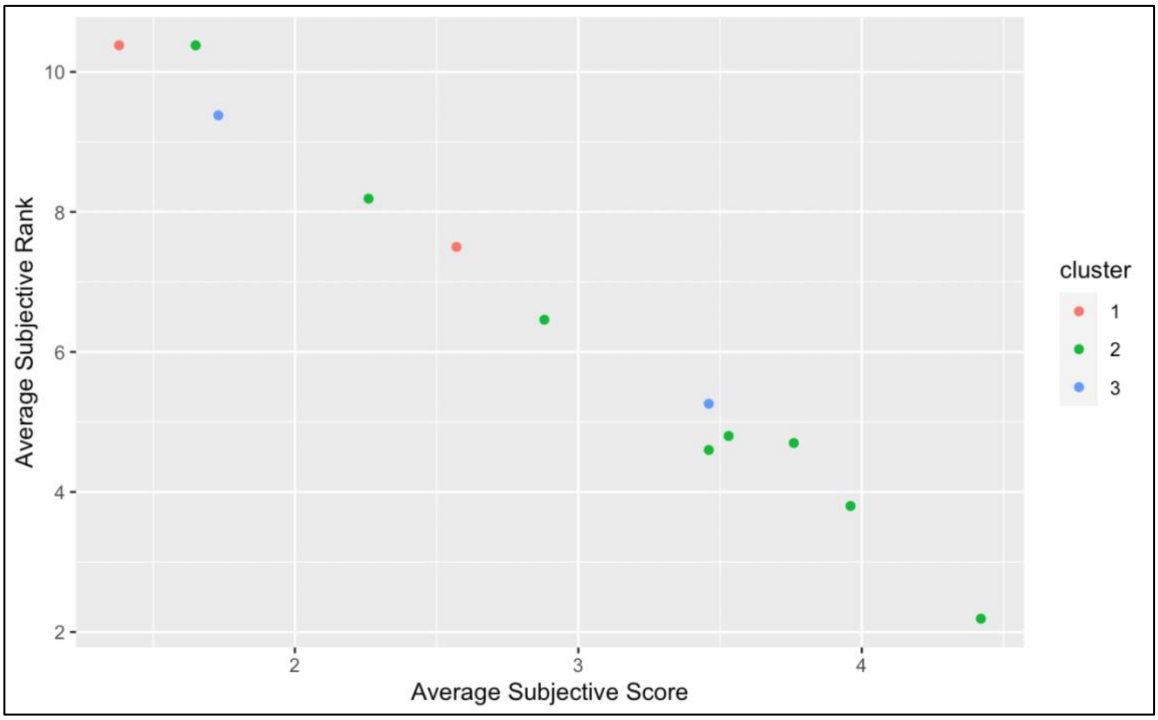


Figure 21: *Average Subjective Rank and Score by accuracy metric cluster.*

6. Discussion

The methods introduced in this dissertation by no means figure as the definitive answer to quantifying survey view building sketch maps. However, they provide an approach that is both adaptable to the specific requirements of spatial cognition research as well as coherent enough to produce a landscape of comparable literature, thus addressing Gardony, Brunyé and Taylor's (2015) concern of congruency within the field. Importantly, **chapter 4** provides a tangible process through which to test and improve the reproducibility of the comparative-gamma map method. It is apparent that in a next step sketch map modelling and data collection rules need developing. However, this chapter will take a step back from the obvious considerations on methodological testing and refinement to discuss the bigger picture, that is, what this method could mean for the field of spatial cognition research and how to deliver on that potential.

Sketch map coding and spatial cognition theory have been treated as two separate things. Sketch map coding is considered a function that returns some quantifiable sketch map data from sketch maps and spatial cognition theory as providing the framework through which to analyse the sketch map data. Accordingly, sketch map coding approaches have made no attempt to emulate the theories on the structure of spatial knowledge in the mind (e.g. the cognitive map or cognitive graph theory). Consequently, experiments have been confined to collecting behavioural and explicit evidence that displays properties in favour of one or the other theory, rather than testing the theories themselves.

This is where the comparative-gamma map method is different to past coding approaches. It is a coding and analysis approach which almost embodies the cognitive graph theory. The hybrid-axial graphs abstracted from the sketch maps can be considered simple cognitive graphs minus rough distal and angular relationships (which are suggested as further properties of cognitive graphs by Ericson and Warren (2020)). From this perspective, the adaptation of the original comparative gamma map approach to complex buildings also figures as a step towards abstracting and analysing cognitive graphs from survey view sketch maps, as it results in a method more sensitive to the topo-configurational structure of buildings. The question then is whether the comparative gamma map method can be progressed towards fully emulating cognitive graphs and what this emulation means for spatial cognition research.

To answer this, Hillier and Iida's (2005) paper 'Network effects and psychological effects: a theory of urban movement' needs to be considered. At the urban scale of space syntax research, pedestrian movement flows have been reliably predicted from the graph properties of street networks by considering topological and angular distance between street segments. The street network graph underlying these computations is essentially a cognitive graph. Hillier and Iida's (2005) work figures as a first attempt in space syntax to establish a connection between how the street network is modelled in graph form and the kind of individual cognitive assumptions it embodies. This permitted them to make the leap from predicting aggregate pedestrian movement flows, to forming assumptions about individual cognitive decisions. They considered different ways of representing network distances in urban systems as abstract, but testable, cognitive theories, and accordingly modelled aggregate pedestrian movement flows through network distances, such as metric distance, topological distance, and angular distance. As they were able to best explain the pedestrian movement flows using a combination of topological and angular network distance, Hillier and Iida (2005) could postulate assumptions about the individual cognitive decisions behind the emergent pedestrian movement flows. This is encouraging, as it suggests two things. Firstly, the space syntax approach to coding space has some cognitive purchase. Secondly, that sketch map coding can and should be treated as a testable emulation of a cognitive theory.

However, as touched upon above, the current comparative gamma map method does not strictly emulate the entire cognitive graph theory. Angular relationships between graph nodes are missing. The real challenge lies in ascertaining whether the angular relationships between features in the sketch map mean anything. Sketch maps are often distorted representations of the reality, which means angular relationships are distorted too (Golledge and Stimson 1997). One solution is to increase the tolerance between comparisons of angular relationships in sketch maps. This is possible by weighting the edges in the hybrid-axial graph by canonical relationships such as up, down, left and right. Methodological considerations aside, what kind of research could be conducted to test the cognitive graph theory through the comparative gamma map method? A proposal is forwarded in the following paragraph.

The comparative gamma map method has the potential to test how the cognitive graph behaves, that is, how the mind may use and adapt it. Following Hillier and Iida's (2005) logic of deriving insights on individual cognitive mechanisms from aggregate observational data, a research design is suggested. Buildings with the same salient topological structures, e.g. a fixed number of cycles and sequential thoroughfares, could be scaled up by a factor of rooms, or by the area oft the building footprint. At each scale of the building, the comparative gamma map method could be used to assess whether sketch map errors have changed the salient topological structures or not. If, for example, at one scale some spaces are missing but the number of cycles and sequential throughfares is preserved, this could indicate that humans can conceptually understand the salient topological structures of the building as such that a cognitive graph of these structures, not the building size, is reproduced. In other words, the comparative gamma map method could be used to asses what structures of the cognitive graph for buildings are reproduced/maintained for different transformations of an identical building. This may provide evidence on how the mind uses the cognitive graph.

7. Conclusion

Reliably processing sketch maps in cognitive mapping research presents a notorious challenge (Kitchin 2000). This dissertation has further developed Bruce's (2022) initial effort to overcome this challenge through the invention of the comparative gamma map method, which figures as a topo-configurational sketch map coding and analysis approach.

Firstly, the comparative gamma map method was adapted to the analysis of complex building systems. Two fundamental shifts - differentiating survey view sketch maps of simple buildings from complex buildings - are identified which necessitate this adaptation: An increase in representational ambiguity stemming from the heightened graphical confoundedness of the sketching task and a decrease in labelled sketch map features. Accordingly, the *hybrid-axial graph* was introduced as a graph that can both capture the topo-configurational structures of complex buildings as well as meet requirements of spatial cognition research. Furthermore, novel sketch map accuracy metrics were developed. These consist of global metrics summarising the overall topo-configurational similarity between baseline and sketch graph, as well as a local metric quantifying the local topological neighbourhood similarity of labelled features in the sketch and baseline graph. For the local metric, Milenković and Pržulj's (2008) *signature similarity between nodes metric* was applied and tested on a hybrid-axial graph of a complex building. This resulted in the discovery of a conceptual extension of Hillier's (2007) topological ABCD-space classification which differentiates more sensitively between distributed and non-distributed building spaces.

Secondly, the methodological reproducibility of the comparative gamma map method was tested and compared to conventional ranking and scoring approaches. Some disagreement in abstracting hybrid axial graphs from sketch maps of complex buildings is evident and consensus improves by perceived "goodness" of map quality. Importantly, initial findings suggest that sketch map quality consensus behaves more consistently across building scales for the comparative gamma map method than conventional ranking and scoring approaches, in which two different consensus patterns are identified for simple and complex buildings ("bad quality consensus" and, presumably, "quality extremes consensus").

Finally, the possibility of classifying the sketch maps according to their sketch graph measures and accuracy metrics was explored. A sample of complex building sketch maps was classified using the K-means clustering algorithm. For the sketch map measures, clusters distinguishing by "ringiness" and size of the building systems were identified. Similarly, for the accuracy metrics, clusters distinguishing by "ringiness" accuracy and building size accuracy were identified.

This three-pronged approach sees the comparative gamma map method mature into a more building-scale-robust sketch map coding and analysis approach that is both adaptable to the specific requirements of spatial cognition research as well as coherent enough to produce a landscape of comparable literature. Importantly, the benchmarking of the method provides a tangible process through which to improve the reliability of the method moving forward. In a next step, sketch map modelling and collection rules need devising. However, considering the bigger picture, the comparative gamma map method presents an exciting prospect to overcome the longstanding divide between theory and coding approach in spatial cognition research. This promises to open the field to new research in which theory is testable at the level of the theoretical model, delivering new insights that may not just validate cognitive mapping theories themselves, but also explain specific aspects of how they work.

8. References

Bicanski, A. and Burgess, N. (2020) 'Neuronal vector coding in spatial cognition', *Nature,* 21, 453-470

Billinghurst, M. and Weghorst, S. (1995) 'The use of sketch maps to measure cognitive maps of virtual environments', in *Proceedings Virtual Reality Annual International Symposium*', 95, 1995 IEEE, 40-47.

Blades, M. (1990) 'The Reliability of Data Collected From Sketch Maps', *Journal of Environmental Psychology*, l10, 327-339

Boehm, A.E. and Weinberg, R.A. (1997) *The classroom observer: Developing observation skills in early childhood settings* (3rd ed.)*,* New York, NY: Teachers College Press

Boyle, M.J. and Robinson, M.E: (1979) 'Cognitive mapping and understanding', in D.T. Herbert and R.J. Johnston (Eds.), *Geography and the Urban Environment: Progress in Research and Applications* (59-82)*,* London: Wiley

Bruce, T. (2022: Unpublished) 'Space Syntax and Spatial Cognition: Comparing and extracting measures of spatial cognition from survey view sketch maps', a research project for the BARC0026 module of the MSc Space Syntax: Architecture and Cities master at the Bartlett School of Architecture, UCL (**see appendix A for full paper**)

Brunyé, T. T. and Taylor, H. A. (2008a). 'Extended experience benefits spatial mental model development with route but not survey descriptions', *Acta Psychologica*, 127, 340-354

Brunyé, T. T. and Taylor, H. A. (2008b). 'Working memory in developing and applying mental models from spatial descriptions', *Journal of Memory and Language*, 58, 701-729

Carassa, A., Geminiani, G., Morganti, F. and Varotto, D. (2002) 'Active and passive spatial learning in a complex virtual environment: The effect of efficient exploration', *Cognitive Processing*, 3, 65-81

Chrastil, E.R. and Warren, W.H. (2014) 'From cognitive maps to cognitive graphs', *PLoS One,* 9, 11

Clark, J. and Holton, D.A. (1991) *A first look at Graph Theory,* Singapore: World Scientific Publishing

Coluccia, E., Bosco, A. and Brandimonte, M. A. (2007) ‘The role of visuo-spatial working memory in map learning: new findings from a map drawing paradigm’, *Psychological Research*, 71, 359-372

Dalton, R. (2005), 'Space syntax and spatial cognition', *World Architecture: Space Syntax Monograph*, 11, 185, 41 - 45.

Dalton, R. C. and Kirsan, C. (2008) ‘Small-Graph Matching and Building Genotypes’, *Environment and planning. B, Planning & design*, 35, 5, 810-830

Dalton, R. C., Hölscher, C. and Alasdair, T. (2012) ‘Understanding Space: The Nascent Synthesis of Cognition and the Syntax of Spatial Morphologies’, *Environment and planning. B, Planning &* design, 39, 1, 7–11.

Day, R.A. (1976) ‘Urban distance cognition: review and contribution’, *Australian Geographer,* 13, 193-200

Downs, R. M. and Stea, D. (1973) *Image and environment: Cognitive mapping and spatial behavior*, Chicago: Aldine Pub. Co.

Downs, R.M. (1985) ‘The presentation of space: Its development in children and in cartography’, in R. Cohen (ed.), *The Development of Spatial Cognition* (323-345), Hillsdale, NJ: Erlbaum Lawrence

Ericson, J.D. and Warren, W.H. (2020) ‘Probing the invariant structure of spatial knowledge: Support for the cognitive graph hypothesis’, *Cognition,* 200, 1-20

Gallistel, C.R. (1990) *The organization of learning,* Cambridge: MIT Press

Gardony, A.L., Taylor, H.A. and T.T. Brunyé (2015) ‘Gardony Map Drawing Analyzer: Software for quantitative analysis of sketch maps’, *Behav Res,* 48, 151-177

Golledge, R. G. and Stimson, R. J. (1997) ‘Spatial Cognition, Cognitive Mapping, and Cognitive Maps’ in *Spatial Behavior A Geographic Perspective*, 224-266, New York: The Guilford Press

Golledge, R.G. (1987) ‘Environmental cognition’, In D. Stokols and I. Altman (eds) *Handbookof Environmental Psychology*, 1, New York: Wiley

Golledge, R.G., Smith,T.R., Pellegrino, J.W., Doherty, S. and Marshall, S.P. (1985) 'A conceptual model and empirical analysisof children'sacquisition of spatial knowledge', *Journal of Environmental Psychology*, 5, 125-152

Hammerly, G. and Elkan, C. (2003) 'Learning the k in k-means', in *Advances in neural information processing systems*

Haq, S. and Girotto, S. (2003) 'Ability and intelligibility: Wayfinding and environmental cognition in the designed environment', in *Proceedings of the 4th International Symposium on Space Syntax*, 68.1-68.20

Hillier (2007) *Space is the Machine,* London: Space Syntax

Hillier, B. (1989) 'The architecture of the urban object', *Ekistics*, 56 (334/33), 5-21

Hillier, B. and Hanson, J. (1984) *The Social Logic of Space*. Cambridge: Cambridge University Press

Hillier, B. and Penn, A. (1991) 'Visible Colleges: Structure and randomness in the place of discovery', *Science in Context*, 4, 1, 23-50

Hočevar, T. and Demšar, J. (2016) 'Computation of Graphlet Orbits for Nodes and Edges in Sparse Graphs', *Journal of Statistical Software*, 71, 10, 1–24

Jacobs, L.F. and Schenk, F. (2003) 'Unpacking the cognitive map: the parallel map theory of hippocampal function', *Psychological Review*, 110, 285-315

Jeffery, K., Guo, W., Ball, D. and Rodriguez-Sanchez, J. (2021) 'Visual imagination and cognitive mapping of a virtual building', *The Journal of Navigation*, 1–14

Kim, Y. O. (2001) 'The role of Spatial Configuration in Spatial Cognition', in J. Peponis, J. Wineman and S. Bafna (eds.), *Proceedings of the 3rd International Symposium on Space Syntax*, 49.41- 49.21

Kim, Y.O., Penn, A. (2004) 'Linking the spatial syntax of cognitive maps to the spatial syntax of the environment', *Environment and Behavior*, 36, 483-504

Kitchin, R. (2000) 'Collecting and analysing cognitive mapping data', in R. Kitchen and S. Freundschuh (Eds.), *Cognitive mapping: Past, present and future* (9-23), London UK: Routledge

Kuipers, B. (1983) ‘The Cognitive Map: Could It Have Been Any Other Way?’, In Pick, H.L., Acredolo, L.P. (eds) *Spatial Orientation*, Boston: Springer

Kuipers, B., Teccuci, D.G., and Stankiewicz, B.J. (2003) ‘The skeleton in the cognitive map: A computational and empirical exploration. *Environment and Behaviour,* 35, 81-106

Lauderdale, E.B. (2021) *Pragmatic Social Measurement, https://uclspp.github.io/POLS0013/readings/pragmatic-social-measurement.pdf*

Lynch, K. (1960) *The Image of the City*, Cambridge, Massachusetts: MIT Press

Milenković, T. and Pržulj, N. (2008) ‘Uncovering Biological Network Function via Graphlet Degree Signatures’, *Cancer Informatics*

Moeser, S. D. (1988) ‘Cognitive Mapping in a Complex Building’, *Environment and behavior*, 20, 1, 21–49

Montello, D. R. and Raubal, M. (2012) ‘Functions and applications of spatial cognition’ in D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (249–264), Washington, DC: American Psychological Association.

