The Influence of an Art Gallery's Spatial Layout on Human Attention to and Memory of Art Exhibits

Jakub Krukar

PhD

2015

The Influence of an Art Gallery's Spatial Layout on Human Attention to and Memory of Art Exhibits

Jakub Krukar

A thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of Doctor of Philosophy

Research undertaken in the Faculty of Engineering and Environment

May 2015

Kasi, w całości.

Dziękuję.

ABSTRACT

The spatial layout of a building can have a profound impact on our architectural experience. This notion is particularly important in the field of museum curation, where the spatial arrangement of walls and artworks serves as a means to (a) strengthen our focus on individual exhibits and (b) provide non-obvious linkages between otherwise separate works of art. From the cognitive viewpoint, two processes which can describe the relevant aspects of the visitor experience are visual attention and memory.

This thesis presents the results of 3 studies involving mobile eye-tracking and memory tests in a real-life task of unrestricted art gallery exploration. The collected data describing attention and memory of the gallery visitors is analysed with respect to the spatial arrangement of artworks. Methods developed within the architectural theory of Space Syntax serve to formalise, quantify and compare distinct aspects of their spatial layouts.

Results show that the location of individual works of art has a major impact on the dynamics and quantity of visual attention deployed to the artworks, as well as the memory of their content and of their spatial location. This spatial influence, in many instances, is proven to be more impactful than that of the content of the artworks. Some gallery arrangements amplify the impact of the studied spatial factors to a higher degree than others.

The results are discussed with respect to the distinct role played by the built environment (and, indirectly, by its designer) in our everyday cognitive experience. The thesis contributes to the field of museum curation by demonstrating how the aesthetic experience of museum visitors is affected by the decisions made by the curator. It also contributes to the fields of architecture and spatial cognition by demonstrating and quantifying the linkage between the formally described spatial layout and its impact on human cognitive processes.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF MODEL EQUATIONS

LIST OF PUBLICATIONS

Following works were published between the years 2012 and 2015 as part of research conducted within this thesis:

Krukar, J. & Conroy Dalton, R. (2012). Towards a Unified Model of Building Usability. In *22nd international association people-environment studies (iaps) conference*. Glasgow, UK

Krukar, J. & Conroy Dalton, R. (2013a). Spatial Predictors of Eye Movement in a Gallery Setting. In P. Kiefer, I. Giannopoulos, M. Raubal & M. Hegarty (Eds.), *Eye tracking for spatial research, proceedings of the 1st international workshop (in conjunction with cosit 2013)* (pp. 14–19). Scarborough, UK

Krukar, J. & Conroy Dalton, R. (2013b). Walk, Look, Remember: Art Galleries as Spaces Facilitating Memory. In Y. O. Kim, H. T. Park & K. W. Seo (Eds.), *Proceedings of ninth international space syntax symposium* (074:1–19). Seoul: Sejong University Press

Krukar, J. (2014a). Spontaneous Memory of Art Exhibitions: Explanations from Eye-Tracking. In C. Freksa, B. Nebel, M. Hegarty & T. Barkowsky (Eds.), *Spatial cognition 2014: poster presentations* (pp. 60–63). Bremen, Germany

Krukar, J. (2014b). Walk, Look, Remember: The Influence of the Gallery's Spatial Layout on Human Memory for an Art Exhibition. *Behavioural Sciences*, 4(3), 181–201

ACKNOWLEDGEMENTS

First and foremost I would like to express my gratitude to what many researchers associate with little more than a source of bureaucratic procedures, stressful hierarchical structures, and internal conflicts - my University institution. Over the 3 years I spent at Northumbria, I rarely met with anything else than kindness, willingness to help and true faith in the value of my work. I wish to thank for many enriching training opportunities and the generous financial support, but primarily for all the small input contributed by many of the Northumbria's staff and faculty members at various stages of this project.

One particular person who deserves the highest praise a student can give on the pages of his doctoral dissertation is Professor Ruth Conroy Dalton. Looking back at the recent years I am fairly confident there is no single thing my supervisor could have done any better. Ruth was calm when I was too stressed, strict when I was too relaxed, forgiving when the mistakes where less serious than they looked in my head, and demanding when there was no time to spare. Her flexibility and competence in supervising this work were an art in itself and I will be recalling her attitude at various turning points of this project with the highest appreciation. The scientific component of her supervision was exemplary, too. It amazes me how each single question in such an interdisciplinary project could have been answered by one person. Ruth's ability to draw connections between different disciplines and speak the language understandable to all of them is one of many of her skills I wish I can ever master to a similar degree. Me and this project could not have had a better supervisor and your competence, experience, and understanding truly impressed me. Thank you, Ruth.

Perhaps it is a family affair to nurture the growth of scientific progress but I also feel obliged to thank Ruth's husband, Sheep and her mother, Dru, for their tremendous help during my stay in Milton Keynes.

I also wish to thank Professor Chris Dorsett for many inspiring discussions and organisational help with arranging art gallery studies; Susi Bellamy for kindly providing her artworks and opening my eyes to many curatorial details I would surely miss on my own; the BALTIC Team, especially Hayley Duff, for their support so far-reaching that it has truly surprised me.

Thank you to Dan Montello and Mary Hegarty for many valuable advice and co-hosting ISCSI 2013 which changed my life and made me pursue the academic career.

Thank you to Adrian Chey for his help with the computer script and many other things which surrounded this discussion.

But above all, this thesis is about the *influence of space*. And as its author I feel obligated to acknowledge spaces which impacted the creation of this work the most. Many ideas here presented had either been conceived, discussed, or written down: at Emin Durak's balcony in Kurtulus, Istanbul (I probably would never apply for a PhD if not it), Pink Lane and Flat Caps cafes in Newcastle, public libraries in Vigo, Spain (the one with eyesight-draining lighting for focused work and the one with the most inspiring harbour view for those creative parts), and a completely random but modestly priced cafe in Paris, with napkins long enough to facilitate a comprehensive discussion on string-matching algorithms. The process of writing this thesis was, in itself, an exercise in the influence of space. Major parts of the written work were conducted in more than a dozen distinct indoor spaces and I thank all those responsible for their diversity.

DECLARATION

I declare that the work contained in this thesis has not been submitted for any other award and that it is all my own work. I also confirm that this work fully acknowledges opinions, ideas and contributions from the work of others.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by the Faculty Ethics Committee on 24.07.2012.

I declare that the Word Count of this Thesis is 86,835 words.

Name:

Signature:

Date:

Part I

INTRODUCTION

1

INTRODUCTION

We spend about 90% of our lives in various indoor environments (Fisk & Rosenfeld, 1997). But even before this was the case and long before creating buildings was an industry as well organised as it is today, humans struggled to make buildings suitable for the function they were serving (Rudofsky, 1987). We never create just 'buildings', but almost always 'schools', 'offices', 'apartments', 'churches', or 'market halls'. The function of a building is inextricably linked with it from the very first stages of the planning process. This is not equal to saying that we do not allow the thought of using the same building for different functions. But when such a thing happens, we talk of 're-usage', 're-adaptation', and 'renovation'. The function a building serves is the most defining element of its lifecycle, right after its creation and demolition.

Each of the many functions we give to buildings is an art and science of its own. Entire companies, research fields and industry standards were developed around—often narrow—functional niches that we give to our spaces. It would not be an overstatement to say that after ensuring it is safe and economically viable, warranting that the new building will be able to serve its function well should be the next priority of its developer[1].

There are some minimal requirements we assume to be 'sufficient' for any building. They must protect us from the external elements, ideally provide a hygienic, comfortable environment, and organise space. And it is this last factor that tends to be specific for individual building types, historical periods, and cultural spheres. After all, 'it is this ordering of space that is the purpose of building, not the physical object itself. The physical object is the means to the end. (...) Buildings are not just objects, but transformations of space through objects.' (Hillier & Hanson, 1984, p. 1). While the progress of civilisation allowed us to take the first two aspects for granted, the organisation of space is a factor where the actual challenge of designing successful buildings lies. In the developed and developing countries, it is rarely an issue for a newly created building to be unsafe, unclean, unstable, or faulty in any 'physical' way. And yet, it is not uncommon to experience dysfunctional buildings (Conroy Dalton, Kuliga

1 Which not always is true in the economic arrangement where the priorities of building's developers rarely match the priorities of their future users.

& Hölscher, 2013). This dysfunctionality can arise from many different factors and largely depends on the function we expect the space to perform. A hospital for instance is generally expected to provide rest (to patients) and efficient working environment (to the staff), while a library serves as a large book storage and a place for various intellectual activities. Each building type can be therefore dysfunctional in its own way.

This thesis will focus on the cognitive aspect - on how buildings make us attend to, think about, and experience things in space. In this respect, an architectural space which is created to facilitate our attention and thinking in a particular manner, but fails to do so, would be considered dysfunctional. And to notice that buildings do that at all is not new. If architects were not aware of this fact, they would not create experiences which are so memorable. We would not marvel at the play of lights in Istanbul's Hagia Sophia, the enormous naves of Rome's basilicas, or the aesthetically pleasing order of Parthenon in Athens. Architects design the experience as much as they design the brick walls. By designing the walls they influence how we move through, observe, and think in that space. They influence our experience.

Over generations, architects learnt how to do it well, if not better than the Ancients. Buildings of Alvar Aalto or Frank Lloyd Wright are today no less popular attractions for tourists willing to experience their interior than the world's oldest surviving human-made structures. And yet, one of the main reasons for why they are so popular is because they are unique. Singled out, visionary creations of genius minds. Of few special individuals, out of hundreds or thousands who left the same architectural school, had the same knowledge, tools and methods to tackle the design challenges of their times. Good design which merges form and function in a single object giving equal justice to both is not common. And the large majority of all other buildings that we visit everyday can be considered average - we do not tend to complain about them that often, but we do not marvel at each single visit to a new structure, either. And once complains occur, they rarely mention purely 'engineering' flaws . It is not uncommon, however, to hear an opinion that a building is 'sad' and 'depressing', 'confusing' and 'weird', or simply 'dull' and 'boring'.

With the relatively recent[2] development of experimental psychology (Wundt, 1874) we are learning to better understand how humans perceive, conceptualise and mentally react to their external environment. Since 1967 (Neisser, 1967) we are using the term 'cognition' to explain how various mental processes important to our everyday lives relate to each other. We are slowly building a more complete understanding of the human mind and the rules which guide its interaction with the world. And just as knowing how the preferred

2 Compared to the history of architecture.

way of walking allows us to design corridors wide enough for the human body, understanding how we perceive or think, allows us to design spaces in which we perceive and think in a more relevant manner. What has been termed 'cognitive engineering' in the studies of Human-Machine Interaction[3], has been present in architecture for centuries, but rarely formalised or studied in a rigorous manner.

Most recent advancements in research technology have expanded our possibilities of studying this aspect of the spatial experience that we probably care about the most - the visual aspect. Mobile eye-tracking devices make it possible to precisely measure how people explore their visual environment and what they attend to. Understanding the nature of this process can enable new generations of architects to incorporate this knowledge into their design thinking. And as a result - improve the quality and suitability of our everyday visual experience inside buildings.

For this change to be possible, at least one common assumption must be shared by researchers and practitioners: that human architectural experience is an interaction. It is an interaction between the will of an unpredictable visitor who is free to explore and think about space in any way desired and the space itself, which can guide, attract, or deter - and thus have an *influence* on human behaviour and cognition through the way it is organised.

Understanding this interaction requires at least two sets of methods. One is needed to understand and measure the cognitive experience a user might have while walking through a 3-dimensional environment, which not always can be directly inferred from the results of laboratory-based studies of single isolated cognitive processes. Additionally, a second set of methods is required to measure the organisation of space. Space is not a physical entity - it cannot be touched, smelled, or even seen. It is the layout of walls inside buildings which gives us some preliminary information on how space is structured inside, but the concept of space cannot be studied directly [4]. This thesis will build on the methods known from experimental psychology to formalise and measure the concept of experience and on Space Syntax theory in respect to the architectural aspect. This is not a new connection and we will later review previous attempts made in between these two fields by other researchers.

It must be emphasised at this point that this way of thinking about architectural design process does not contradict it as a creative endeavour. As Davies (1957) described it, "[k]nowledge is the raw material for design. (...) It is not a substitute of architectural imagination: but it is necessary for the effective exercise of imagination and skill in design. Inadequate knowledge handicaps and trammels

3 The most generic definition of which can be designing with the consideration of how users think; (see e.g. Norman, 1986).
4 Similarly to the concept of cognition.

the architect, limits the achievements of even the most creative and depresses the general level of design.”. This thesis will aim to expand the state of knowledge in an understudied, still developing field where new research technology only recently provided the sufficient precision of measurement. It will aim to explain the influence of space on the level of individual interactions between the user and the space, within a narrow subset of selected cognitive processes and spatial contexts.

The particular context being here studied is a visit to an art gallery. This setting has been selected for multiple reasons which later will be described in detail. At this point it is sufficient to say that art galleries are a unique type of environment where the impact of space on human cognition can be noticed even by a naive interested observer - it can be done by simply asking a friend how s/he felt about a visit to two separate art exhibitions. Equally, an art gallery is a building type for which a large proportion of the effort spent during the design phase goes into the consideration of the spatial arrangement of objects creating the locus of interaction between the art and the human mind.

1.1 RESEARCH QUESTION AND CONTRIBUTION TO KNOWLEDGE

This thesis aims to answer the question of whether the spatial arrangement of an art gallery has an effect on the visitor experience. The main hypothesis states that the influence of artwork’s spatial location is often larger than that of the artwork’s content and that it can be used to partially predict the cognitive impact of an exhibition. This is a particularly fascinating problem considering that very often thematically-grouped, and visually similar modern artworks are displayed in aesthetically-neutral, ‘white-cube’ galleries. This thesis will aim to explain whether such a situation enhances the potential impact of the spatial arrangement, shifting the decision about the outcome of the visitor-artwork interaction from the artist creating the work, to the curator responsible for its arrangement.

The current work aims to make original contributions to the following disciplines:

- *Visitor Studies;* by expanding the existing state of knowledge on visitor behaviour (and, most importantly, cognition) during the interaction with the art, in a real-life gallery settings;
- *Curatorial Theory;* by empirically verifying which aspects of the curatorial design activity bear the strongest impact on the final perception and experience of the exhibition;
- *Applied Cognitive Psychology and Spatial Cognition;* by investigating cognitive processes contributing to the creation of the subjective experience inside an art exhibition, including processes

responsible for the dissimilar understanding of space across individuals;

- *Architectural Theory of Space Syntax;* by empirically linking formally defined characteristics of space (used for studying and planning spatial arrangements) to their influence on the human experience.

1.2 METHODOLOGICAL APPROACH

Research approach employed in this thesis (and consequently all methods used) comes from the positivist perspective. Therefore it is further assumed that phenomena such as the visual experience can be quantified and described with numbers. There are two assumptions that this thesis makes in relation to the consequences and limitations of using such a methodological approach, namely:

1. That these numbers need not represent maximum engagement with the art, or maximum efficiency of the visitor-artwork interaction. The relevant numbers can often refer to the variance of the experience, as the goal of art is often to stimulate novel, unique experiences.

2. That there are many of such 'optimal' numbers and it should be at the discretion of the curator to choose what type of experience the exhibition shall embrace. For instance, the curator might wish to create a highly dynamic visual experience, where the visitors quickly switch their engagement between multiple exhibits. This would be considered undesired within many study-based exhibition-planning guidelines in the past but is nevertheless measurable with methods presented in this thesis. The current work will not aim to put value on more or less 'desired' types of experience, but rather to classify and describe their variety.