Montello, D.R. (2016) ‘Behavioural methods for spatial cognition research’, in R. Gifford (ed.), *Research Methods for Environmental Psychology* (161-182), West Sussex: John Wiles & Sons, Ltd

NetworkX (2022) https://networkx.org (accessed: 13th of May 2022)

Newcombe, N. (1985) ‘Methods for the Study of Spatial Cognition’, In R. Cohen (ed) *The Developmentof Spatial Cognition*, Hillsdale, NJ: Erlbaum

O’Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*, Clarendon Press

Peer, M., Brunec, I.K., Newcombe, N.S. and Epstein R.A. (2021) ‘Structuring Knowledge with Cognitive Maps and Cognitive Graphs’, *Trends in Cognitive Sciences,* 25, 1, 37-54

Peponis, J., Zimring, C. and Choi, Y. K. (1990) ‘Finding the building in wayfinding’, *Environment and Behavior,* 22, 5, 555-590

Pržulj, N., Wigle, D. , and Jurisica, I. (2004) ‘Functional topology in a network of protein interactions’, *Bioinformatics*, 20, 3, 340-348

Saarinen, T.F. (1988) ‘Centring of mental maps of the world’, *National Geographic Research,* 4, 112-127

Shervashidze, N., Schweitzer, P., Jan van Leeuven, E., Mehlhorn, K. and Bogwardt, K.M. (2011) ‘Weisfeiler-Lehman Graph Kernels’, *Journal of Machine Learning Research,* 12, 2539-2561

Siegel, A.W. and Cousins, J. H. (1985) ‘The symbolizingand symbolized child in the enterprise of cognitive mapping’, In R. Cohen (ed) *The Developmentof Spatial Cognition*, Hillsdale, NJ: Erlbaum

Tolman, E. C. (1948) ‘Cognitive maps in rats and men’, *Psychological Review*, **55**(4), 189–208

Traag, V.A., Waltman, L. and van Eck, N.J. (2019) ‘From Louvain to Leiden: guaranteeing well-connected communities’, *Sci Rep,* 9, 5233, 1-19

Wood, D, and Beck, R. (1976) ‘Talking with environmental A: An experimental mapping language’, in G.T. Moore and R.G. Golledge (Eds.), *Environmental Knowing* (351-361)*,* Stroudsberg, PA: Dowden, Hutchinson and Ross

Woollett, K. and Maguire, E. A. (2010). ‘The effect of navigational expertise on wayfinding in new environments’, *Journal of Environmental Psychology*, 30, 565-573

Zanbaka, C. A., Lok, B. C., Babu, S. V., Ulinski, A. C. and Hodges, L. F. (2005) ‘Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment’, *IEEE Transactions on Visualization and Computer Graphics*, 11, 694-705

Zhang, M. and Chen, Y. (2017) ‘Weisfeiler-Lehman Neural Machine for Link Prediction’, *in Proceedings of KDD,* 13-17

9. Appendix

A. Bruce, T. (2022: Unpublished) 'Space Syntax and Spatial Cognition: Comparing and extracting measures of spatial cognition from survey view sketch maps', a research project for the BARC0026 module of the MSc Space Syntax: Architecture and Cities master at the Bartlett School of Architecture, UCL

Space Syntax and Spatial Cognition: Comparing and extracting measures of spatial cognition from survey view building sketch maps.

An ADRP Project By Timothy Lawrence Bruce
Supervised by Professor Alan Penn

MSc Space Syntax: Architecture and Cities
Bartlett School of Architecture, UCL
May 2022

Abstract

Cognitive maps are mental models used by humans and animals alike to navigate the world. The process of developing such an internal representation is referred to as cognitive mapping. Scholars of the built environment sciences have studied cognitive mapping since the 1960s, as it has implications for designing more intelligible environments. Sketch mapping is the primary technique used to collect data on a participant's environmental knowledge and is applied widely in cognitive mapping research. However, the information contained in sketch maps is often distorted and incomplete as well as confounded by the participant's graphical skills. This presents a challenge to cognitive scientists seeking to score the accuracy of these maps. Up until now, qualitative approaches have been taken. However, it is difficult to maintain objectivity and strict reproducibility with such an approach. Accordingly, this study forwards a topological graph matching method inspired by Hillier and Hanson's (1984) gamma map method to produce seven accuracy measures for survey view building sketch maps. The method's efficacy is tested on a sample of 156 sketch maps collected from Jeffery *et al.'s* (2021) cognitive mapping research. The analysis with this novel method suggests that the accuracy measures are highly applicable to cognitive mapping research, by not only mirroring the findings of Jeffrey *et al.'s* (2021) study but also uncovering novel spatial retention trends.

Word Count: 4'438

Table of Contents

Acknowledgments

I would like to extend my gratitude to Professor Alan Penn for supervising my project. Without his insights, encouragement, and willingness to connect me with other people in the architecture and psychology department at UCL this research would not have been possible. I would also like to thank Professor Kate Jeffery from the UCL Department of Experimental Psychology for advising me and sharing the sketch map data from her experiments.

1. Introduction

To navigate the real world, humans develop mental, or cognitive, maps that serve as an internal representation of an experienced or sensed external environment (Tolman 1948). Cognitive mapping describes the process in which such internalised representations are formed by placing environment sub-spaces in configurationally correct relationships (Downs and Stea 1973). When designing the built environment, whether at the city, neighbourhood, or building-scale, understanding how people construct cognitive maps is important, as it presents opportunities to design environments that are more intelligible and that structure their functionality according to that purpose (Gärling and Golledge 1989).

Spatial cognition research on cognitive mapping often uses sketch maps to elucidate information about a subject's environmental knowledge (Golledge and Stimson 1997). Compared to conventional maps, sketch maps contain no reliable metric information or sense of direction. Instead, information is often incomplete, distorted, and reliant on the individual's graphical skills and ability to transfer knowledge from an internal representation to a two-dimensional sketch (Golledge and Stimson 1997). This presents a challenge to spatial cognition researchers seeking to evaluate sketch map accuracy. Up until now, especially qualitative approaches, that distinguish sketches by their conceived completeness, coherency or similarity, have been taken (see Jeffery *et al.* 2021, Moeser 1988). These approaches risk confounding findings by the subject's ability to undertake the transformation process from internal representation to two-dimensional sketch map successfully. Furthermore, the most reliable information underlying sketch maps, such as topology and configuration, is disregarded (Golledge and Stimson 1997). Accordingly, there is a need to develop alternative methods that harness the topology and configuration of sketch maps to produce multidimensional quantitative accuracy scores for cognitive mapping.

This paper proposes a new method to score survey view building sketch maps. Underlying the challenge of scoring sketch map accuracy lies a historical focus of spatial cognition research on the subject's cognitive processes rather than the description of the research environment (Peponis, Zimring and Kyung 1990). There is a need to incorporate analytical descriptions of space in spatial cognition research, and space syntax provides a toolset for spatial description that will enable a greater precision in defining the research environment (Dalton, Hölscher and Turner 2012). Hillier and Hanson's (1984) gamma map, a graph representing the interior layout of a building according to permeability relationships between spaces, provides a fundamental building block for the spatial description of buildings. Accordingly, this research is guided by one main research question:

> How can Hillier and Hanson's (1984) gamma map method be adapted to extract meaningful measures of spatial cognition from survey view building sketch maps?

Chapter 2 reviews the literature on cognitive mapping, sketch maps and related space syntax research. In **Chapter 3** the proposed method to score sketch map accuracy is presented. This is followed by a chapter on the trial data provided by Jeffery *et al.'s* (2021) cognitive mapping research (**Chapter 4**) and an analysis chapter (**Chapter 5**) in which the proposed method is tested. Finally, the method is discussed in **Chapter 6** and conclusions are drawn in **Chapter 7**.

2. Literature Review

2.1 Cognitive Maps and Cognitive Mapping

Tolman (1948) first conceived of the concept of the cognitive map by suggesting that animals form mental representations of their environments that condescend simple sequences of associations. The idea that animals and humans alike store spatial information in their mind as a model-like entity is widely accepted now, however, how these internal representations are structured remains contested (Peer *et al.* 2021). Cognitive mapping is the process of forming a cognitive map. Downs and Stea (1979:7) define it as "[…] a process composed of a series of psychological transformations by which an individual acquires, stores, recalls and decodes information about the relative locations and attributes of the phenomena in his everyday spatial environment". Cognitive mapping is widely considered as a subordinate process of spatial cognition, which is defined by Hart and Moore (1973:248) as "the knowledge and internal or cognitive representation of the structure, entities and reflections of space".

Cognitive mapping research stretches back to the 1960s, when behavioural geographers and built environment researchers first started working with the concept of the mental map as spatial information represented in the mind in some maplike form (Golledge and Stimson 1997). Lynch's (1960) pioneering work on understanding how the city is represented in the mind through the analysis of sketch maps is one of the earliest and most prominent examples.

Cognitive maps can be made up of different types of information (Peer *et al.* 2021). Golledge purposefully distinguishes between the cognitive map, as an internal representation, and "cognitive configuration" as the information gleaned from them (see Golledge and Stimson 1997). Kuipers (1983) groups the information contained in cognitive maps into five categories: Topological, route descriptions, fixed features, metric and sensory images. Depending on the research purpose different types of information are of interest, for which different techniques need to be used to retrieve them.

2.2 Externalising Cognitive Map Information Through Sketch Maps

In cognitive mapping research sketch mapping is among the most proliferated techniques with which information on a subject's environmental knowledge is collected (see Golledge *et al.* 1985, Moeser 1988, Jeffery *et al.* 2021). However, this practice has been widely criticised (see

Golledge 1987 or Siegel and Cousins 1985). It is posited that the environmental knowledge gleaned from sketch maps is confounded with the subject's graphical skills and that sketch maps require the subject to undertake an unnatural transformation from their egocentric view of the world to an allocentric two-dimensional view of the world. The difficulty of interpreting and quantifying sketch maps has also been highlighted. Nevertheless, Blades (1990) found sketch maps to be reliable sources of environmental information and Newcombe (1985) assessed them to be no less accurate than other techniques in spatial cognition research.

The analysis of sketch maps can broadly be categorised into three distinct approaches (Kim 2001). The first is to understand how the perception of local configurational elements varies between environments. Such as Sadalla and Montello's (1988) research on the numbers of turns in a path and perceived distance, or Evans *et al.'s* (1980) research on sketch map distortions. The second it to establish correlations between exogenous factors, such as the participant's socio-economic background, and the content presented in the sketch maps (Appleyard 1970). Lastly, the frequency of certain sketch map features has been analysed (Lynch 1960, Haq 2003).

It is widely accepted that topology, configuration, and feature occurrence and frequency are the most reliable types of information that can be gleaned from sketch maps on a subject's environmental knowledge (Golledge and Stimson 1997). Lynch (1960), for instance, found sketch maps to be particularly useful for measuring a subject's topological knowledge. Nevertheless, most research working with sketch maps has relied on categorically distinguishing them based on qualitative criteria or feature counts (see Lynch 1960, Moeser 1988, Jeffery *et al.* 2021). There have been attempts to incorporate topology more prominently in the evaluation procedure, such as in Billinghurst and Weghorst's (1995) research, however these remain unconvincing, lacking the necessary tools to describe more complex spatial configurations. This disregard of topological and configurational information underlying the sketch maps forms part of Peponis, Zimring and Kyung's (1990) wider critique of spatial cognition and navigation research. They assert that an historical focus on psychological processes has led to a lack of methods with which to reliably describe the environment. Accordingly, as space syntax research is transitioning out of its infancy, a flurry of interdisciplinary research has been undertaken seeking to introduce the space syntax toolkit of environmental description to spatial cognition research (Dalton, Hölscher and Turner 2012).

2.3 Space Syntax and Sketch Maps

The syntactical properties of sketch maps have been studied within the interdisciplinary field of space syntax and spatial cognition. Much of this work stems from a growing consensus that wayfinding performance can be reliably predicted from topological variables of the environment, which can also be found in sketch maps (Haq and Girotto 2003). Kim (2001) investigated the role of intelligibility in mediating configurational differences between the reality and cognitive representations. This revealed that better sketch maps were produced for inhabitants living in configurationally more intelligible areas. Similarly, Kim and Penn (2004) linked the spatial syntax of sketch maps to that of the environment through axial line analysis. At the building scale, Haq and Girotto (2003) analysed sketch maps of hospitals for their intelligibility and local topological correctness.

In these examples adapting space syntax methodology to the analysis of sketch maps has been pioneered. However, the focus has mainly been on large scale environments, and accordingly axial analysis was deployed as the main method. In instances in which buildings were examined, axial line analysis also figured as the main method. To the author's awareness there have been no attempts to analyse sketch maps based on Hillier and Hanson's (1984) gamma map method. As the axial line models, compared to the gamma map, require drawing the longest and fewest lines of sight through a plan, which can become an inherently subjective process (Ratti 2004), and don't permit distinction between topological entities, there is a research gap in developing a suitable building scale sketch map scoring method based on the gamma map.

3. Methodology

3.1 Hillier and Hanson’s Gamma Map

Hillier and Hanson (1984) invented the gamma map to represent and research the configurational properties of a building’s interior spaces. It is a topological graph abstracted from buildings by representing each distinct space as a node and the permeability relationships between them as edges. When representing topological graphs, the node location is usually of no informational relevance, which is why there are many stylistic options. Force-directed layouts are particularly popular, as they can effectively visualise communities of closely connected nodes (in graph theory referred to as vertices) in topological graphs with many nodes (Brandenburg, Himsolt and Rohrer 1996). Hillier and Hanson’s (1984) gamma map structures the nodes in the graph by depth, in terms of topological steps, from the building entrance (see **Figure 1**).

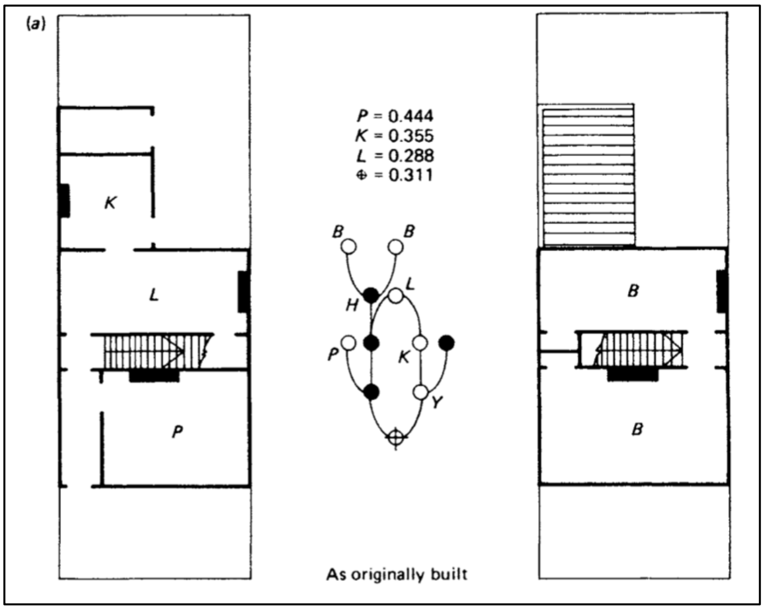


Figure 1: *Hillier and Hanson’s Gamma Map as abstracted from the example building. Source: Hillier and Hanson (1984: 156)*

3.2 Calculating Accuracy Measures from Survey View Building Sketch Maps

The method developed in this paper is based on Hillier and Hanson's (1984) gamma map. Topological graphs are abstracted from the sketch maps, from here on referred to as sketch graphs, which are then compared to the graph of the original floorplan, from here on referred to as the baseline graph. As in the gamma map, nodes represent distinct spaces and edges the permeability relationships between them. This approach processes the most reliable information contained within the sketch maps, topology and configuration (Golledge and Stimson 1977), and is less confounded by the subject's ability to produce an aesthetically pleasing floorplan sketch. A small graph matching approach for building floorplans has been proposed by Conroy Dalton and Kirsan (2008) in which the similarity between topological graphs is evaluated by the overall cost associated with transforming one into the other. This application is useful for deriving building genotypes from datasets of floorplans. However, for spatial cognition research, multiple related similarity measures capturing different aspects of the cognitive mapping process is deemed more beneficial. Accordingly, five different measures have been developed (see **Table 1**).

The *node count variance* measures how well the subject can remember the number of spaces present in the building. The location, relation to other spaces or existence of a particular space is irrelevant to the measure.

To understand whether the subject remembers specific spaces, a *node accuracy* measure is used. This measure does not account for non-existent spaces included in the sketch map or the relationships between spaces. Considered together, the *node count variance* and *node accuracy* account for how well the subject retained information about the existence of spaces in a building.