Part II

LITERATURE REVIEW

2

DEFINING A 'USABLE' MUSEUM

The popularity of physical art galleries prevails despite the ease of access to all kinds of information—including art—from the comfort of one's own living room, through the internet. This fact suggests that exhibiting art in a physical setting can bear a unique function, difficult or impossible to fulfil with alternative means of displaying. In this chapter we will look at how this function is understood by artists and curators. We will briefly look at how art exhibiting changed throughout history, and ask the question of how museums can be evaluated for serving their function well. This chapter will introduce the role of museums in the modern society and try to establish how a museum exhibition can be 'usable'. By this term, architecture and other disciplines define how well a physical object or setting is suited to perform the function it was designed for.

2.1 HISTORICAL CONTEXT

The first known mention of the word 'museum' occurred in the Hellenistic Alexandria, however art collecting as we know it became popular much later - in the Italian Renaissance (Pevsner, 1976; Newhouse, 2005). The fifteenth century saw pieces of antique statuary displayed in courtyards, loggias and around private properties; but the very first place specially designed for presenting art was a square cloister in the Vatican, supposedly designed around the year 1508 (Pevsner, 1976). In 1543 the word 'musaeum' officially appears on a building containing the collections of Paolo Giovio in Como. This way, contrary to the earlier meaning of the term, instead of a place designated to celebrate knowledge and culture (like a library in the ancient Alexandria), 'museum' gained its current meaning. It also quickly became synonymous to a 'gallery' due to the fact museums were often designed as long, narrow rooms (Pevsner, 1976). As museums/galleries grew in number and diversity throughout France and Italy, by the late seventeenth century they became an almost inseparable section of every major palace. Soon, with the Enlightenment and its trend for specialisation, museums were divided into rooms dedicated to various objects. Since then, not only ancient sculptures, but also paintings, or natural history artefacts were considered worthy of displaying in a manner different to furniture (Pevsner, 1976; Newhouse, 2005). Following the ideas of the Enlightenment, a belief that these collections should be made publicly available became more widespread, and as a consequence, the first known dedicated museum programme was printed in 1742 in Dresden.

Openness to the wider audience brought the question of how to group and present objects belonging to discrete categories. Throughout the early history of museums, two main principles where used to present artefacts in a meaningful order. First was, what Pevsner (1976) describes as the concept of 'iconography' - statues and paintings of gods would be located in the upper floor, whereas rulers (as less significant) were placed in a lower—but still central—space. Afterwards, an alternative approach became popular, with the chronology of style being the main principle guiding exhibition arrangement (Pevsner, 1976).

Nevertheless, it was in Munich in the nineteenth century where the way of thinking about exhibiting as an object-centred activity was questioned. Pevsner (1976) describes the following discourse between the architect Leo von Klenze who was appointed by Prince Ludwig to design the Glyptothek, and Johann Martin Wagner - an artist, archeologist and a friend of the prince. Glyptothek was envisioned as the greatest ancient sculpture museum of its times. Wagner suggested a very modest decor, as his standpoint was that 'any ornament, anything gay in colour and glittering does damage to works of ideal art' and that 'if you visit a collection of ancient sculpture you go because of the ancient sculpture', so consequently 'one recognizes the merit and talent of an architect by the strict coincidence of a building with its function' (Pevsner, 1976, p. 124). He made it explicit that 'it is my principle ... to prefer utility to beauty in case the two cannot be united' and even made a statement that 'polished marble walls and floors are an attraction only for the common rabble' (Pevsner, 1976, p. 126). Leo von Klenze had quite the opposite approach: 'a museum is not a place for artists' training, (...) but a place in which to show a number of treasures of art to all kinds of visitors in a manner to be worthy of the objects and to create pleasure in them' (Pevsner, 1976, p. 126). The prince agreed with von Klenze and the ideas aiming at delivering 'a grandiose architectural effect of whole parts of the building' (Pevsner, 1976, p. 126) were decisive. This dispute is an excellent example of how the two postulates underpinning the museums' primary function: to give justice to the works collected, and to display them in a manner somewhat attractive to a viewer, expanded to a debate still ongoing in modern museology.

2.2 OBJECT-CENTRED AND VIEWER-CENTRED CURATION

One voice in this debate is Michael Baxandall's text (Baxandall, 1991) which takes a closer look at the viewers' goals in art museums. As he points out, the main attractor for visiting museums, which makes the experience superior (or at least different) to reading a book about the subject of the exhibition, is the visual character of the experience (see also Baxandall, 1991; Alpers, 1991; Newhouse, 2005). As a result, the museum experience arises not only from the content of the objects (since what the artist originally intended to say might be interpreted or misinterpreted in many different ways), but also from the way of displaying them (Gurian, 1991; Baxandall, 1991; Alpers, 1991; Newhouse, 2005). Baxandall (1991) then notices that it is impossible to exhibit objects in a neutral way, without enforcing another meaning upon them, and that every curatorial decision contributes to a 'label' imposed on an object. This label consists of such interventions as lighting, placement, or segregation of objects and it bears explanatory ideas of the curator, which are often different from those intended by the artist (Baxandall, 1991; Newhouse, 2005). The final interpretation of art can also vary depending on the viewer's individual understanding of those relations. The interplay between three agents contributing to the final meaning—the artist, the curator and the viewer—is then clearly apparent. Consequently, taking into consideration the active presence of the viewer in the process of exhibiting and interpreting art is crucial (Gurian, 1991; Baxandall, 1991; Alpers, 1991).

This consideration might be realised differently, depending on the particular museum's context (Screven, 1976). To bring these opinions together, Vergo (1989b) suggested that one thing important to all curators, artists or exhibition planners, regardless of the theme and the context, is making sure that the visitor will 'get something out of it' (Vergo, 1989b, p. 46). According to him, this can happen in two ways. Firstly, there is the 'aesthetic' view, suggesting that objects are best shown in isolation and that the whole interpretation can only emerge from the visitor's experience alone ('object-centred curation'). Displaying objects in a way glorifying their presence is a craft in itself. Alpers (1991) calls this manner of displaying, which turns objects into attention-drawing artefacts, 'the museum effect'. Secondly, the 'contextual' approach proposes that each museum can (and perhaps should) be a set of labels and instructions barely illustrated by examples of prominent artworks ('viewer-centred curation'; Vergo, 1989b).

Concurrently, Vergo (1989b) notes that there is no method for predesigning an exact experience in a complex exhibition. This is because not only every viewer differs significantly from any other, but also every single visit of the same person will inevitably vary from any other visit. Therefore there is a limit to the influence a curator can consciously have on the visitor experience, but this influence is inevitable (Newhouse, 2005). This is true for both approaches as each of them involves intentional curatorial interventions which should be planned. Most importantly however, regardless of what the 'correct' approach is, none of them will let the visitors to 'get something out of it' without ensuring that the initial visual (or tactile, or sonic for other types of art) encounter between the person and the object has taken place. Thus the recurring call for greater consideration of display methods in curation (Alpers, 1991; Baxandall, 1991; C. S. Smith, 1989; B. L. Taylor, 2010; Vergo, 1989b; P. Wright, 1989; Gurian, 1991; Newhouse, 2005).

2.3 HOW TO EVALUATE ART GALLERIES?

Weil (2004) goes back to Joseph Nobel's *Museum Manifesto* (1970) to remind us of five basic tasks of every museum: to *collect*, *conserve*, *study*, *interpret*, and to *exhibit*. As he suggests, these roles have since changed. He claims museums nowadays rather ought to: *preserve*, *study*, and *communicate*. The last function is considered a combination of Nobel's *interpretation* and *exhibiting*.

This change of paradigm does not necessarily occur instantly but is rather a realisation of the fact that has gradually become obvious (Weil, 2004). It is the result of accepting that there is no way of exhibiting that would be deprived of *interpretation* and that every attempt of presenting anything in the museum context is inextricably linked with (explicit or implicit) presumptions. Weil (2004) continues the argument to point that if it is the museum's responsibility to *communicate* anything, it should be relatively easy to assess the museum's performance on that task as soon as we realise what the intended message is. As he then notes, however, this way of thinking about museum's role is in fact very passive and does not consider diverse agendas with which visitors arrive to museums, or their own active interpretations of the exhibited content. A new paradigm should, in his mind, take all this into consideration to avoid thinking about exhibiting as a one-way process. This, in his view, is yet to emerge (Weil, 2004).

In relation to Weil (2004), L. C. Roberts (2004) takes a closer look at the interpretative role of museums with the focus on art galleries. She quotes John Walker—a curator and the director of the National Gallery of Art in the United States between 1956 and 1969—who emphasised the role of art gallery in presenting masterpieces with no interference, so that it can communicate with the viewer on its own. One view contrary to that is that the lack of contextual background might leave an uninformed viewer confused, and far away from the pleasure of fully experiencing the artwork as it was intended to. As she writes, the *Visitor Studies* movement which followed this argument, aimed at evaluating the activities providing the context inside museums. However, as early as in the mid 1950s, it was realised that such an evaluation might never be able to conclude about the effectiveness of museums' educating activities. L. C. Roberts (2004) draws a conclusion that museums' interpretative role is not to educate in a top-down manner. But rather to allow the visitors to find their own connections between the presented works, based on their own agendas and value systems. Roberts calls this 'a shift in focus from the *object* to the *process of looking at it*' (L. C. Roberts, 2004, p.220). Accepting that visitors will look at exhibits from very diverse viewpoints (both literally and metaphorically) bears a responsibility of allowing them to set their own priorities and contributing to the experience in a manner not less creative than that of the curator's or the artist's. As she concludes, the role of modern museums is to empower visitors to take part in the critical dialogue about what and how exhibits are presented (L. C. Roberts, 2004).

2.4 reminder: 'to get something out of it'

This section of the Literature Review has presented publications which make it evident that:

1. There is a recognition of the visitor's central position in the role museums play in the modern society. Times when museums served barely to showcase private collections of the noblemen have passed. Museums are built for their users and as such should consider their goals and expectations with greater care.

2. This being said, the need for evaluation of museums' performance is an emerging issue. This has been noted not only by the academic community (Vergo, 1989a; Andersen, 2004), but also by the public policy makers, requiring museum institutions to self-assess their performance (Arnsdorf, 2010).

3. Yet, it is not clear which aspects of the visitor experience are in fact deserving deeper consideration. While science museums can more clearly define their goal as 'informal learning', this issue is more problematic in the context of art galleries. One general point that seems to be recursive across different publications is what Vergo (1989b) summarised as 'getting something out of it'. This thesis assumes that for this to happen, museums must ensure they are displaying their content in a cognitively accessible manner. We will focus on the earliest stages of the visitor-artwork interaction, as this—contrary to the viewer's background, preferences, and expectations—is the factor which can be facilitated through the means of curation.

The next section of the Literature Review will look at studies which tried to assess whether museum visitors indeed do 'get anything out of it'. Understanding what is already known and identifying gaps in our knowledge on human behaviour and cognition inside art galleries will be crucial for guiding the scope of this thesis.

3

RELEVANT COGNITIVE PROCESSES

We concluded the previous section by identifying that for a museum to fulfil its function the visitors must 'get something of out' their visits. To formally operationalise this phrase, it can be said that the museum must ensure its visitors are leaving with a different mindset (or state of knowledge) from the one they entered with. How this aim is achieved is a separate issue. It might happen through empowering the visitors to explore, challenge and critique (e.g. L. C. Roberts, 2004), or by teaching them new information (e.g. Falk & Dierking, 2004). Embracing one of those fundamentally different approaches will most likely depend on the type of the institution, its goals, and the overall context of the expected experience. But in each case, the primary priority for any museum should be to ensure that the visitors are able to cognitively engage with the displayed exhibits in a meaningful manner.

In Psychology, the occurrence of such interaction can be described by the processes of *attention* and *memory*. In Spatial Cognition—a branch of Psychology concentrated on the processes of forming and using mental representations of space, as well as embedding mental representations in space—*attention* and *memory* contribute to what is known as *online* and *offline* aspects of cognition (e.g. Waller & Nadel, 2013). *Online* processes are those, which allow us to operate in the dynamically changing spatial surrounding, for example by keeping track of where we are in space, estimating distances, and recognising objects in that space. *Offline* processes on the contrary, are those which allow us to encode, store, retrieve and subsequently use spatial information over longer time periods. This makes it possible for humans to better navigate through a previously visited environment, or to draw a map of it based on the mental representation of that environment (so called 'cognitive map'; Tolman, 1948). The crucial principle underlying the theoretical approach of this thesis is that spontaneously encoded spatial information is automatically embedded in our knowledge structures and becomes a significant component of our retrospective experience. Accounting for both *online* and *offline* processes makes it possible to consider how this phenomena forms during the exploration of the environment, and which of its aspects are best preserved.

Despite the fact memorisation is not the predominant goal of visiting exhibitions, memory performance might serve as a benchmark indicating whether deeper cognitive processing of an object took place. The following section will provide evidence for this statement and justify the importance of *attention* and *memory* as measures of human cognitive engagement within art exhibititons. This thesis treats mainly with the visual aspect of the museum experience, as it is considered to be the most important (Baxandall, 1991). For this reason the review of studies on human *attention* has been limited to *visual attention*. It must be noted however, that the non-visual aspect of the museum experience is attracting a great deal of interest (Levent & Pascual-Leone, 2013b) and need not to be ignored in more generic discussions. Especially if to consider memory as a process of binding features across senses (J. Ward, 2013). It is, however beyond the scope of this thesis to directly account for other senses influencing the visitor experience.

3.1 VISUAL ATTENTION

3.1.1 *What is Attention?*

Contrary to the common assumption, attention by definition *is not* the state of 'exceptional mental effort' deployed to a particular stimulus (for a more detailed discussion on this, see Pashler (1998), who reviewed some central notions of common (mis)understanding of attention). The academic understanding of this process is different from its commonly assumed characteristic. Yet, there is a large variability in its classifications and definitions. Scholl (2001) for example, defines *attention* as the collective name for several various types of selective processing. Chun and Turk-Browne (2007) put emphasis on the distinction between *resources* and *selection*—*to attend* means to *select* what deserves attending, to *allocate* processing resources to this task or object, and to *prevent* competitive stimuli from using these resources (Chun & Turk-Browne, 2007). According to this view, attention is therefore more of a 'filtering' mechanism than a 'focusing' one.

Yet different aspects gain importance when we consider attentional processes engaged in the everyday exploration of the built environment (an extensive review of which is provided by A. Johnson (2011)). Attending to discrete elements of our surrounding can be the result of either: voluntary deployment of one's own effort into looking at it (*endogenous attention*); or of an external stimuli (such as loud noise or a blinking light) automatically apprehending human awareness (Jonides, 1981, *exogenous attention*;). Another important distinction is made between directing the senses (e.g. the eye movement) towards a stimulus (*overt attention*), and mentally focusing on a larger number of potential stimuli, e.g. in the peripheral field of view (*covert atten-*

tion) (R. D. Wright & Ward, 2008; Land & Tatler, 2009). Hence, covert attention can be deployed before overt attention but never the other way around (Land & Tatler, 2009). The focus of this work is on *overt endogenous visual attention*.