The *edge count variance* measures how well the subject can remember the number of permeability relationships between spaces present within the building. The location or existence of a particular relationship between spaces is irrelevant to the measure.

Just like the *node accuracy,* the *edge accuracy* measures the amount of specific permeability relationships the subject can remember. The introduction of non-existent relationships is not accounted for. Together the *edge count variance* and *edge accuracy* capture how well the subject retained information about the existence of permeability

Table 1: *Sketch Map Accuracy Measures*

Metric	Calculation	Description
Node Count Variance	$\sum V_S - \sum V_B$ Where: V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex	Quantifies the deviance of number of spaces in the sketch map from the baseline map
Edge Count Variance	$\sum E_S - \sum E_B$ Where: E_S is Sketch Graph Edge E_B is Baseline Graph Edge	Quantifies the deviance of number of permeability relationships in the sketch map from the baseline map
Node Accuracy	$\sum (V_S \in V_B) / \sum V_B$ Where: V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex	A percentage score quantifying the share of original spaces in the baseline map contained within the sketch map. If all spaces are included the score is 100%
Edge Accuracy	$\sum (E_S \in E_B) / \sum E_B$ Where: E_S is Sketch Graph Edge E_B is Baseline Graph Edge	A percentage score quantifying the share of original permeability relationships in the baseline map contained within the sketch map. If all spaces are included the score is 100%
Total Depth Variance	$\sum_j \sum_i d(V_{Sj}, V_{Si}) - \sum_j \sum_i d(V_{Bj}, V_{Bi})$ Where: d is distance V_{Sj} is Sketch Graph Origin Vertex V_{Si} is Sketch Graph Destination Vertex V_{Bj} is Baseline Graph Origin Vertex V_{Bi} is Baseline Graph Destination Vertex	Quantifies the extent to whicht the configurational characteristics of the spatial system are changed by changes in the sketch map. This is the closest to measuring a form of allocentric spatial understanding

relationships between spaces. Importantly, the node and edge metrics are interdependent, as a node omission will always lead to an edge omission and an edge omission can be coupled with a node omission. Accordingly, the metrics measure different dimensions of a multifaceted cognitive process, and one measure is not more important than another, or able to determine causality over another.

Lastly, the *total depth variance* measures the overall configurational similarity of the sketch map to the baseline map. The use of total depth as a variable that defines the configurational properties of a building is inspired by Hillier and Hanson's (1984) research. It is assumed that the characteristics of the global relationship between spaces will be most similar between floorplans with total depth values that are close to each other. **Figure 2** visualises how the accuracy measures are applied in practice on a baseline graph which has undergone several transformation scenarios that represent sketch graphs.

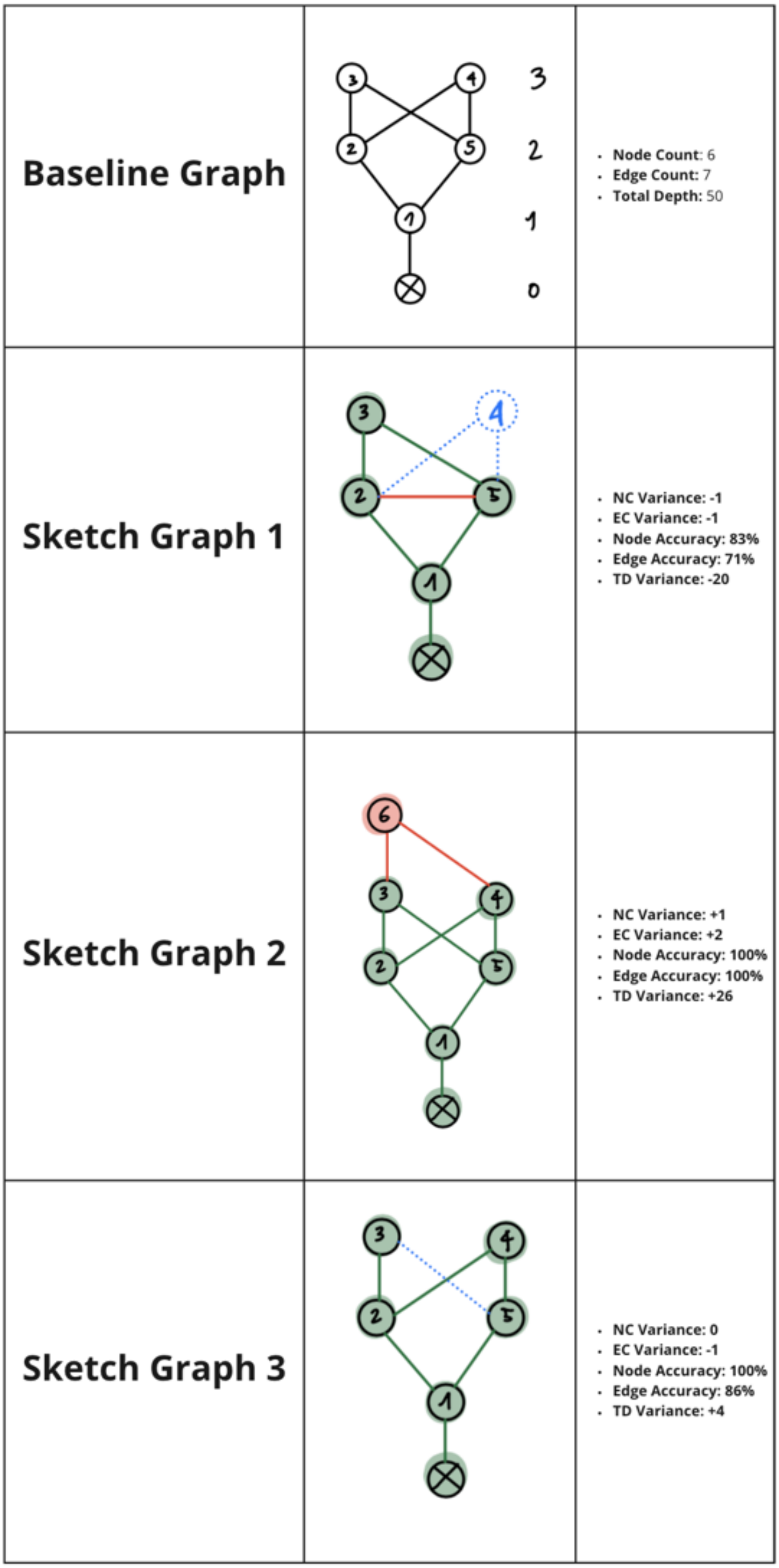


Figure 2: *Comparing sketch graphs to a baseline graph through the proposed accuracy measures.*

3.3 Including Feature Counts

In addition to topology and configuration, environmental feature occurrences and frequencies are another source of reliable information that can be gleaned from sketch maps (Golledge and Stimson 1997). By considering environmental features as subordinate elements of spaces, both local and global feature variance measures can be calculated. In a graph representation this entails treating feature counts as contextual data of the nodes to be compared between graphs. Two different global feature counts have been developed (see **Table 2**). The first is a global window variance measure which compares the number of windows in the sketch map to the baseline map. The second is a global convex space error measure. The errors comprise of omissions and additions. A convex space is a fundamental two-dimensional geometrical building block of a space in which each point is visible from all other points in space. Larger spaces can be broken down into the fewest and fattest convex spaces (Hillier and Hanson 1984). This metric should provide a measure of the subject's approximate geometrical understanding of the baseline map.

Table 2: *Feature Count Accuracy Measures*

Metric	Calculation	Description
Global Window Variance	$\frac{\sum W_s - \sum W_b}{\sum W_b}$ Where: W_s is Sketch Map Window W_B is Baseline Map Window	A percentage score quantifying the deviance of number of windows in the sketch map from the baseline map
Global Convex Space Errors	$\frac{\sum_{V_S \in V_B} \lvert CV_s - CV_B \rvert + \sum_{V_S \notin V_B} CV_s + \sum_{V_B \notin V_S} CV_B}{\sum CV_B}$ Where: V_S is Sketch Graph Vertex V_B is Baseline Graph Vertex CV_s is Convexity Count for Sketch Graph Vertex CV_B is Convexity Count for Baseline Graph Vertex	A percentage score quantifying the number of concex errors made in the sketch map as a share of all convex spaces in the baseline map

3.4 Digitizing the Process

The accuracy measures are particularly interesting to analyse when calculated for a large sample of sketch maps, so that aggregate differences and trends can be observed. This is exemplified in the analysis chapter (**Chapter 5**). As the process of calculating the accuracy measures becomes time intensive for large sample sizes, it needs to be automated. Generally, graphs can either be digitized as an adjacency matrix of nodes or an edge list in which edges are specified as node pairings (McNulty 2022). For the analysis conducted in this paper all sketch graphs were expressed as edge lists which were then transformed into graph objects in python using the NetworkX library (NetworkX). The advantage of storing all graphs as NetworkX graph objects is a suite of graph computation algorithms that can be called as methods on them. Having digitized all the graphs, a graph analyser was programmed in python which outputs the accuracy measures and local error statistics for all sketch map graphs as csv files (see **A2**).

3.5 Limitations

Whilst the proposed method is more objective than a qualitative scoring system, there are limitations. Firstly, a component of subjectivity remains when abstracting the sketch graphs from the sketch maps. This is a manual process and, as was experienced whilst trialling the method, there are still cases in which sketches were “messy” enough to involve a large amount of interpretation. Secondly, the method relies on the sketch maps being collected in the correct format. Comprehensive labelling of map features is required. If the labelling is missing, it is not possible to calculate the *node* and *edge accuracy* measures, as there is no way to verify whether the subject is referring to the specific element being matched for. Accordingly, this imposes restrictions on the experimental environments of spatial cognition research by requiring each space to possess distinct identifiable features that can be referred to in a sketch map.

4. Data

To test the proposed methods, the sketch map data collected from Jeffery *et al.'s* (2021) study "Visual imagination and cognitive mapping of a virtual building" was used (see **A1**). In this study passive explorations of two building layouts (see **Figure 3**), a rotational and mirror symmetry one, were simulated through different modalities. The modalities being video walk-through, verbal, and written description. The aim of the study was to ascertain the importance of verbal versus visual representations of the built environment to cognitive mapping performance. Furthermore, the effects of different environmental symmetries, as factors that may introduce confusion, and gender on cognitive map quality was examined for. The sample consisted of 80 participants, from which 156 sketch maps were collected.

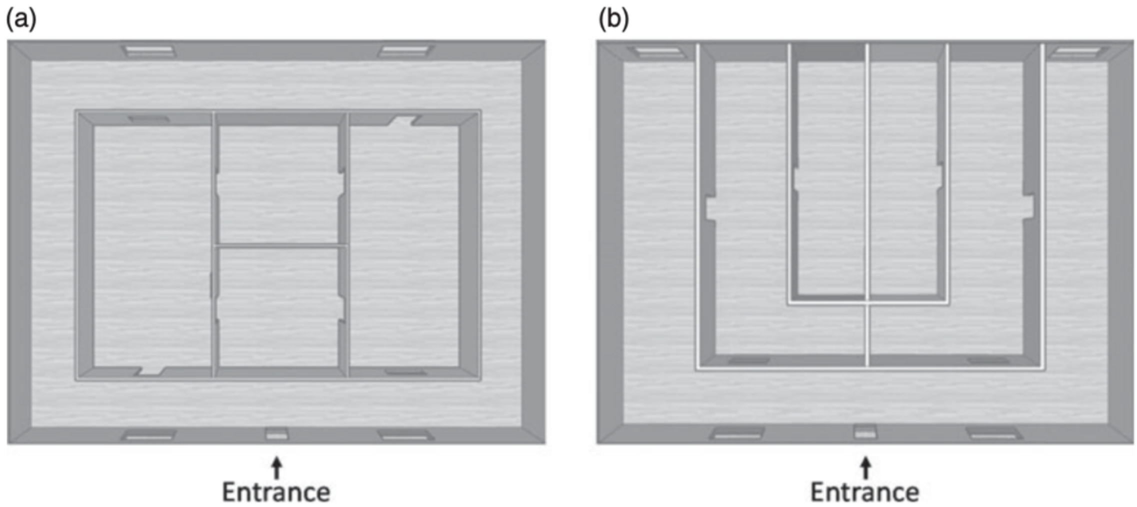


Figure 3: *Two building layouts in Jeffrey et al.'s study. Rotational symmetry (left) and mirror symmetry (right) Source: Jeffery et al. (2021: 4)*

5. Analysis

5.1 Floorplan Differences

Table 3: *High Level Comparison between Floorplans*

	Mean Edge Accuracy	SD	Mean Node Accuracy	SD	Mean Node Count Variance	SD	Mean Edge Count Variance	SD	Mean Total Depth Variance	SD
Rotational	0.86	0.20	0.99	0.04	0.07	0.38	-0.19	**0.61**	5.08	16.89
Mirror	0.88	0.19	0.98	0.06	0.06	0.64	-0.37	1.14	4.88	22.55

N = 151 (74 Miror and 77 Rotational)

Beginning with a high-level perspective, **Table 3** compares average accuracy measures between the rotational and mirror symmetry layout. 151 out of 156 sketch maps were processed, 74 of which were mirror symmetrical and 77 mirror rotational. The best performing metric is marked out in green. Neither floorplan consistently performs better across the accuracy metrics and all differences are statistically insignificant at a confidence level of 95% (two-sided Welch's t test). This suggests that environmental symmetry does not affect cognitive mapping quality.

Beyond the absence of significant aggregate differences, the accuracy measures reveal some interesting insights. Firstly, the subjects were better at remembering specific rooms (average node accuracy of 99% and 98%) than specific permeability relationships between them (average edge accuracy of 86% and 88%). This difference is statistically significant at a 95% confidence level (one-sided Welch's t test). The count variances reveal that on average subjects were slightly more likely to introduce more spaces to their sketch maps and include less permeability relationships than in the baseline maps. Lastly, the total depth variances are both rather low, suggesting that global configurational properties were well understood.

5.2 Modality Differences

Table 4: *Modality Comparison*

	Mean Edge Accuracy	SD	Mean Node Accuracy	SD	Mean Edge Count Variance	SD	Mean Node Count Variance	SD	Mean Total Depth Variance	SD
Spoken	0.81	0.23	0.98	0.07	-0.37	1.22	0.16	0.86	10.84	33.09
Written	0.84	0.21	0.99	0.06	-0.25	1.16	0.08	0.55	4.28	16.88
Video	0.93	0.15	0.99	0.04	-0.25	0.57	0.01	0.20	2.42	9.71

N = 151 (Spoken: 38, Written: 36, Video: 77)

Moving on to cognitive mapping performance by modality, **Table 4** compares the average accuracy measures between modalities for all floor plans. 38 spoken, 36 written and 77 video sketch maps were processed. The video mode consistently performs the best across all metrics except for the edge count variance. This suggests that mental map quality is better for environments experienced visually compared to written or spoken. Performance differences between modes are as expected, with spoken performing the worst, followed by written and video. Just as in the high-level analysis, the subjects were better at remembering specific rooms than specific permeability relationships, more likely to omit permeability relationships and more likely to include more spaces in their sketch map than in the baseline map.

However, apart from the edge accuracy measure for the video compared to written and spoken mode, all performance differences between modalities are insignificant at a confidence level of 95% (one-sided Welch's t tests). This suggests that subjects experiencing the layouts visually were significantly better at retaining specific permeability relationships in their sketch maps. It is possible that with a larger sample size the other differences may become significant.

5.3 Gender Differences

Table 5: *Gender Comparison*

	Mean Edge Accuracy	SD	Mean Node Accuracy	SD	Mean Edge Count Variance	SD	Mean Node Count Variance	SD	Mean Total Depth Variance	SD
Female	0.88	0.19	0.99	0.05	0.44	0.82	0.14	0.54	7.96	22.19
Male	0.87	0.20	0.98	0.06	0.50	0.87	0.12	0.47	5.57	12.93

N = 76 (Female: 48, Male: 28)

Table 5 compares differences in performance between gender for a subset of the participants that experienced the floorplans in visual mode. The differences are small, and none are statistically significant at a 95% confidence level (two-sided Welch's t test). This suggests that gender does not affect mental map quality.

5.4 Convex Space Errors and Window Count Variance

Table 6: *Convex Space Errors and Window Count Variance*

	Convex Space Errors	SD	Window Count Variance	SD
Rotational	0.06	0.2	0.03	0.24
Mirror	0.15	0.26	-0.06	0.21

Table 6 shows the average convex space errors and window count variances for the two floorplans. The rotational floorplan performs significantly better across both metrics at a 95% confidence level (two-sided Welch's t test). In the rotational symmetry floorplan, the convex space errors were 6% compared to 15% in the mirror symmetry layout. Arguably the rotational symmetry floorplan is less geometrically complex, consisting of more single-convex spaces. Subjects were also more likely to introduce new windows in the rotational symmetry floorplan

(+3% variance) and more likely to omit windows in the mirror symmetry floorplan (-6% variance).