Eye movement is indicative of the deployment of overt visual attention (Buswell, 1935; Holmqvist et al., 2011) and has been used for this purpose in a plethora of studies (e.g. Buswell, 1935; Heidenreich & Turano, 2010; Henderson, 2008; Hollands, Patla & Vickers, 2002; Locher, Krupinski, Mello-Thoms & Nodine, 2007; B. W. Tatler, Baddeley & Gilchrist, 2005; Locher et al., 2007; Land & Tatler, 2009). However, some questions about the nature of the process remain open. Especially problematic is defining the point at which the perceptual stimuli is converted into a cognitive concept further incorporated into the existing knowledge framework. It is therefore unclear when visual attention is a location-based process, scanning the visual field like a 'spotlight' (Posner, Snyder & Davidson, 1980), and when it classifies the perceived stimuli into objects of some level of complexity (Scholl, 2001). The latest findings show the latter might be the case at much earlier stages of processing than it was previously assumed (B. W. Tatler, Gilchrist & Land, 2005). As it is beyond the scope of this thesis to investigate the basic nature of the process, this work will refer to 'visual attention' as to the equivalent of where a person looked. For simplicity, it will often treat the terms of 'visual attention' and 'eye movement' interchangeably, although it must not be ignored that (a) the latter is only to some extent indicative of the former, and (b) that overt endogenous visual attention is never the sole source of information acquired from our surrounding at any given moment.

3.1.2 *The Nature of Eye Movement*

To facilitate the uptake of visual information from the surrounding, our eyes consistently scan the environment. Rapid movement of deploying overt attention onto an object is called a *saccade*. A period of inhibited movement of the eye, or a pause in motion, is a *fixation*. Fixations can last from 25 ms to 200 ms, depending on the task (Ashcraft & Radvansky, 2010).

Saccades constitute 10% of all movement and during that time the observer remains practically blind. Information is only absorbed during fixations (Holmqvist et al., 2011). When a very long saccade is required to fixate on an object distant from the current point of fixation, usually a single, long, imprecise saccade is followed by a more accurate, correcting one (Land & Tatler, 2009).

Subsequent fixations and linking them saccades are described as a *scanpath*. When eye-movement is analysed in respect to discrete objects, a single scanpath within an object can be described as a *dwell* (Holmqvist et al., 2011). Figure 3.1 visualises a sample dwell.

Figure 3.1.: A sample 'dwell'.

Other elements of the oculomotor behaviour have also been classified and are comprehensively reviewed by Holmqvist et al. (2011). It is important to remember however, that they are all sub-parts of a larger system, the primary goal of which is to gather visual information required for current actions (Land & Tatler, 2009). For this reason, the exact definitions of these sub-parts and their operationalisations vary across literature - Holmqvist et al. (2011) describe about 120 of them. Their usage depends on the research aims, and equally often on the technological limitations of the—still imperfect—eye-tracking devices. The specific measures used in this thesis, the rationale for their choice, and its technological aspect, will be described in detail in the Method Overview section (p. 103). But the reader should be aware that this study's understanding of *visual attention* and its operationalisation is limited (similarly to other works on the subject) and by no means comprehensively represents this complex aspect of human cognition.

3.1.3 *The Role of Visual Attention During Exploration of the Environment*

The main aim of this work is to identify environmental correlates of human cognition in the gallery setting. Where each person looks is a combination of many, mostly uncontrollable factors, such as personal preferences. The rationale behind the current thesis is the assumption that those variable factors coexist with more basic principles guiding visual attention, which form repeatable patterns. Above all, visual attention evolved to allow us effectively and safely move around our environments. As such, some patterns of the oculomotor behaviour, by principle, must be present at the level of the entire species. This section provides a review of works aimed at identifying such patterns of human eye movement across diverse contexts of our everyday life.

Some suggestions about the way visual attention acts during the exploration of our surrounding can be drawn from eye-tracking studies on the perception of natural scene images. Velichkovsky, Joos, Helmert and Pannasch (2005) and later Pannasch, Helmert, Roth, Herbold and Walter (2008) for example, suggested two viewing modes of a natural scene: a global (ambient) mode when the general layout is extracted over the first few seconds, and focal scanning of detailed scene elements. Henderson (2008) reviews many other findings from the 2-D scene viewing paradigm. In general, viewing patterns vary considerably across participants and it is difficult to predict them, especially after longer periods of viewing. However, they are more consistent at the beginning of the viewing process. Interestingly, one such pattern has been observed as early as in the 1935, by perhaps the very first eye-movement researcher. Grounding his assumption on 2-D picture viewing paradigm Buswell (1935) suggested 'it is probable that most of the visitors to an art gallery look at the pictures with this [quick survey] type of perception and that they see only the main centers of interest' (Buswell, 1935, p. 142). So it might be the case that people first scan the scene and then view it in detail.

However, there has been a considerable criticism regarding direct applications of 2-D scene viewing patterns into our understanding of the nature of vision in the 3-D environment (B. W. Tatler, 2008). Even though some patterns are similar, their effect sizes tend to be much larger in the real world (B. W. Tatler, Gilchrist & Land, 2005). On the contrary, mobile eye-tracking allowing researchers to record oculomotor behaviour in 3-D is still a novel technique, and in many cases careful interpretations of the natural scene viewing studies is the best available source of data for generating hypotheses. The current thesis will refer to such works where necessary, but their interpretation in the context of real-life eye-movement should be conducted with caution.

The recent book by Land and Tatler (2009), to date, presents the most diversified collection of mobile eye-tracking studies from multiple real-life situations. As they note, the context of the considered task is crucial when studying human visual attention, because eye gaze is primarily directed to the places where information required for a particular action is likely to be present. In fact, studying attention without defined goals is undesired, as larger uncontrolled variation occurs when participants are prompt to select their own goals, even if implicitly (Land & Tatler, 2009). Since *spatial exploration* and *viewing objects within space* are the primary purposes of an art gallery visit, museums constitute an ideal research setting for studying eye-movement in the real-life context. Human 'cognitive aims' are rarely so uniform in any other building type and in consequence, the variability of confounding interpersonal factors is decreased.

As noted above, goal-free viewing is improbable. However, as Land and Tatler (2009) write, the largely accepted assumption is that there should exist some 'default viewing mode'. Laboratory-based studies for long suggested it might be the object's visual salience that guides visual attention during a goal-free exploration. This is known as the *bottom-up* strategy, since 'bottom' factors embedded in the environment would steer the 'higher' cognitive processes. The concept of *Saliency Map* (Itti & Koch, 2000)—a method of computing which elements of the viewing field 'stand out' from the background—gained a considerable popularity for making these predictions. This has been since questioned, as the correlations between Saliency Maps and eye fixations were shown to be the effect of a systematic bias in the saliency metric (B. W. Tatler, Baddeley & Gilchrist, 2005). Saliency Maps were also problematic in their theoretical aspect. If true in the real-life context, this bottom-up strategy of attention being allocation based on the stimulus' salience would be rather ineffective and dangerous (Land & Tatler, 2009; McCarley, Steelman & Horrey, 2014).

For this reason, a more pleasing explanation (both theoretically and empirically), is the *top-down* strategy. According to it, the eye movement is dependent on the current behavioural and cognitive goals, and allocated into regions which are relevant, rewarding, or explain uncertainty (Land & Tatler, 2009; B. W. Tatler, Hayhoe, Land & Ballard, 2011; Henderson, 2008). In other words, people look where they expect to find information important for their current objective (Land & Tatler, 2009; B. W. Tatler et al., 2011). In the context of a simple walking task, this phenomena is expressed by participants' attempts to obtain information about the desired walking direction very shortly before the turn of direction happens (Hollands et al., 2002). This top-down, task-dependent explanation is consistent with the early work of Buswell (1935), who observed many other top-down phenomena of oculomotor behaviour, such as its dependence on the participant's cultural background, expertise, or experimental instruction. Over 70 years later, behavioural and cognitive goals (the top-down strategy) are fully reinstated as the primary factor guiding our oculomotor behaviour (Land & Tatler, 2009). And yet, once these top-down cognitive goals are identical, some bottom-up characteristics of the environment are likely to affect viewing behaviour of multiple participants in a consistent way. It can be said, that the oculomotor behaviour consists of a set of context-specific routines, deployment of which is modulated by the cognitive objectives (Shinoda, Hayhoe & Shrivastava, 2001). Therefore, the everyday functioning of our attention seems to be the combination of both strategies as it has been described formally in the Contextual Guidance model by Torralba, Oliva, Castelhano and Henderson (2006).