5.5 Local Error Distributions and Path Effect

The previous analyses focused on global measures of sketch map accuracy. These can be disaggregated into local error distributions, such as the edge omission distribution, and further analysed. **Figure 4** visualises the local error distribution of all edge omissions. 66 edges were omitted in the rotational symmetry layout and 59 edges in the mirror symmetry layout. There is an interesting pattern in which edges that are symmetrical to other edges in terms of the permeability relationship they represent have not been omitted the same amount. This may be due to a path effect in which edges encountered later in the passive walk through of the floorplan are more likely to be omitted by the subject. **Figure 5** visualises this relationship with the sequence position of the edge on the x-axis and the corresponding omission rate on the y-axis. There is a strong positive correlation with an R-value of 0.7 that is statistically significant at a confidence level of 95%. This insight may reveal something interesting about the cognitive mapping process. The normal assumption is that edges encountered later are omitted less as they are more current in the memory of the subject. However, this finding suggests that in creating the sketch map the subject may be attempting to construct a mental representation sequentially starting off from the beginning of the tour.

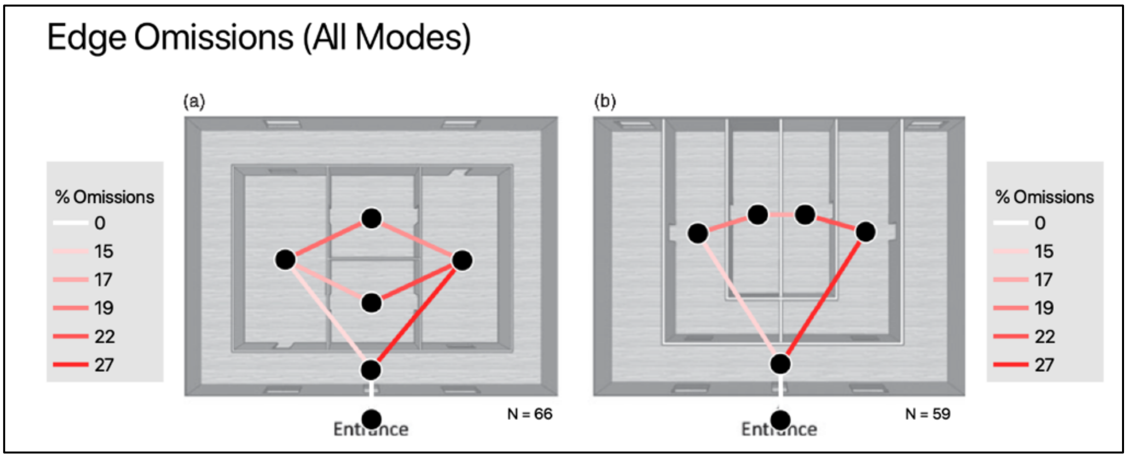


Figure 4: *Local error distribution of edge omissions*

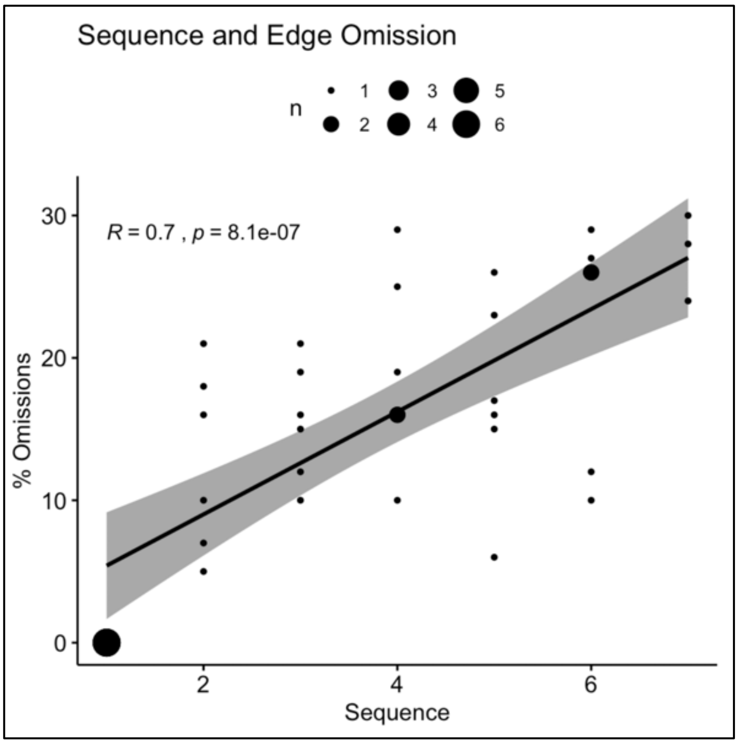


Figure 5: *Correlation between sequence position and omission of edges.*

6. Discussion

6.1 Comparing the Graph Matching Method to Qualitative Accuracy Scores

Analysing the sketch map data through the proposed graph matching method produces insights that corresponded with Jeffery *et al.'s* (2021) findings. Namely, that there is no difference in map quality between building type, between males and females and between the spoken and written mode, and that map quality is better for video than verbal mode (see **Table 5**). This evidences that the method can be used to at least arrive at similar conclusions as with a qualitative scoring system.

However, the analysis also showcases how the accuracy measures enable more specific insights. For instance, higher map quality for the video presentation is attributed specifically to more of the original permeability relationships being retained (see **section 5.2**). Off course a qualitative scoring system (such as accurate vs. inaccurate map) may also account for how well geometrical properties are preserved in the sketch maps, which this topology-based method cannot. However, what exactly is factored into a qualitative evaluation and how consistently remains unclear. This makes research results not strictly reproducible. Therefore, whilst a qualitative scoring system may be more holistic in terms of considering several domains of sketch map accuracy at once, the graph matching method is specific about the aspects in which sketch map quality is significantly different. This will ultimately help pinpoint those aspects of the cognitive mapping process that are different across modes of representation, gender, or floorplan layouts.

Furthermore, the method can provide new insights beyond those of Jeffery *et al.'s* (2021) study. For example, that distinct spaces are retained better in sketch maps than the permeability relationships between them (see **section 5.1**), that there is a path effect that determines the likelihood of permeability relationships being retained (see **section 5.5**), or that feature frequency information is better retained for the rotational symmetry layout (see **section 5.4**). Accordingly, these bring to light new questions for which experiments can be devised to examine them in more depth.

Table 5: *Comparing findings to Jeffery et al. (2021)*

No.	Jeffery *et al.'s* (2021) findings	Alignment with this research	Relevant (examined for)
1	Map scores correlated with multiple choice scores	NO	NO
2	Males and Females did not differ in Map quality or MCQ scores	**YES**	**YES**
3	Map quality was higher for video than verbal presentation	**YES**	**YES**
4	No differences in spoken versus written map scores	**YES**	**YES**
5	MCQ scores and map quality did not differ between building typ	**YES**	**YES**
6	No interaction between presentation type and building symmetr	NO	NO

6.2 Trial Limitations

The floorplans used in Jeffery *et al.'s* (2021) research are rather similar and simple spatial systems. Accordingly, only 40% of all sketch maps deviated in some way from the baseline map. It is possible that the full potential of the accuracy measures could not be established through the analysis. There remains a need to test the accuracy measures in experiments with different floorplans that are more complex configurational iterations of each other. Furthermore, if comparisons of cognitive mapping performance between different spatial systems, in terms of complexity and size, are undertaken there may be a need to standardise the accuracy measures to make them comparable.

6.3 Future Research Agenda

As touched upon in **section 6.1,** the accuracy measures can produce new findings alongside which the potential for new research is unlocked. The method may be extended to a three-dimensional framework. It may be used to explore the relationship between building-scale spatial configurations and allocentric cognitive performance or to explore the relationship between route sequencing and mental mapping in buildings. If enough sketch map data is collected, the method could even provide a basis for a machine learning model which predicts mental mapping difficulties in building layout proposals of equivalent spatial complexity.

7. Conclusion

This paper has outlined how Hillier and Hanson's (1984) gamma map method can be adapted to arrive at a novel topological graph-based approach to measuring the accuracy of survey view building sketch maps in spatial cognition research. This method provides a more precise, reproducible, and objective alternative to the qualitative scoring systems applied to measuring sketch map accuracy up until now. By testing the method on the data collected from Jeffery *et al.'s* (2021) cognitive mapping research, its efficacy was demonstrated. This has shown that by comparing sketch and baseline graphs, measures that usefully quantify a subject's ability to recall aspects of the experimental environment could be arrived at. Specifically, the analysis of Jeffery *et al.'s* (2021) data produced no evidence of gender biases in cognitive mapping quality. However, it evidenced that spatial information in video mode is better processed than in written or verbal descriptions. Furthermore, the analysis of local error distributions showed that recall is related to the path sequence of the subjects within the experimental environment. This raises questions for future research on whether such a path effect is equally observable in active explorations of virtual environments and if configurational constraints on possible path sequences might lead to systematic results. Importantly, these findings suggest that there is fertile ground for the application of this method in future spatial cognition research. The general philosophy of the approach to harness the topological information embedded in sketch maps leaves potential to tailor and develop the accuracy measures to fit the specific needs of future spatial cognition research.

8. References

Appleyard, D. (1970) ‘Styles and Methods of Structuring a City’, *Environment and behavior*, 2, 1, 100–117

Billinghurst, M. and Weghorst, S. (1995) ‘The use of sketch maps to measure cognitive maps of virtual environments’, in *Proceedings Virtual Reality Annual International Symposium*’, 95, 1995 IEEE, 40–47.

Blades, M. (1990) ‘The Reliability of Data Collected From Sketch Maps’, *Journal of Environmental Psychology*, l10, 327-339

Brandenburg, F. J. , Himsolt, M. and Rohrer, C. (2005) ‘An experimental comparison of force-directed and randomized graph drawing algorithms’, in *Graph Drawing*, Berlin, Heidelberg: Springer Berlin Heidelberg, 76–87

Dalton, R. C. and Kirsan, C. (2008) ‘Small-Graph Matching and Building Genotypes’, *Environment and planning. B, Planning & design*, 35, 5, 810–830

Dalton, R. C., Hölscher, C. and Alasdair, T. (2012) ‘Understanding Space: The Nascent Synthesis of Cognition and the Syntax of Spatial Morphologies’, *Environment and planning. B, Planning &* design, 39, 1, 7–11.

Downs, R. M. and Stea, D. (1973) *Image and environment: Cognitive mapping and spatial behavior*, Chicago: Aldine Pub. Co.

Evans, G.W., Fellows, J., Zorn, M., Doty, K. (1980) ‘Cognitive mapping and architecture’, *Journal of applied psychology*, 65, 4, 474–478

Gärling, T. and Golledge, R.G. (1989). ‘Environmental Perception and Cognition’, In Zube, E.H., Moore, G.T. (eds) *Advances in Environment, Behavior, and Design*, vol 2. Boston: Springer

Golledge, R. G. and Stimson, R. J. (1997) ‘Spatial Cognition, Cognitive Mapping, and Cognitive Maps’ in *Spatial Behavior A Geographic Perspective*, 224-266, New York: The Guilford Press

Golledge, R.G. (1987) ‘Environmental cognition’, In D. Stokols and I. Altman (eds) *Handbookof Environmental Psychology*, 1, New York: Wiley

Golledge, R.G., Smith,T.R., Pellegrino, J.W., Doherty, S. and Marshall, S.P. (1985) ‘A conceptual model and empirical analysisof children'sacquisition of spatial knowledge’, *Journal of Environmental Psychology*, 5, 125-152

Haq, S. (2003) ‘Investigating the Syntax Line: Configurational Properties and Cognitive Correlates’, *Environment and planning. B, Planning &* design, 30, 6, 841–863

Haq, S. and Girotto, S. (2003) ‘Ability and intelligibility: Wayfinding and environmental cognition in the designed environment’, in *Proceedings of the 4th International Symposium on Space Syntax*, 68.1 – 68.20

Hart, R. A. and Moore, G. T. (1973) ‘The Development of Spatial Cognition: A Review’, In R. M. Downs & D. Stea (Eds.), *Image & environment: Cognitive mapping and spatial behavior* (pp. 246–288)

Hillier, B. and Hanson, J. (1984) *The Social Logic of Space*. Cambridge: Cambridge University Press

Jeffery, K., Guo, W., Ball, D. and Rodriguez-Sanchez, J. (2021) ‘Visual imagination and cognitive mapping of a virtual building’, *The Journal of Navigation*, 1–14

Kim, Y. O. (2001) ‘The role of Spatial Configuration in Spatial Cognition’, in J. Peponis, J. Wineman and S. Bafna (eds.), *Proceedings of the 3rd International Symposium on Space Syntax*, 49.41- 49.21

Kim, Y.O., Penn, A. (2004) ‘Linking the spatial syntax of cognitive maps to the spatial syntax of the environment’, *Environment and Behavior*, 36, 483-504

Kuipers, B. (1983) ‘The Cognitive Map: Could It Have Been Any Other Way?’, In Pick, H.L., Acredolo, L.P. (eds) *Spatial Orientation*, Boston: Springer

Lynch, K. (1960) *The Image of the City*, Cambridge, Massachusetts: MIT Press

Mc Nulty, K. (2022) *Handbook of Graphs and Networks in People Analytics,* Routledge

Moeser, S. D. (1988) ‘Cognitive Mapping in a Complex Building’, *Environment and behavior*, 20, 1, 21–49

NetworkX (2022) https://networkx.org (accessed: 13th of May 2022)

Newcombe, N. (1985) ‘Methods for the Study of Spatial Cognition’, In R. Cohen (ed) *The Developmentof Spatial Cognition*, Hillsdale, NJ: Erlbaum

Peer, M., Brunec, I.K., Newcombe, N.S. and Epstein R.A. (2021) ‘Structuring Knowledge with Cognitive Maps and Cognitive Graphs’, *Trends in Cognitive Sciences,* 25, 1, 37-54

Peponis, J., Zimring, C. and Choi, Y. K. (1990) ‘Finding the building in wayfinding’, *Environment and Behavior,* 22, 5, 555-590

Ratti, C. (2004) ‘Space Syntax: Some Inconsistencies’, *Environment and planning. B, Planning & design*, 31, 4, 487–499

Sadalla, E. K. and Montello, D. R. (1989) ‘Remembering Changes in Direction’, *Environment and behavior*, 21, 3, 346–363

Siegel, A.W. and Cousins, J. H. (1985) ‘The symbolizingand symbolized child in the enterprise of cognitive mapping’, In R. Cohen (ed) *The Developmentof Spatial Cognition*, Hillsdale, NJ: Erlbaum

Tolman, E. C. (1948) ‘Cognitive maps in rats and men’, *Psychological Review*, **55**(4), 189–208

B. Supporting Evidence

B1. Similarity Signature Tables

Node Similarity Signatures Compared to Node 21 (Hybrid-Axial Graph)

Node	Similarity Value
1	0.975
2	0.823
3	0.975
4	0.975
5	0.975
6	0.975
7	0.975
8	0.975
9	0.844
10	0.844
11	0.851
12	1
13	1
14	0.989
15	0.989
16	1
17	1
18	0.844
19	0.844
20	0.851
21	1
22	1
23	1
24	0.989
25	0.989
26	1

Node Similarity Signatures Compared to Node 2 (Hybrid-Axial Graph)

Node	Similarity Value
1	0.82
2	1
3	0.82
4	0.82
5	0.82
6	0.82
7	0.82
8	0.82
9	0.918
10	0.918
11	0.91
12	0.823
13	0.823
14	0.826
15	0.826
16	0.823
17	0.823
18	0.918
19	0.918
20	0.91
21	0.823
22	0.823
23	0.823
24	0.826
25	0.826
26	0.823

Node Similarity Signatures Compared to Node 1 (Hybrid-Axial Graph)

Node	Similarity Value
1	1
2	0.82
3	1
4	1
5	1
6	1
7	1
8	1
9	0.858
10	0.858
11	0.839
12	0.975
13	0.975
14	0.971
15	0.971
16	0.975
17	0.975
18	0.858
19	0.858
20	0.839
21	0.975
22	0.975
23	0.975
24	0.971
25	0.971
26	0.975

Node Similarity Signatures Compared to Node 10 (Hybrid-Axial Graph)

Node	Similarity Value
1	0.858
2	0.918
3	0.858
4	0.858
5	0.858
6	0.858
7	0.858
8	0.858
9	1
10	1
11	0.937
12	0.844
13	0.844
14	0.843
15	0.843
16	0.844
17	0.844
18	1
19	1
20	0.937
21	0.844
22	0.844
23	0.844
24	0.843
25	0.843
26	0.844

Node Similarity Signatures Compared to Node 21 (L-Node Graph)

Node	Similarity Value
1	0.543
2	0.422
3	0.666
4	0.654
5	0.605
6	0.529
7	0.617
8	0.632
9	0.529
10	0.541
11	0.576
12	0.517
13	0.485
14	0.58
15	0.554
16	0.575
17	0.638
18	0.701
19	0.649
20	0.908
21	1
22	0.717
23	0.727
24	0.54
25	0.72
26	0.788

Node Similarity Signatures Compared to Node 2 (L-Node Graph)

Node	Similarity Value
1	0.478
2	1
3	0.586
4	0.443
5	0.47
6	0.638
7	0.497
8	0.572
9	0.624
10	0.608
11	0.421
12	0.441
13	0.586
14	0.499
15	0.477
16	0.546
17	0.44
18	0.411
19	0.558
20	0.44
21	0.422
22	0.449
23	0.474
24	0.523
25	0.438
26	0.4

Node Similarity Signatures Compared to Node 1 (L-Node Graph)

Node	Similarity Value
1	1
2	0.478
3	0.719
4	0.845
5	0.899
6	0.519
7	0.671
8	0.673
9	0.639
10	0.782
11	0.795
12	0.859
13	0.757
14	0.79
15	0.838
16	0.709
17	0.686
18	0.797
19	0.403
20	0.57
21	0.543
22	0.412
23	0.381
24	0.297
25	0.574
26	0.493

Node Similarity Signatures Compared to Node 10 (L-Node Graph)

Node	Similarity Value
1	0.782
2	0.608
3	0.785
4	0.674
5	0.696
6	0.646
7	0.655
8	0.658
9	0.812
10	1
11	0.593
12	0.659
13	0.799
14	0.743
15	0.736
16	0.712
17	0.637
18	0.595
19	0.577
20	0.557
21	0.541
22	0.57
23	0.449
24	0.447
25	0.523
26	0.546

Node Similarity Signatures Compared to Node 21 (Adjacency Graph)

Node	Similarity Value
1	0.887
2	0.659
3	0.875
4	0.866
5	0.875
6	0.875
7	0.866
8	0.875
9	0.67
10	0.67
11	0.71
12	1
13	0.836
14	0.831
15	0.831
16	0.836
17	1
18	0.67
19	0.67
20	0.71
21	1
22	1
23	0.836
24	0.831
25	0.831
26	0.836

Node Similarity Signatures Compared to Node 2 (Adjacency Graph)

Node	Similarity Value
1	0.699
2	1
3	0.63
4	0.641
5	0.63
6	0.63
7	0.641
8	0.63
9	0.773
10	0.773
11	0.736
12	0.659
13	0.687
14	0.657
15	0.657
16	0.687
17	0.659
18	0.773
19	0.773
20	0.736
21	0.659
22	0.659
23	0.687
24	0.657
25	0.657
26	0.687

Node Similarity Signatures Compared to Node 1 (Adjacency Graph)

Node	Similarity Value
1	1
2	0.699
3	0.913
4	0.905
5	0.913
6	0.913
7	0.905
8	0.913
9	0.751
10	0.751
11	0.688
12	0.887
13	0.771
14	0.744
15	0.744
16	0.771
17	0.887
18	0.751
19	0.751
20	0.688
21	0.887
22	0.887
23	0.771
24	0.744
25	0.744
26	0.771

Node Similarity Signatures Compared to Node 10 (Adjacency Graph)

Node	Similarity Value
1	0.751
2	0.773
3	0.692
4	0.688
5	0.692
6	0.692
7	0.688
8	0.692
9	1
10	1
11	0.852
12	0.67
13	0.787
14	0.766
15	0.766
16	0.787
17	0.67
18	1
19	1
20	0.852
21	0.67
22	0.67
23	0.787
24	0.766
25	0.766
26	0.787

B2. Analysis, Custom Function and Output Files

R. files:

- *ExpAnalysis_ComplexBuildings.Rmd*
- *SignatureSimilarity.Rmd*
- *Survey_analysis.Rmd*

Jupyter Notebooks:

- *Graph_Analyser.ipynb*
- *Signature_Similarity.ipynb*

Outputs:

- msim_orbits.csv
- msim_adj_orbits.csv
- mirrorL_orbits.csv
- as_accuracy.csv
- ms_accuracy.csv
- rs_accuracy.csv
- as_ax_accuracy.csv
- ms_ax_accuracy.csv
- AS_AX_stats.csv

- MS_AX_stats.csv
- RS_AX_stats.csv

Data and files are available upon request (bruce.timothy@hotmail.com)

B3. Survey

Introduction

In spatial cognition research sketch maps are often used to elucidate knowledge about a subject's environmental knowledge. Evaluating the accuracy and informational completeness of these maps is challenging and could undermine the reproducibility of many studies. This survey is designed to evaluate the variance of qualitative scoring approaches used in current practice as well as components of a novel quantitative approach being developed as part of my master thesis. **You will be asked to complete three different tasks** involving sketch maps collected from recent research. **The entire survey should take no longer than 20 minutes to complete.** Thank you for your participation!

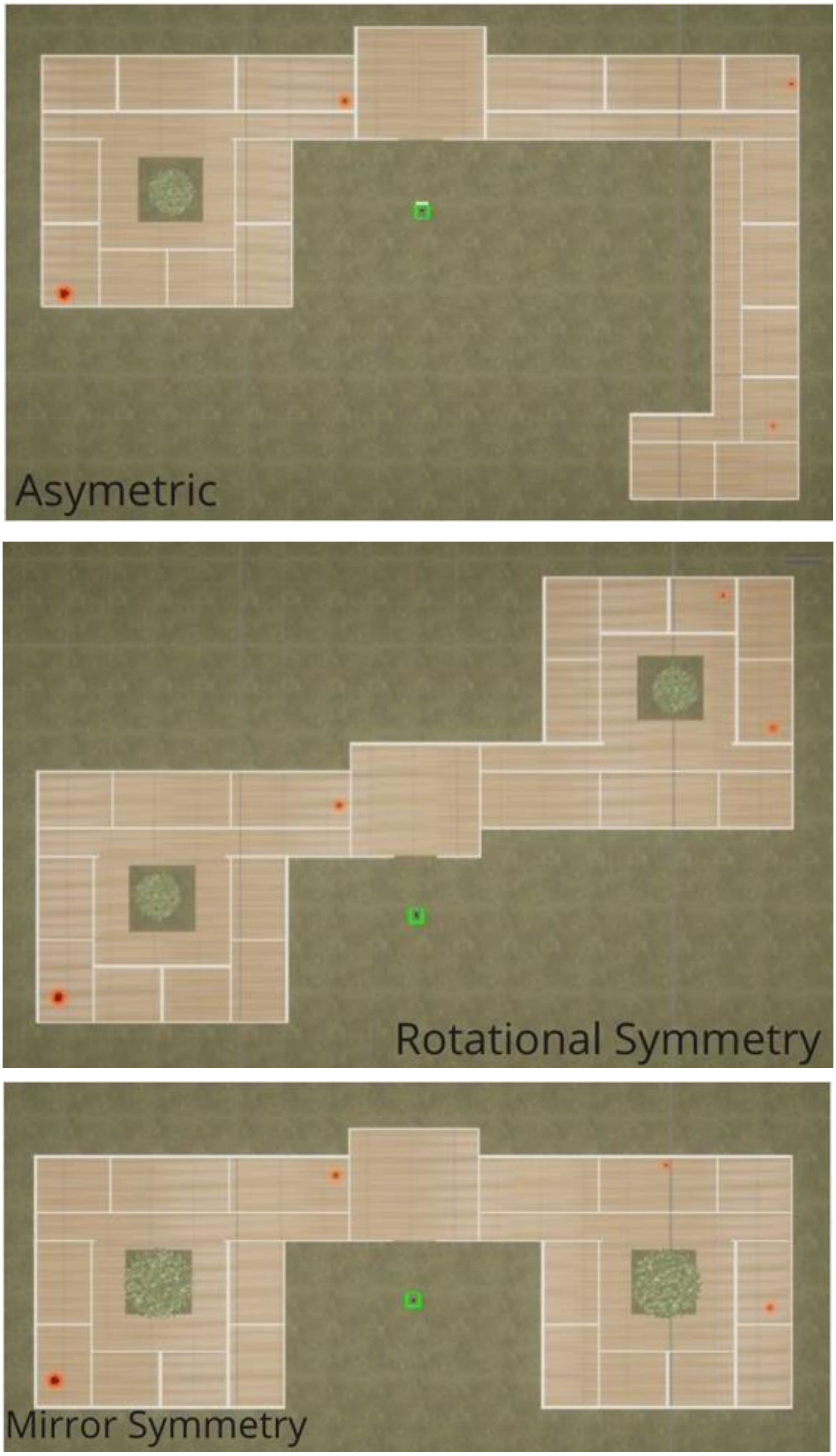
Participant Information

Name, Surname:
Email Address:
Age:
Gender:
Current Occupation:
Education in a field of the built environment (Yes/No):

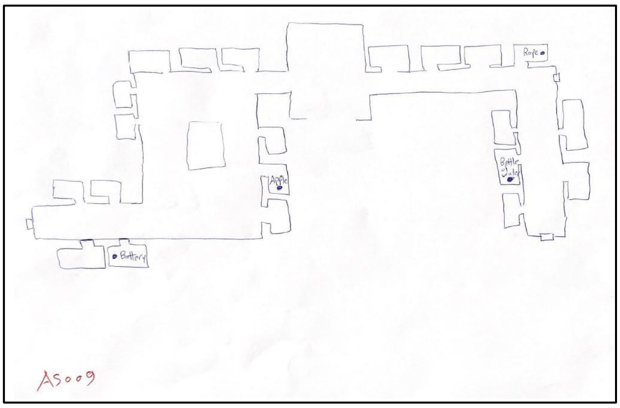
Task 1 *(5mins)*

1.1
In this task you will be presented with a sample of 12 sketch maps collected from spatial cognition research. There are three different building types represented in the sample that are all distinguished by their symmetrical properties (each occur four times): Rotational Symmetry, Mirror Symmetry and Asymmetric (see **Figure 1**). Your task is to rank all the sketch maps by accuracy in order from best to worst, with **1. being the best** and **12. being the worst.** The ranking task is supposed to be based on intuition, so please do not spend too much time deliberating over your choice. A table will be provided with the sketch map codes where you can enter the rank you wish to allocate to each sketch map.
1.2
After ranking the sketch maps, you will be requested to score each of them individually on a scale from 1-5, with **5 denoting a perfect map** and **1 no resemblance** to the original floorplans. This task is supposed to be based on intuition, so please do not spend too much time deliberating over your choice. A table will be provided with the sketch map codes where you can enter the score you wish to allocate to each sketch map.

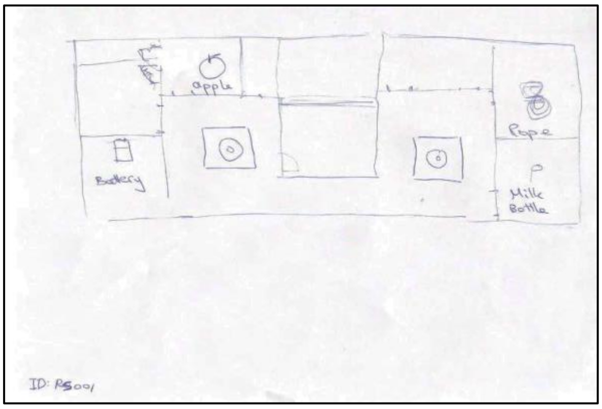
Figure 1: *Original Floorplans*



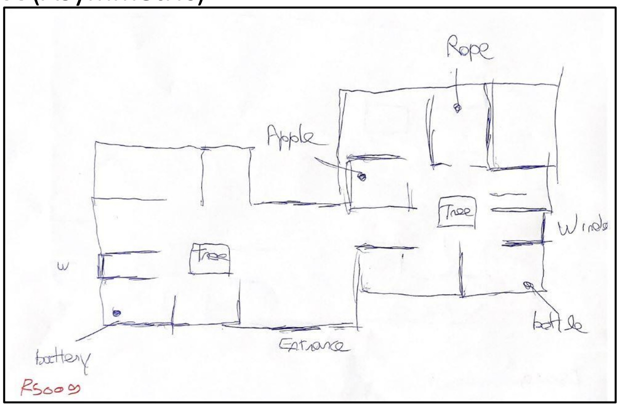
Sketch Maps: Task 1



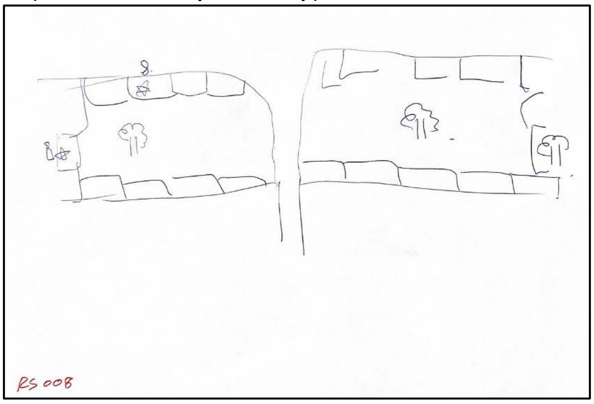
A (Asymmetric)



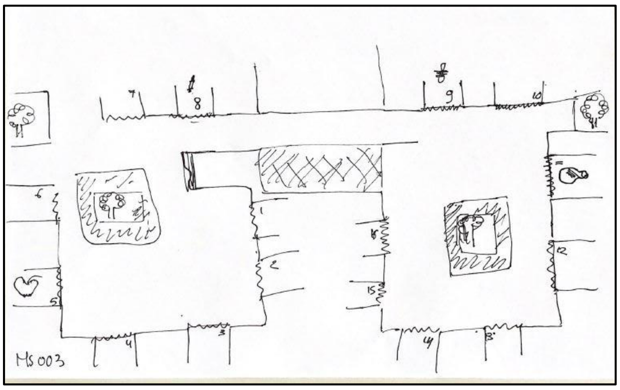
B (Rotational Symmetry)

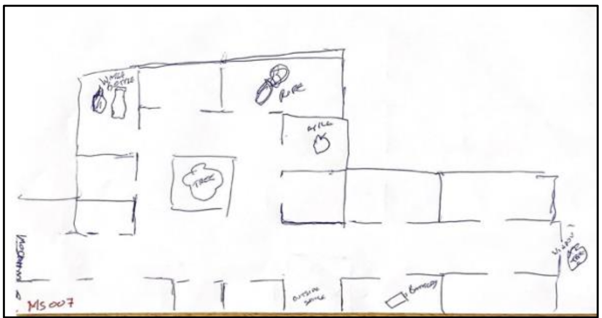


C (Rotational Symmetry)

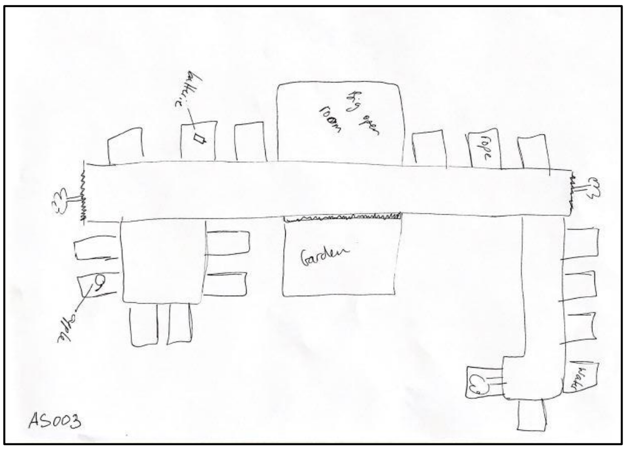


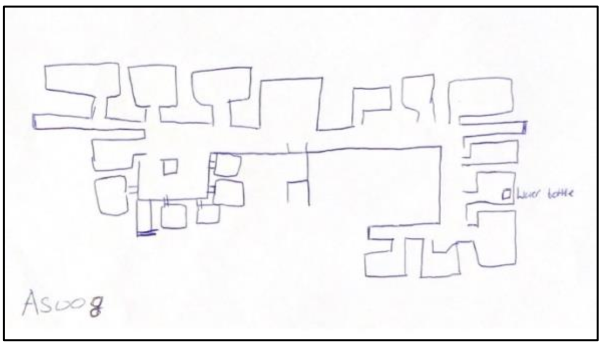
D (Rotational Symmetry)





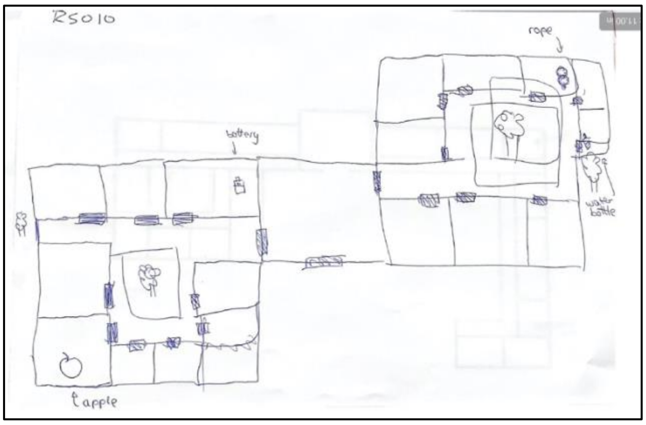
F (Mirror Symmetry)