3.1.4 *Connections Between Attention and Memory*

The main reason why the two processes of *visual attention* and *memory* will be analysed together is the fact that their functioning is inseparable (Henderson, 2008). Our visual acuity (i.e. resolution and colour sensitivity) is limited to a very narrow centre area of our visual field and a single fixation would not allow us to fully comprehend any larger object (Henderson, 2008). For this reason, multiple quick saccadic eye movements and fixations are made across the stimulus. At this point, attention and memory are already interacting - because what we see (or rather what we think we see) is the whole image, at once. This is only possible as the memories of the image's elements are sustained for tens of milliseconds (*trans-saccadic memory*), conveyed across the saccadic movements (*active online memory*), and combined to create the final representation of that image in our mind (Henderson, 2008). This combined representation is then unconditionally sustained in so-called *sensory memory* until vanishing, unless we actively maintain it in our *short-term/working memory*. When the latter happens, the representation can either be incorporated into more stable *long-term memory* structures, or discarded (Henderson, 2008).

The reader might now wonder how our cognitive mechanisms 'know' the connection between individual fixations. In order to link information obtained from separate fixations with a larger, discrete element of the visual field, our visual system does not use retinal coordinates, or spatial coordinates of the environment, but *object files*. This term is used to describe temporary representations of individual objects which are robust to the dynamic changes in the environment (Kahneman, Treisman & Gibbs, 1992; B. W. Tatler, Gilchrist & Land, 2005). Attention and memory therefore work together from the earliest stages of perception to overcome the physical limitations of our visual system and provide us with useful representations of discrete objects from our surrounding.

Attending to an object is by definition equivalent to processing it (Chun & Turk-Browne, 2007), assuming that attention is overtly deployed (R. D. Wright & Ward, 2008; Land & Tatler, 2009). To see, therefore, means to memorise. At least briefly. In which case it is not surprising that multiple studies showed a linear relation between the number of looks (or total fixation times) on an object, and the memory of it (for an overview see: Henderson, 2008). This arises from the fact that longer and more frequent viewing provides more opportunities for encoding (Henderson, 2008). Consequently, the existence of an object in our long-term memory structures gives evidence for deeper cognitive processing having taken place, or in the museum context: more absorbing cognitive engagement with it inside the gallery. Importantly, those objects which were search targets (i.e. it was the participants' explicit goal to identify them), are remembered even better than it could be predicted from the linear regression based on the number and length of fixations (Henderson, 2008). This demonstrates the strong influence of cognitive goals on the relation between our visual attention and memory.

The object file theory also accounts for this interaction of memory and attention. Once an object file is created, any information additionally gathered on it from eye fixations can be associated with it by updating the object file. B. W. Tatler, Gilchrist and Land (2005) showed that some visual aspects of objects within a scene accumulate over viewing time (i.e. with additional fixations), whereas some do not. The fact that participants' memory of object's *presence* and *colour* did not accumulate, suggests that this particular information can be gathered outside the fovea. Memory of object's *position* in the environment however, seems to increase with the number of fixations. The authors' suggested explanation is that each time the viewer's eye returns to an object, information on its *position* is updated in the object file, while information of its *presence* and *colour* are not.

It is important to note this interrelation between memory and attention is a two-directed process, as working memory is also used by attention to select and maintain relevant perceptual information from our surrounding (Chun, 2011). As a consequence, working memory has been named 'an interface' for attention (Chun, 2011). Henderson (2008) for example, notes how our oculomotor behaviour during 2-D scene viewing is initially guided by the known associations of the scene's identity. The subsequent eye movement is then directed by the object layout, learnt at the initial stages of visual exploration (so called 'gist'). In the real-world scenario Land and Tatler (2009) give an example of making a tea to illustrate this point. Participant's fixations proceed the next required action, but the knowledge of the kitchen layout makes it much more efficient to deploy eye fixations almost directly into the next relevant area of the environment.

3.2 MEMORY

In this section we will review the studies of memory of visual objects (or *visual memory*). Firstly, we will briefly look at how people memorise pictures. As it will be shown, in spatial context, the identity of an object is encoded together with its spatial properties. This is done within the structures of *spatial memory*, literature on which will be reviewed afterwards. A distinct case in this group of studies is the tradition of landmark research, which gathered evidence on spontaneous real-world human spatial behaviour. It has shown that the memory *of an object* is linked to, but often dependent on distinct factors from the memory *of its spatial location*.

3.2.1 *Visual Memory*

In general, human long-term memory of pictures, even those presented briefly is very good (Mulligan, 2013). Hollingworth (2008) reviews a number of studies where up to thousands of images were presented to participants and they still were able to correctly recognise 85-90% of them within days following the learning phase. The process of memorising pictures, and the biases it is sensible to, is similar to our verbal memory which forms the basis of most of findings on human memory (Mulligan, 2013). The only differences between our verbal and picture memory seem to be quantitative (i.e. in the size of reported effects) rather than qualitative. Mulligan (2013) reviews these known effects in the context of memory of pictures and lists the most impactful ones:

- *Age Effect* - older adults remember less pictures.
- *Serial Position Effects* - the order of stimuli presentation during learning affects memory. The first few items are usually remembered better, which is known as *the primacy effect*. Also the corresponding effect for the last few pictures—known as *the recency effect*—can occur if the test phase immediately follows the learning phase. If there is a temporal gap, or a distraction provided between learning and testing, the recency effect usually is eliminated but the primacy effect still applies. In some circumstances the recency effect can be considerably smaller for pictures than it would be for verbal stimuli.
- *Level-of-Processing Effect* - objects which were processed deeper are remembered better. To establish that, some studies asked for perceptual vs semantic attribute of an object. Whether the size of the effect for pictorial stimuli is equal, or attenuated in relation to its verbal memory counterpart, is unclear.
- *Divided Attention Effect* - attentional distraction greatly impairs consecutive memory. However, this effect mostly relates to studies which used various modifications of dual-task paradigm. This could be done, for example, with tonal distractors played in the background that had to be monitored during viewing pictures. In a study by Talmi, Schimmack, Paterson and Moscovitch (2007), in the presence of auditory distractors, participants were told to either focus fully on pictures, half on pictures and half on sounds, or mostly on sounds. Their attention happened to be easily controllable and the group instructed to focus on pictures did not get distracted very much. In the context of an art gallery visit, external distractors are largely eliminated, and it seems that *divided attention effect* does not apply to a situation where we are free to pick the object of our interest (or in other words: it does not seem to apply to endogenous, contrary to exogenous attention). Mulligan (2013) notes inconsistency in the results available to date, although one finding particularly interesting in the context of this thesis has been made by Reinitz and Hannigan (2001). In their series of experiments, recog-

nition memory of faces decreased when stimuli were presented simultaneously on the screen. The effect did not exist during sequential presentation. The limitation of this work is the fact that only conjunction errors were measured, and human faces constitute a rather distinctive type of stimuli on its own.

As mentioned previously, during viewing of natural scenes, visual objects are remembered based on the binding with their location in the scene's layout (Hollingworth, 2008; Akdal, Hodgson, Hill, Mannan & Kennard, 2002). The information about this layout is extracted before the information on the object's surface properties (such as colour or identity; B. W. Tatler, Gilchrist & Rusted, 2003). Moreover, the memory of an object's position accumulates with the increasing number of fixation, contrary to the memory of its surface properties (B. W. Tatler, Gilchrist & Land, 2005). This shows an interesting interplay between the traditionally viewed memory of pictorial stimuli and their spatial setting. As mentioned previously however, applying 2-D scene viewing task directly into our understanding of 3-D, real-world viewing situation can be deeply problematic (see Section 3.1.3, p. 44). B. W. Tatler and Land (2011) suggest that real-world actions require sparser and more task-specific memory representations of objects, which operate in conjunction with the larger models of the surrounding to stabilise the experience and ensure more efficient behaviour. In the previously mentioned work, B. W. Tatler, Gilchrist and Land (2005) also showed that the effect of better location memory accumulating with the number of fixations was stronger for real-life situation than for 2-D image viewing.