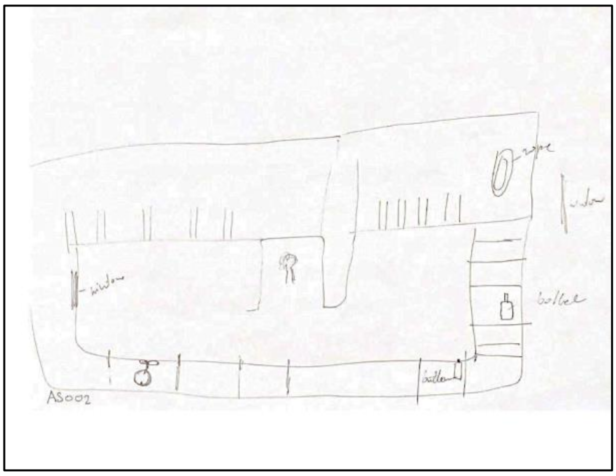


H (Asymmetry)

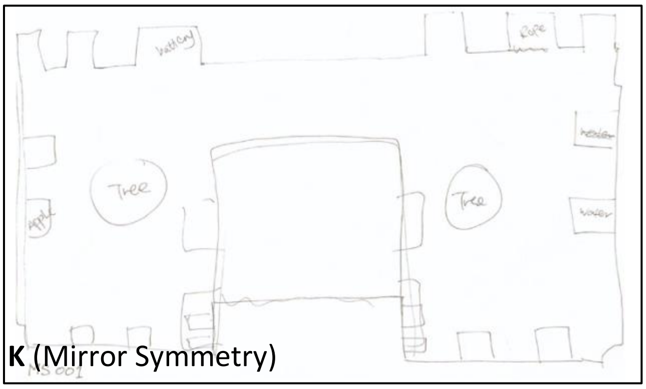
G (Asymmetry)



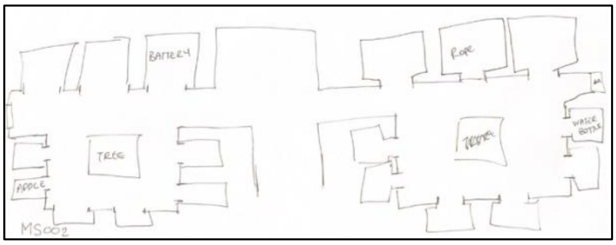
I (Rotational Symmetry)



J (Asymmetry)



K (Mirror Symmetry)



L (Mirror Symmetry)

Table for task 1.1 and 1.2

Sketch Map Code	Rank (1-12)	Score (1-5)
A		
B		
C		
D		
E		
F		
G		
H		
I		
J		
K		
L		

<u>Task 2 *(5mins)*</u>

2.1
In this task you will be presented with another sample of 12 sketch maps collected from spatial cognition research. This time there are two different and smaller building types represented in the sample which also distinguish themselves by their symmetrical properties: Rotational Symmetry and Mirror Symmetry (see **Figure 2).** As in **task 1.1**, you will be required to rank all the sketch maps by accuracy in order from best to worst, with **1. being the best** and **12. being the worst.** The ranking task is supposed to be based on intuition, so please do not spend too much time deliberating over your choice. A table will be provided with all the sketch map codes where you can enter the rank you wish to allocate to each sketch map.
2.2
After ranking the sketch maps, you will be requested to score each of them individually on a scale from 1-5, with **5 denoting a perfect map** and **1 no resemblance** to the original floorplans. The task is supposed to be based on intuition, so please do not spend too much time deliberating over your choice. A table will be provided with the sketch map codes where you can enter the score you wish to allocate to each sketch map.

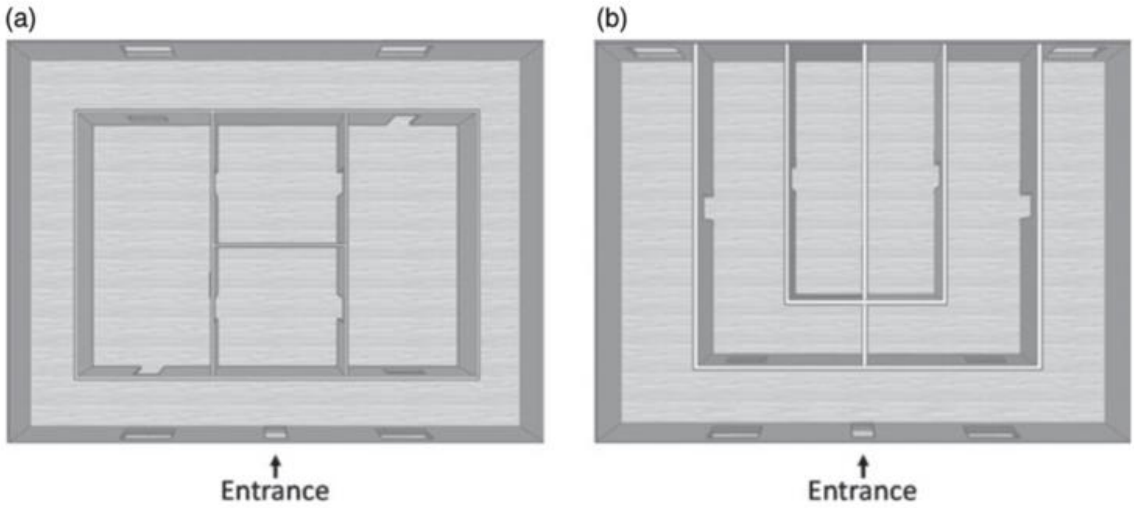
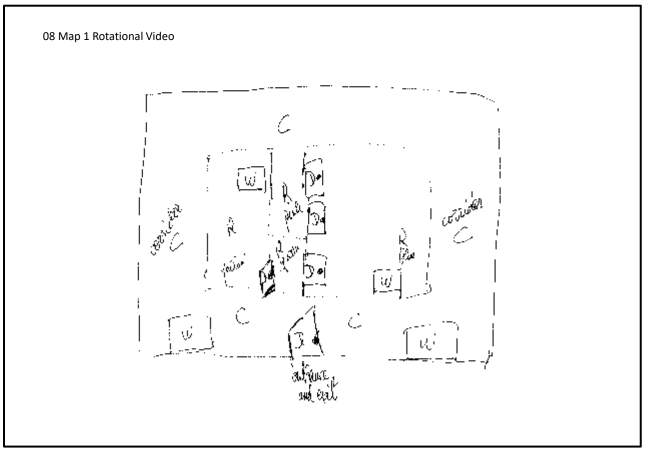
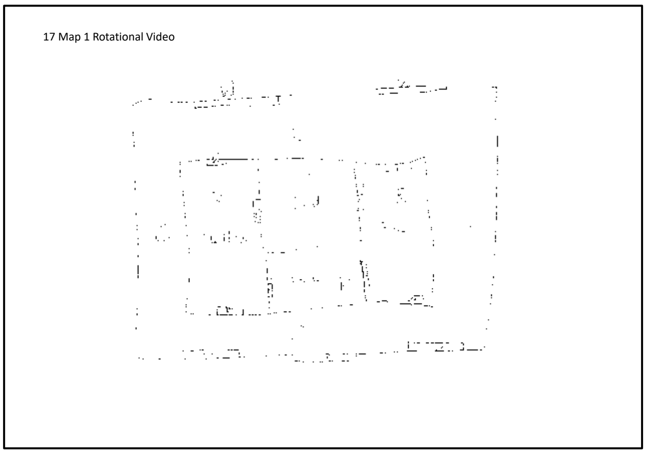


Figure 2: *Original Floorplans*

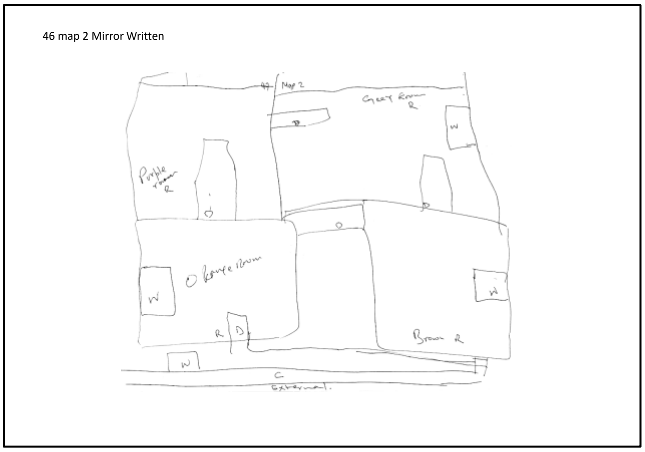
Sketch Maps for Task 2



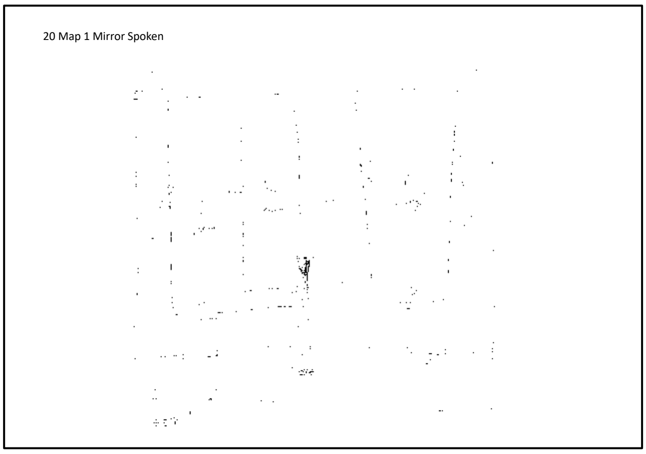
A (Rotational Symmetry)



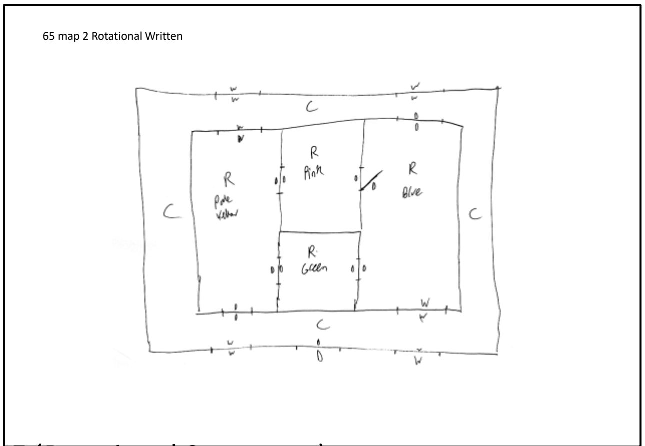
B (Rotational Symmetry)



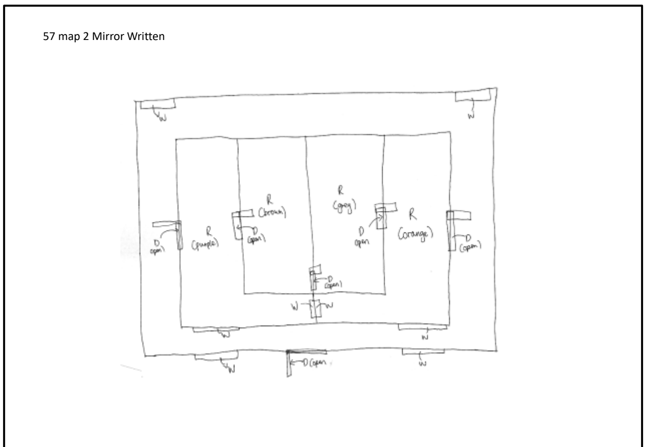
C (Mirror Symmetry)



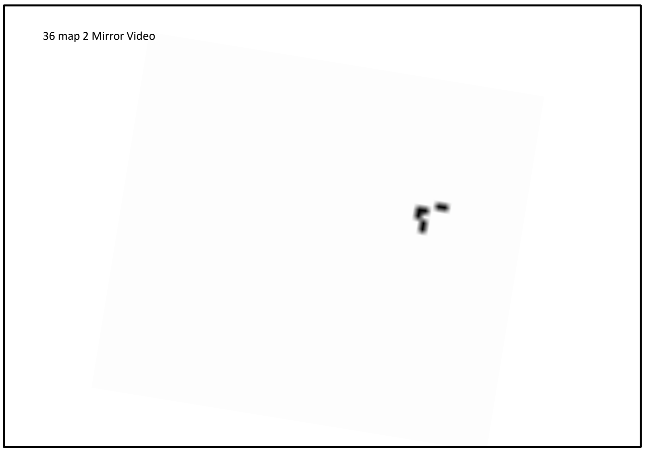
D (Mirror Symmetry)



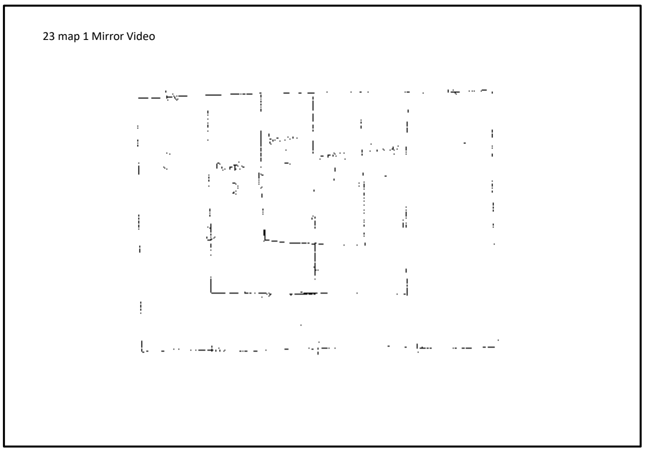
E (Rotational Symmetry)



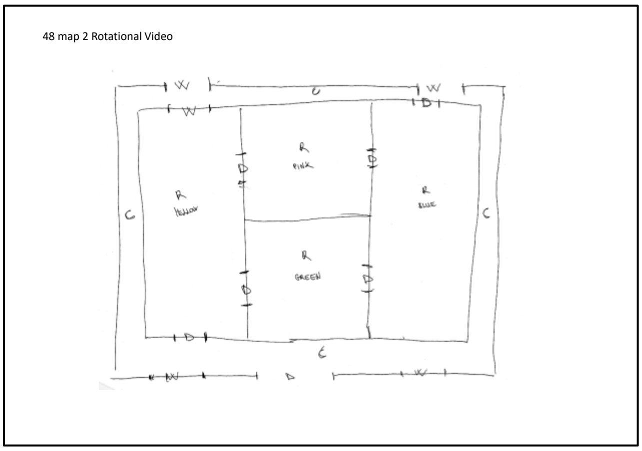
F (Mirror Symmetry)



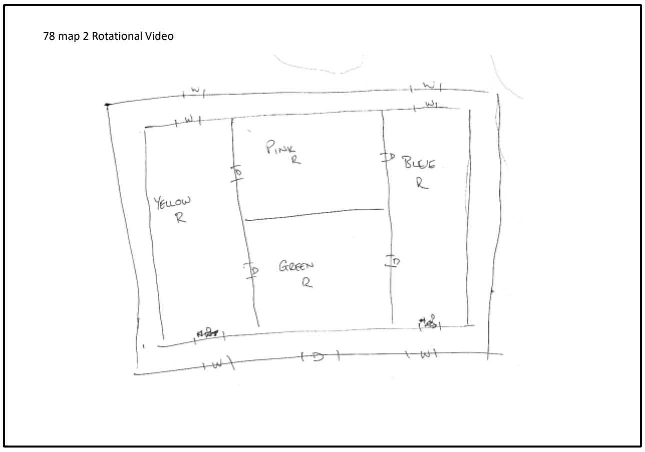
G (Mirror Symmetry)



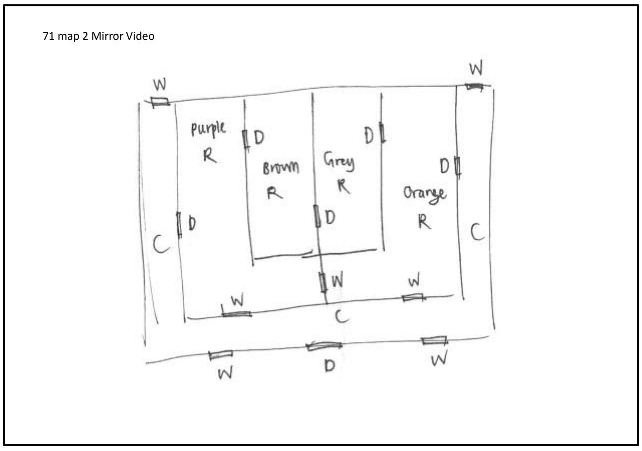
H (Mirror Symmetry)



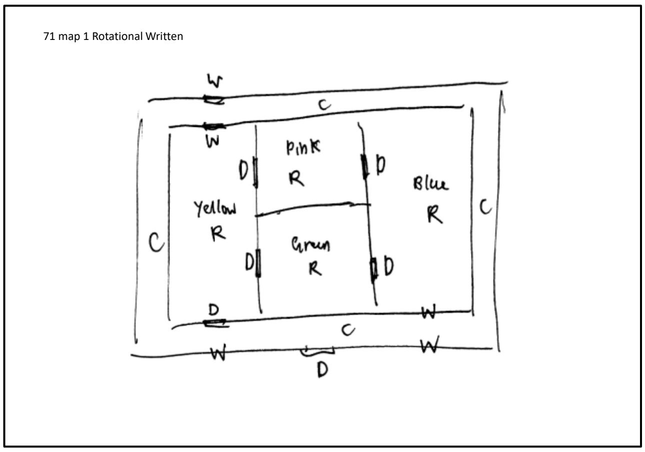
I (Rotational Symmetry)