This difference in cognitive patterns is fully understandable. It would be purely ineffective for our spatial functioning to memorise each visual element of the environment according to exactly the same patterns we employ for learning pictures in a rather specific 'picture learning task'. Many of the described research further detached it from the real-life situations by presenting pictures of standardised, neutral, individual objects. However, the specificity of the art gallery experience, where attending to pictures is the 'default' cognitive task, advocates making careful predictions based on the aforementioned effects.

We will now take a close look at the phenomenon of object-to-location binding, mentioned on a number of occasions throughout the section above. Its application in the real-world setting was explored in depth in the studies of landmarks. However, before we turn into this narrower application, let us review the basic principals of spatial memory - the mechanism which allows us to remember where things are in space.

3.2.2 *Spatial Memory*

Kelly and McNamara (2009) define spatial memory as a set of two mechanisms which allow us to remain oriented with regard to the known locations in our close and distant surrounding. The first of them, *sensorimotor spatial memory* holds directions and distances between the navigator and known objects in his or her immediate environment. It allows us to control our spatial behaviour, such as walking around obstacles or recognising a close landmark. This system is limited in capacity and therefore requires constant updating during navigation: representations of the surrounding objects are added or erased as they become accessible. The second mechanism, *long-term spatial memory* holds distances and directions between all known objects, and is organised with regard to the spatial reference system (Kelly & McNamara, 2009). This thesis will investigate the nature of the latter mechanism in the context of an art gallery visit and the resulting knowledge of the exhibition.

According to McNamara (2013), *long-term spatial memory* consists of:

- *object-place knowledge* of the identity and appearance of distinct elements of the environment, a special case of which is landmark knowledge (when an object is used for navigation; Siegel & White, 1975);
- *route knowledge* of sequences of landmarks;
- *environmental shape knowledge* of the geometrical properties of the environment;
- and *survey knowledge* of the configurational relationships between individual locations.

The acquisition of these memories is spontaneous (Siegel & White, 1975; Montello, 1998) as long as attention is being paid to the environment (Montello, 1998). The above listed sub-types of *long-term spatial memory* usually develop in parallel (Montello, 1998). Importantly (and perhaps counter-intuitively), it is questioned to what extent the metric component is necessary and automatically incorporated into this knowledge (Montello, 1998; Penn, 2003). Lastly, interpersonal differences play a significant role in what spatial knowledge is extracted by two separate individuals from the equal exposure to the same environment (Montello, 1998; Davis, Therrien & West, 2008; Spiers & Maguire, 2008). The most critical stage moderated by these differences is the moment of integrating separately learned places into a hierarchical, complex ('survey') knowledge structure (Montello, 1998; McNamara, 2013).

The importance of integrating individual locations into *survey knowledge* arises from the fact that every unit of space can be remembered as a discrete location, or as a part of a larger set-up. This ability allows us to use our spatial knowledge at different scales (Kelly & McNamara, 2009) and, for example, to draw a layout of an apartment based purely on spatiotemporally linear explorations of its individual rooms. Even then however, spatial knowledge often remains *fragmented*, meaning that while some areas might be remembered in detail, their neighbouring locations need not necessarily be encoded so well (McNamara, 2013).

At this point it becomes important to understand what constitutes a single, psychologically discrete location. This problem has been tackled by Montello (1993) who proposed a classification of psychological spaces. Among others, he distinguished *vista spaces* and *environmental spaces*. *Vista spaces* are spaces which can be visually apprehended from a single point, without considerable amount of travelling. *Environmental spaces*, on the contrary, are larger environments requiring some locomotion to explore them entirely, but this exploration (and the subsequent memories of it) need not be aided by symbolic representations such as maps (Montello, 1993). A *vista space* is therefore usually equivalent to a simple room, while *environmental space* can be understood as a set of spatially organised *vista spaces*. It is this organisation of *vista spaces* into *environmental space* which likely constitutes the boundary between learning discrete locations and integrating them into *survey knowledge*. The focus of this work is on *vista* and *environmental* spaces. In addition, Montello (1993) distinguished *geographical spaces* (requiring symbolic representations to comprehend) and *figural spaces* (which are smaller than the body size; e.g. pictures of natural scenes). He emphasises a qualitative difference between the set of psychological processes involved in learning these spaces. For this reason, this thesis does not review laboratory-based experiments on object-location binding conducted on computer screens in further detail. As they investigate the location encoding *within figural spaces* (usually a computer screen), the underlying cognitive mechanisms responsible for those effects are likely to differ, even if the results are often similar to other situations (e.g. Mather & Nesmith, 2008).

One of the most important difference might be the reference system through which spatial learning takes place. The reference system used to encode and retrieve spatial information can be *egocentric* or *allocentric* (McNamara, 2013). *Egocentric* reference system specifies location in relation to the perceiver. It consists of eye coordinates, head movement, and body orientation (McNamara, 2013). The *allocentric* system specifies spatial relations to external elements of the environment, such as room shapes, or—on the abstract level—geographical latitude and longitude (McNamara, 2013). Spatial reference systems of smaller locations can also be hierarchically interrelated to each other (McNamara, 2013). For example, it has been shown that spatial relations between objects are stored in the allocentric system, while the visual memories of particular landmarks are sustained in the egocentric system (McNamara, 2013). Although it is important to note that the two systems exist in parallel and their relationship is dynamic (Burgess, 2006; B. W. Tatler & Land, 2011; Philbeck & Sargent, 2013).

Putting this into the art gallery context, we can summarise that:

- humans spontaneously acquire long-term knowledge of the spatial properties of the gallery being explored;
- this knowledge consists of information on discrete locations (e.g. how 3 pictures are hung next to each other on one wall), as well as how specific locations relate to each other at the scale of the entire gallery;
- the quality of this knowledge is likely to be fragmented, i.e. different for discrete sub-parts, and will vary depending on (among other factors) the attention being paid, interpersonal differences, and the complexity of the environment;
- potentially the most difficult task for art viewers' spatial memory is to integrate those separately learnt locations (individual art objects, gallery walls, and rooms) into a hierarchically organised structure of the environmental space;
- this is a challenge due to the need of integrating egocentric views of a single artwork, a single room, or a single gallery wall, into the allocentric system containing relations between these sub-parts of the gallery;

In order to investigate how these memory processes combine in real-life situations we will now review the research literature on landmarks.

3.2.3 *Landmark Studies*

The[1] mechanisms of spatial cognition that guide our exploration of space and spontaneous memory of encountered objects have been largely investigated in the landmark literature. If we define landmarks as easily recognisable objects serving as a point of reference in space (Chan, Baumann, Bellgrove & Mattingley, 2012) and we assume that their acquisition is spontaneous (Chan et al., 2012; Janzen & van Turennout, 2004) then the findings from these studies could be applied to artworks in a gallery space. Even during a free exploration of an art gallery, visitors must still use their spatial abilities to orientate themselves in space. If this was not true, their movement would be random, with no mechanisms helping them to avoid revisiting rooms, to explore new spaces, and to find one's way out. Viewing