J (Rotational Symmetry)



K (Mirror Symmetry)



L (Rotational Symmetry)

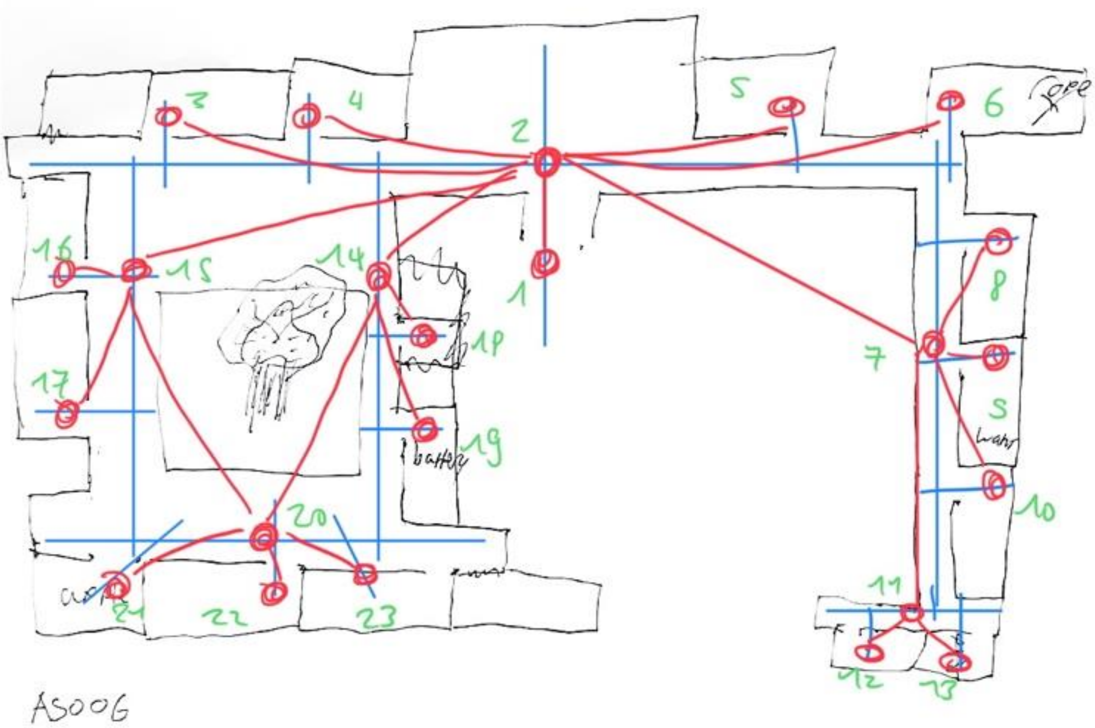
Table for 2.1 and 2.2

Sketch Map Code	Rank (1-12)	Score (1-5)
A		
B		
C		
D		
E		
F		
G		
H		
I		
J		
K		
L		

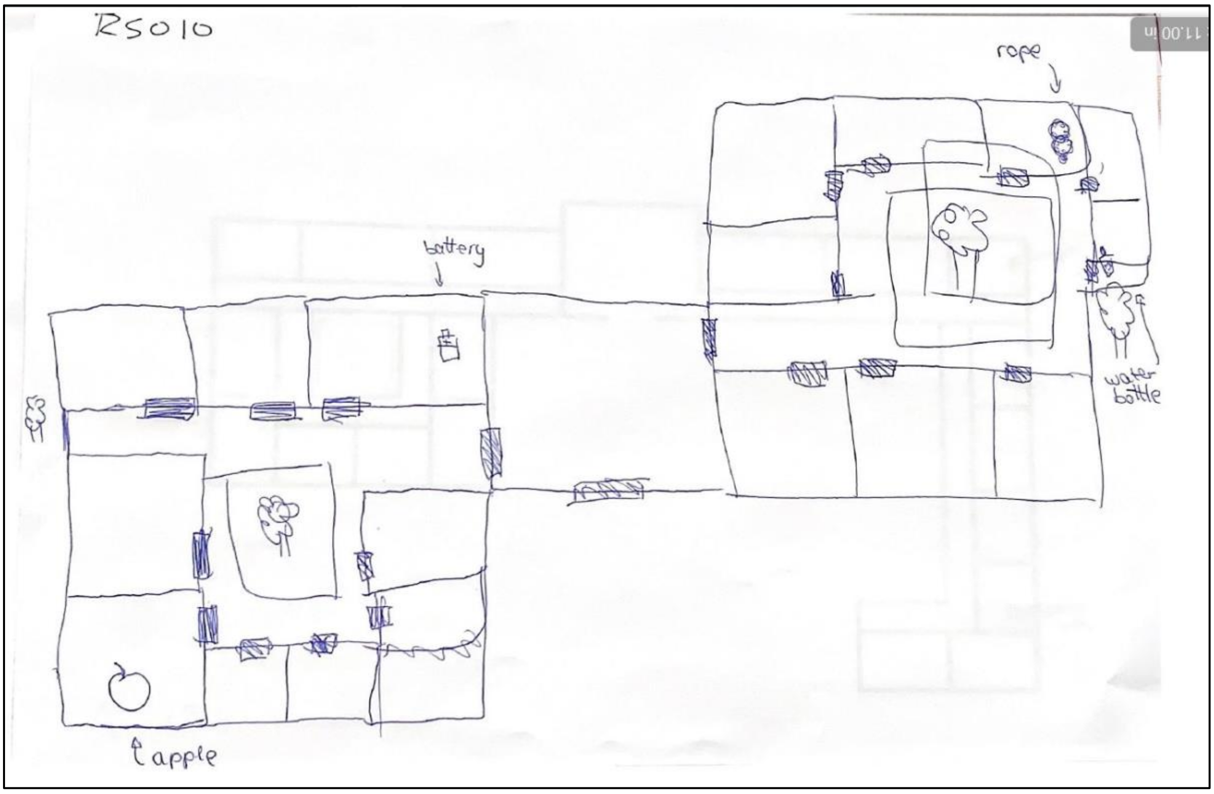
__Task 3 (Space Syntax Practitioners only and semi guided) *(10mins)*__

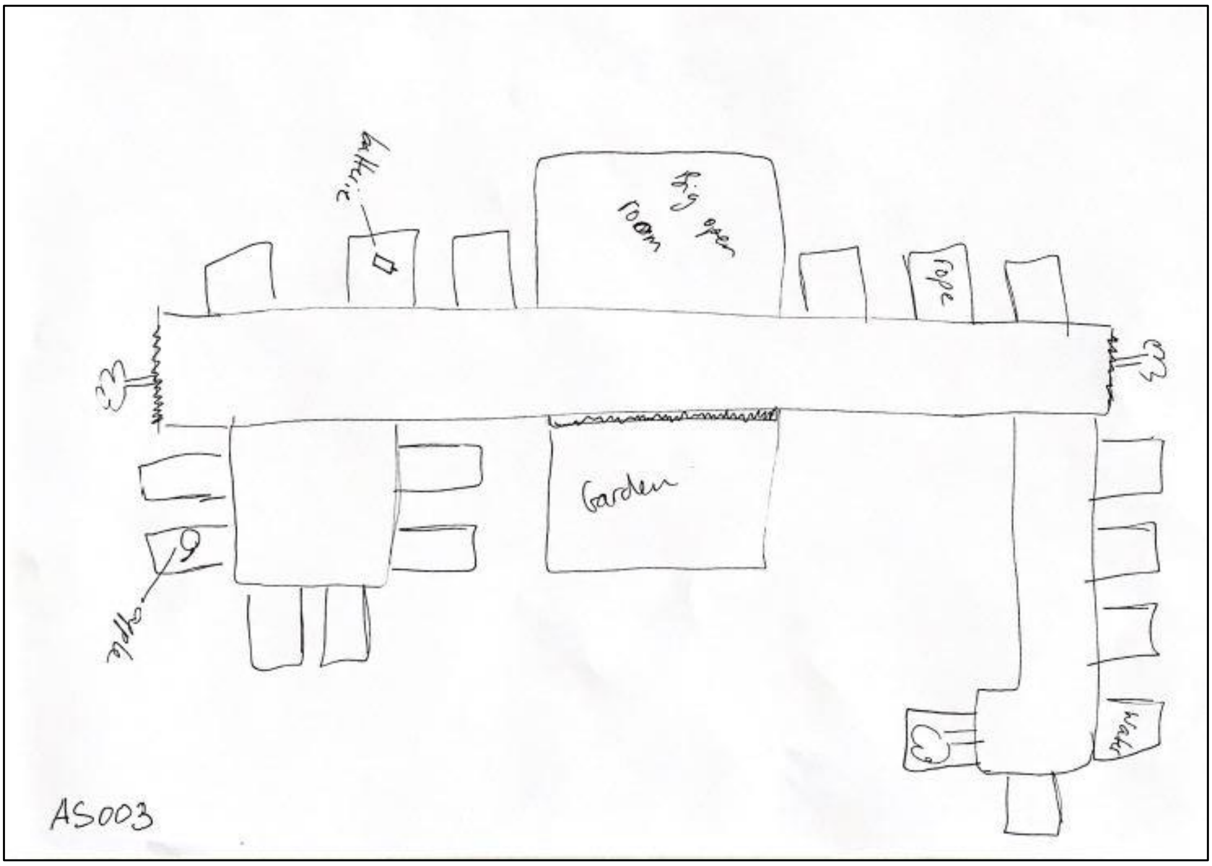
In this task you will be presented with 9 sketch maps. Your task is to draw the axial map representation (least and longest lines of sight) on top of each sketch map and to produce the underlying dual graph by representing each axial line as a node and the intersections between them as edges (see **Figure 3**). Please draw the axial line representation and dual graph in two different colours. In this task you will be guided by the issuer of the survey and it should take no longer than 10 minutes, so feel free to stop drawing axial maps once you have exceeded the time limit.

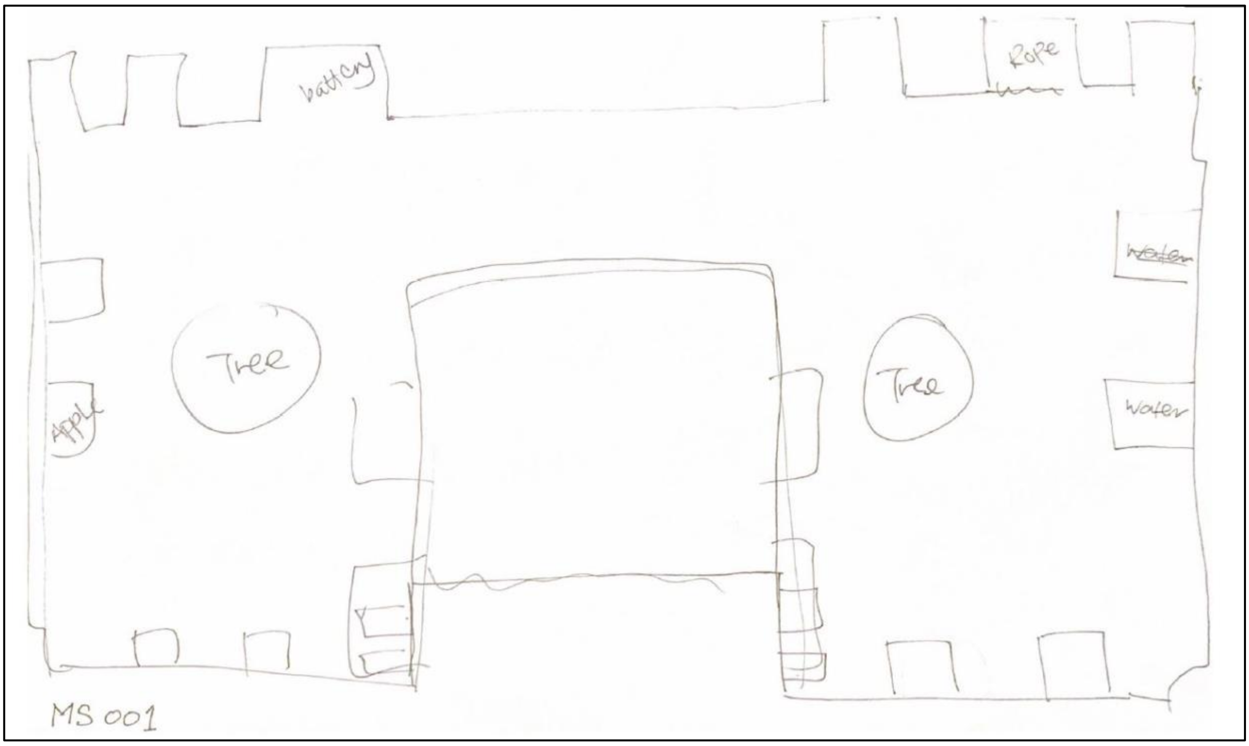
Figure 3: *Example of Axial Line Representation and Dual Graph*



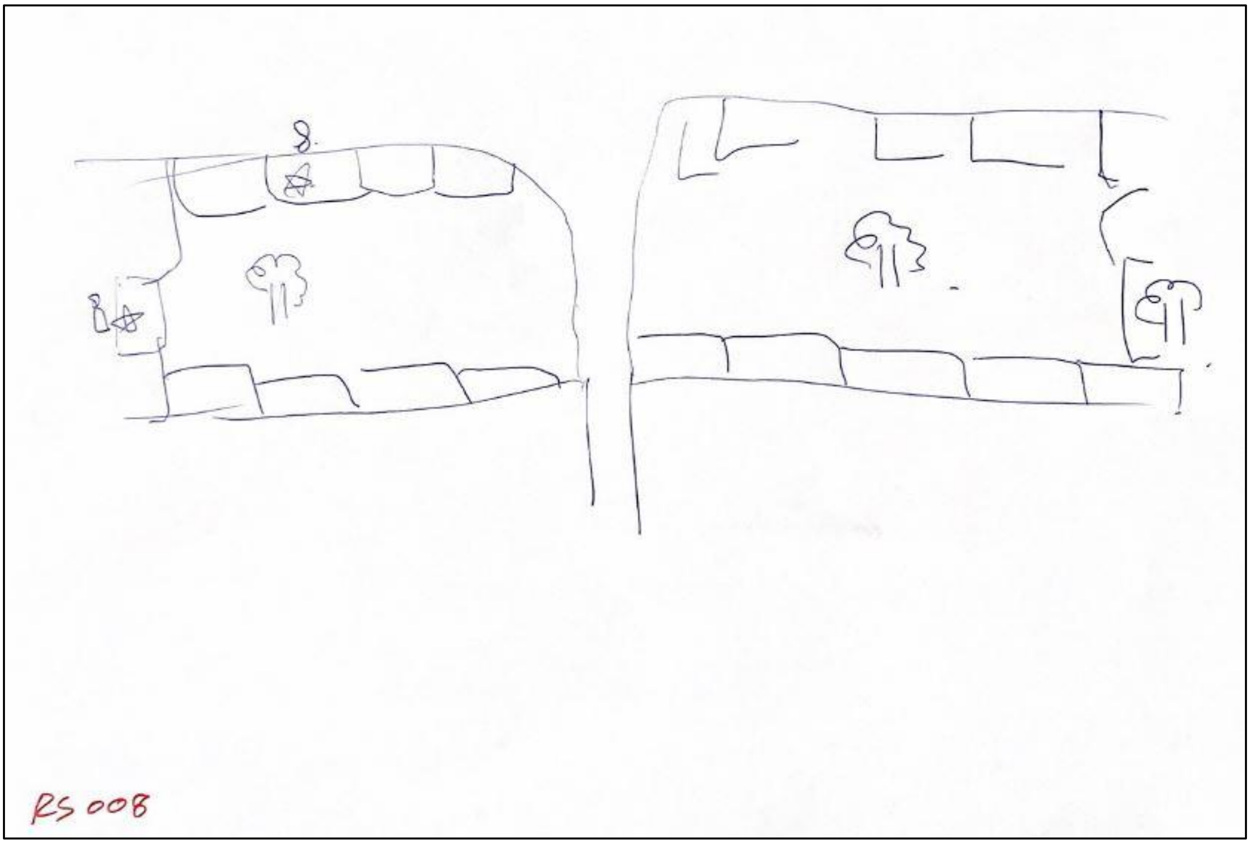
Sketch Maps for Task 3



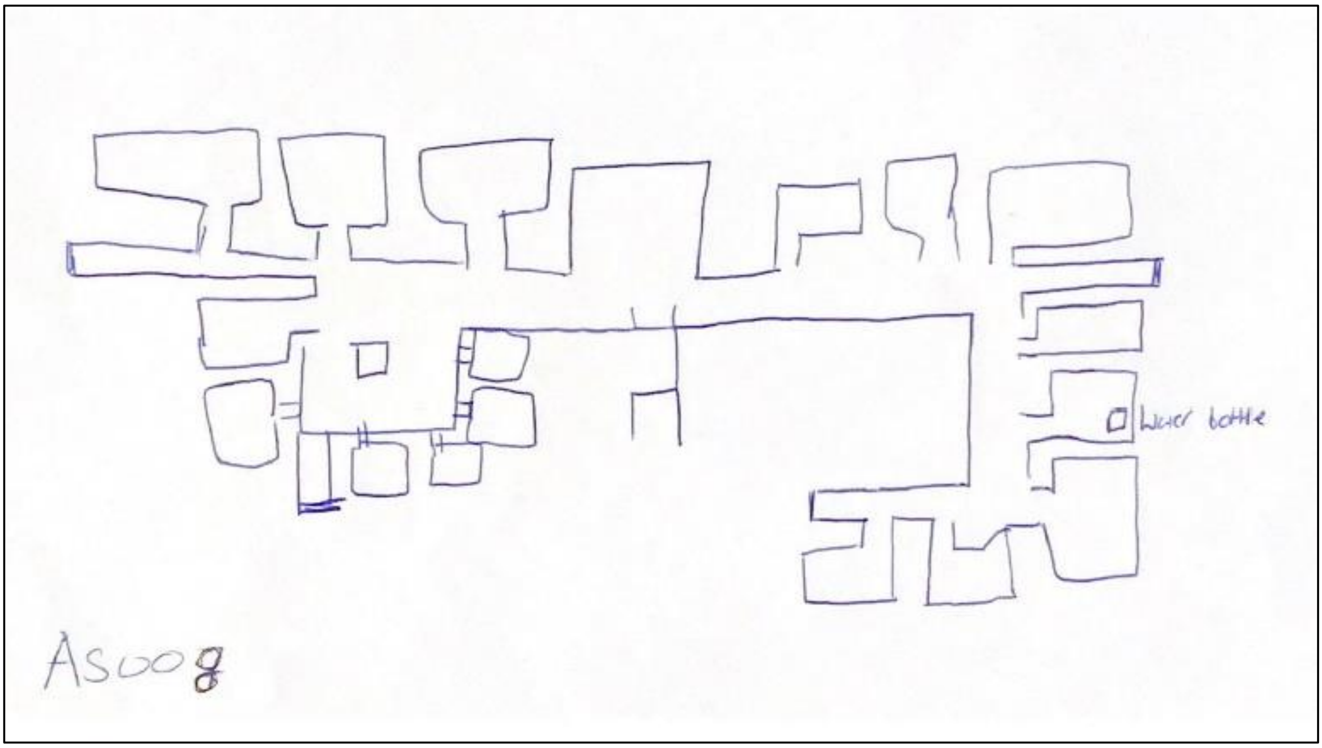


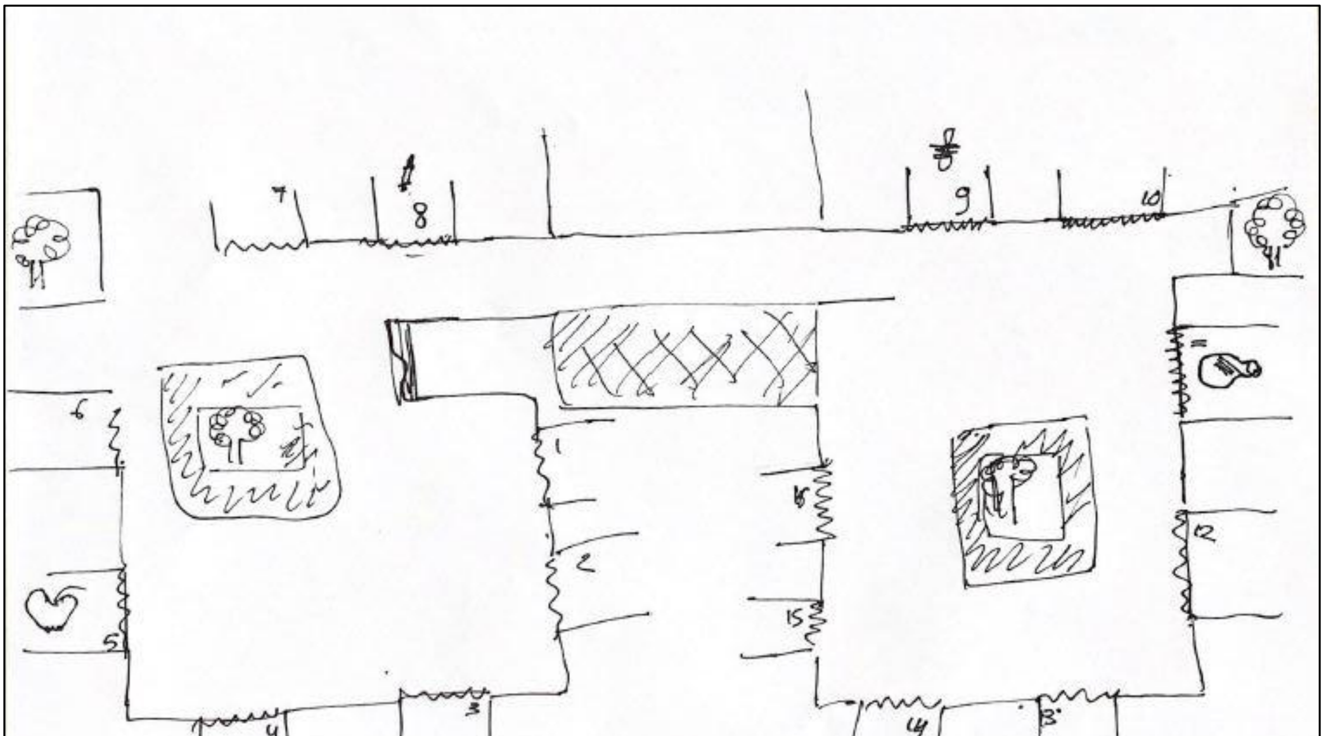


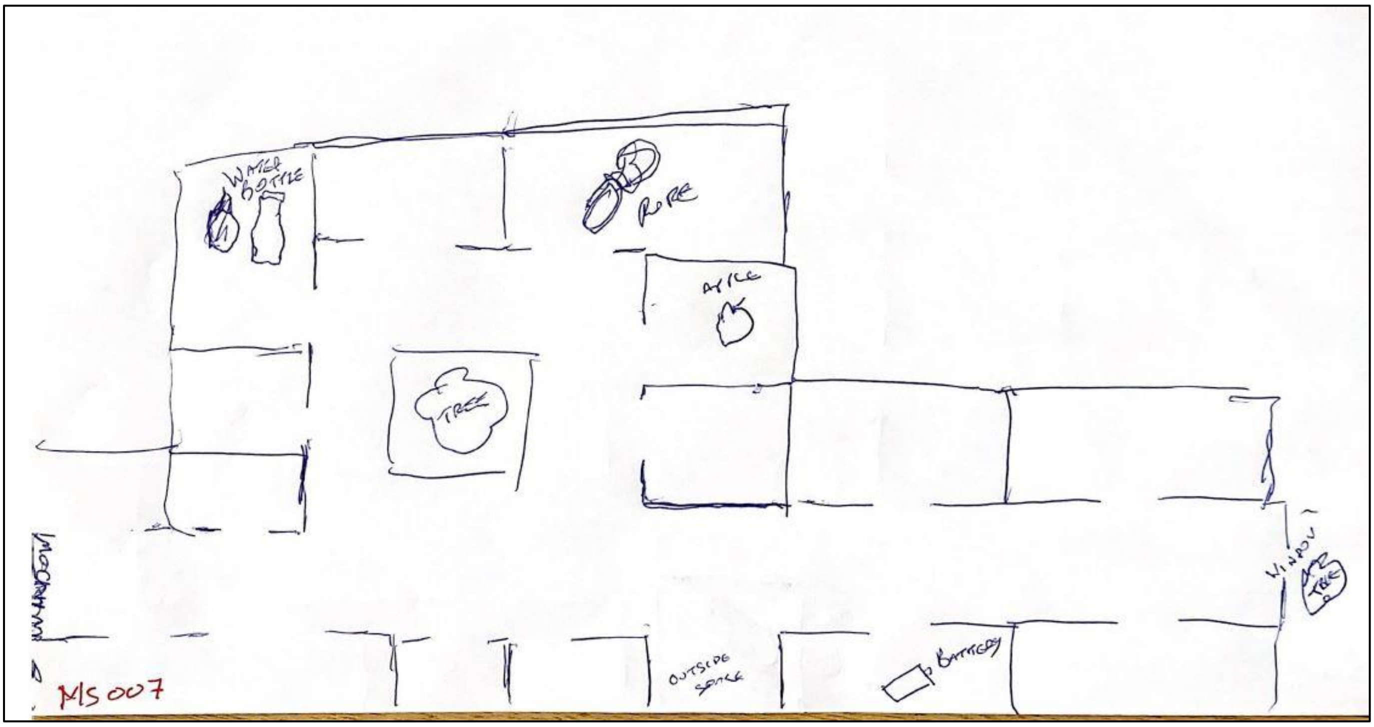
battery
Rope
Water
Tree
Tree
Apple
Water
MS 001

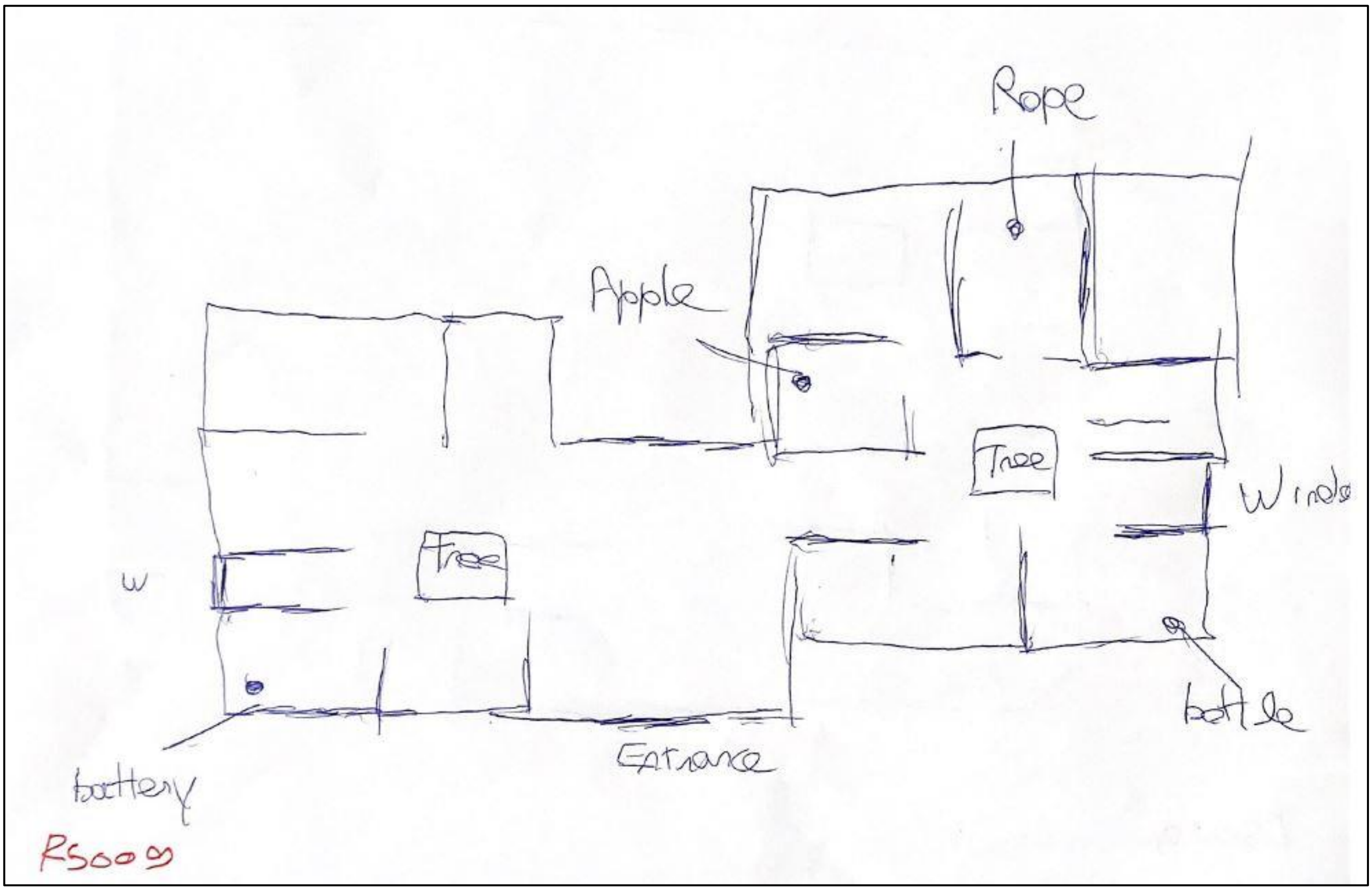


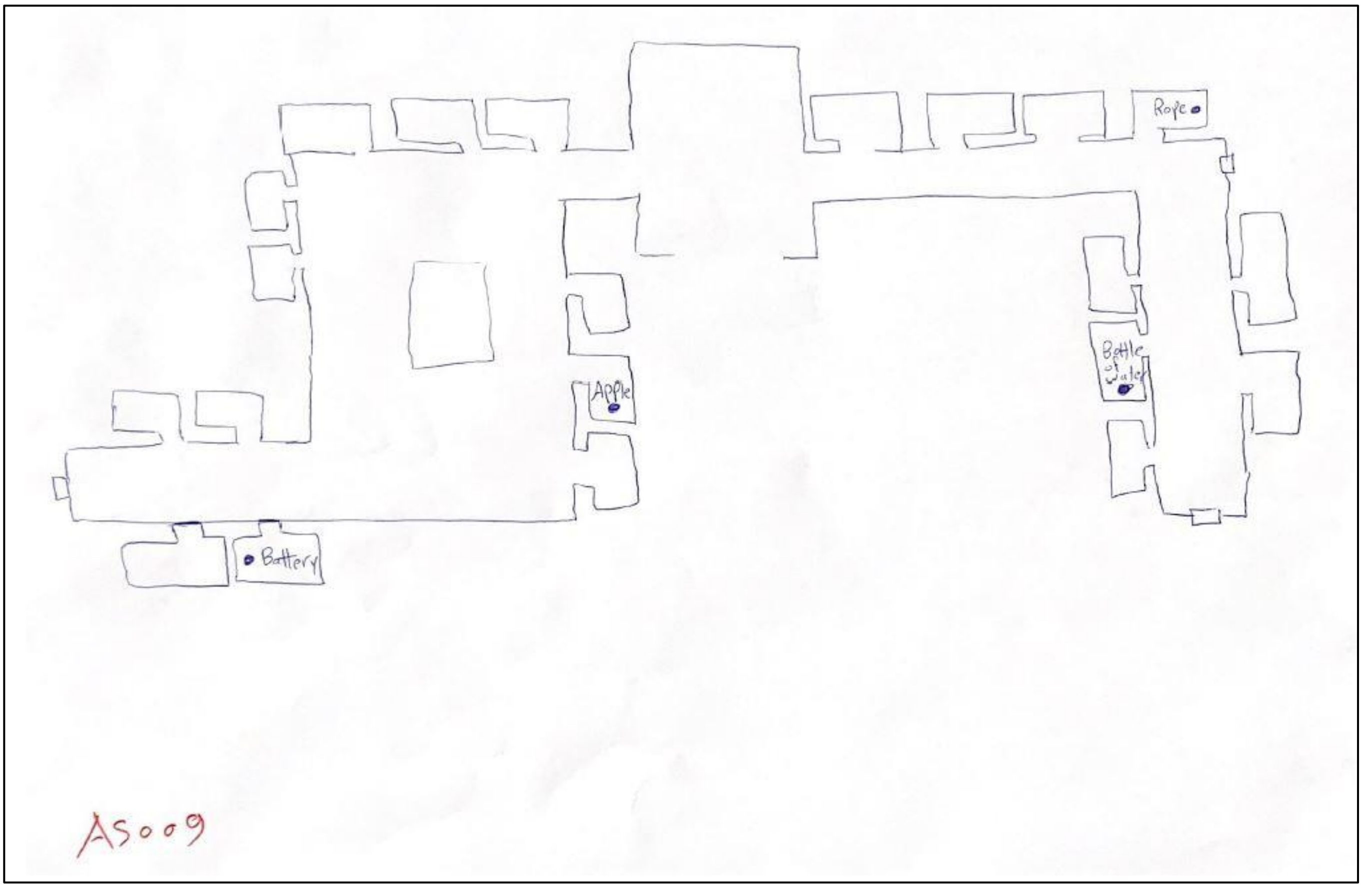
RS 008


Water bottle
AS008


7
8
9
10
6
1
2
5
4
3
15
14
13
12
MS003


WATER BOTTLE
ROPE
APPLE
TREE
WINDOW
TREE
OUTSIDE SPACE
BATTERY
MS007


Rope
Apple
Tree
Tree
Windo
w
bottle
battery
Entrance
RS009



Anonymized survey data is available upon request (bruce.timothy@hotmail.com)