1 Parts of this section were previously published in (Krukar, 2014b)

art in a museum must be inextricably linked with acquiring knowledge about spatial location of certain objects and the gallery's layout. For this reason it is important that studies of landmarks make the distinction between *object-based* and *location-based* attention (Caduff & Timpf, 2008) resulting in the *knowledge of objects* and *knowledge of the spatial relations between them* (Montello, 1998). For instance, Janzen (2006) observed that objects placed next to decision points (junctions) in a virtual museum are recalled faster on a computer-based recognition task than those placed along straight paths. The author suggested that this effect might be the result of a linkage between the memory representation of a particular object and the representation of its location. This explanation would be in line with different neural activity patterns in the parahippocampal gyrus (responsible for place-object mapping), that can be induced by decision-point-based and non-decision-point-based objects (Janzen & van Turennout, 2004). In a later study, Miller and Carlson (2011) designed a similar virtual museum to Janzen's, but in addition aimed to take the objects' perceived salience into account. Perceived salience, in the real-life context, is likely to be dependent on many more factors than just visual dominance (e.g. emotional reaction, or previous knowledge/memories about the object). For this reason, objects placed inside were separately rated for their perceived salience by an independent group of participants. The authors managed to replicate Janzen (2006) results when objects of high perceived salience were placed on decision points, but did not, when highly salient objects were purposefully placed on navigationally irrelevant locations (i.e. on non-decision points). In the latter variation of the experiment, high perceived salience was a factor enhancing response times on the computer-based recognition test, while navigational relevance guided participants' responses in map drawing and route description tasks. As the authors conclude, the encoding of a landmark (object) might be driven by its perceptual features, whereas its selection during spatial tasks seems to be driven by its spatial features (Miller & Carlson, 2011) (see also Miller, Carlson & Hill, 2011). This shows the importance of separating the object-oriented and location-oriented memory in this context. Those results by Miller and Carlson (2011) also suggest that the former should be highly dependent on the objects' perceived salience, while the latter should remain unrelated to it.

While the qualitative distinction between those memory types seems fully applicable to the art gallery experience, the specific factors having an effect on the strength of those memories are likely to differ. Navigational behaviour highly varies depending on the actual context, as well as goals and strategies of the individuals (Steck & Mallot, 2000; H. Taylor, Naylor & Chechile, 1999). A gallery visitor is likely to direct his/her attention to artworks, and in a visually ascetic space these become highly salient reference points, but their navigational importance is secondary. Instead, the spatial relations between them can become an important component of the viewer's understanding of curatorial intentions (either explicitly or spontaneously). A gallery visit therefore incorporates both types of attention: object-based and location-based. These induce memory traces of the individual objects, as well as of the spatial relations between them. Object-based memories should therefore mainly (but not exclusively) be influenced by the picture's perceived salience. Location-based memories are dependent on the pictures' position in the gallery, and therefore derive from initially designed curatorial narrative (Psarra, 2009). A noticeable disadvantage of applying landmark studies to non-navigational research is the importance of goals for spatial behaviour and learning (A. Johnson, 2011; Land & Tatler, 2009; H. Taylor et al., 1999). In principle, landmark studies put the emphasis on learning the route or objects encountered along it (Buchner & Jansen-Osmann, 2008; Miller & Carlson, 2011; Janzen, 2006; Janzen & van Turennout, 2004) which is a goal very different to the purpose of visiting an art gallery. For this reason, if the effect of spatial location in landmark studies depended on its importance for navigational decisions, perhaps in the art viewing context this effect shall be influenced by the location's importance for potential understanding of the narrative of the exhibitions (such as co-visibility of multiple artworks with it).

A related disadvantage is the fact that the above mentioned studies measured spatial memory by asking for route descriptions. In such a case it is predictable that objects placed at decision points are mentioned more often, and this does not necessarily mean they were remembered better. A different task, allowing for expressing spatial relations between nearby, as well as distant objects would be more desired for investigating spatial memory in the art gallery context.

One more important contribution to the rationale behind this work comes from the landmark study conducted by Buchner and Jansen-Osmann (2008). As their results showed, learning of ordinal position of objects in a dynamically experienced virtual spatial environment was superior to learning the same objects presented as a serial, non-spatially embedded list of items. This effect disappeared when corridors along which the objects were located have been transformed to equal lengths (Buchner & Jansen-Osmann, 2008). This result shows that experiencing objects in the spatial context is a qualitatively different experience from a non-spatial one, and that the objects' mental representations are not just an environmentally-embedded version of a standard list-type representation. This fact could be one of the main reasons why the general public still prefers the hassle of travelling to a physical museum building, time commitment, and (often) paying the entrance fee, over viewing the same artworks through the internet or a in a printed catalogue. This argument adds to the finding by Larish and Andersen (1995) who showed that active control of vision and movement led to supreme performance in perceptual task over the passive observation condition. Being able to actively explore the environment on our own makes us attend and remember better.

Importantly, the lack of the effect under the equal-corridor-length condition in the study of Buchner and Jansen-Osmann (2008) signifies the importance of environmental differentiation for enriching the learning experience. This study did not explain whether such an environmental differentiation must be metric (contrary to other potential factors, such as wall colours, or textures). Yet, it raises an important concern in the context of white cube art gallery spaces, where environmental differentiation is often purposefully eliminated.

It is important to notice, that a significant proportion of the the recent landmark studies have been conducted in the virtual set-ups (Buchner & Jansen-Osmann, 2008; Miller & Carlson, 2011; Janzen, 2006; Janzen & van Turennout, 2004). Since metric information can be distorted in virtual environments (Marsh, Chardonnet & Merienne, 2014) some of the effects shown might be underplayed, or overexaggerated compared to real-life experience. Additionally, where alternative forms of navigation are imposed instead of walking involving bodily movements (e.g. watching first-person video, or navigating the virtual reality via a joystick; Buchner & Jansen-Osmann, 2008; Miller & Carlson, 2011; Janzen, 2006; Janzen & van Turennout, 2004), the experience is further detached from the real-life experience and have been shown to activate different spatial learning mechanisms (Hegarty, Montello, Richardson, Ishikawa & Lovelace, 2006). Therefore, the importance of experiments investigating spontaneous spatial learning of objects in real-life settings is noteworthy.

3.3 A NOTE ON IMAGE RECOGNITION

A considerable issue with applying visual attention and visual memory studies to the real-life context is that laboratory-based studies control for the angle of viewing. Contrary to that, during unrestricted[2] spatial explorations people view objects at multiple oblique angles. In the context of art gallery studies, this becomes potentially problematic as visitors explore the space via multiple pathways and therefore viewing pictures from a comfortable straight-on location is often interrupted by less ideal slanted views.

Retinal image of every scene changes as the viewer moves the head even slightly. In a dynamic viewing situation, multiple slanted views of a painting result in greatly distorted retinal images of that picture. Numerous studies on image recognition have shown, however, that

2 Contrary to this thesis' approach, many Spatial Cognition studies restrict participants' freedom to explore the environment by predefining their route, or walking pace. Consequently, the variance in behaviour not directly related to the research question is decreased at the cost of decreasing ecological validity.

this is less of a challenge to human visual systems than it might at first appear. E. Bruce Goldstein extensively studied human ability for recognising slanted pictures. Despite the above described phenomenon, these visual distortions pass largely unnoticed until the viewer's attention is drawn to them (for example, by simultaneously presenting the same image from another viewing angle; Goldstein, 1987). Goldstein (1987) empirically distinguished between various picture properties which are perceived differently across a wide range of viewing angles (between 20 and 160 degrees). Perceived relations between distinct elements of the picture's layout have been identified as the most robust to distortions. In general, it has been indicated that human perception of slanted images is very good, as long as the information on the underlying surface orientation is available (Vishwanath, Girshick & Banks, 2005; Yang & Kubovy, 1999). This is almost always the case in the art gallery context, as well-lit white walls and their orientation tend to be easily recognisable. Therefore, viewing images from oblique angles—as it can often happen during an unrestricted exploration of an art gallery—is unlikely to prevent the visitors from meaningfully attending to them.

This note represents the perceptual explanation of the problem. In later sections we will see how the availability and accessibility of different viewing perspectives can guide the gallery visitor through space. Despite the fact that slanted viewpoints should not have a large impact on how we recognise the artworks' content, they might affect our understanding of the individual exhibits, especially with the respect to their broader surrounding.

4

COGNITION IN THE WILD... ART GALLERY

In the previous section we have looked more closely at *attention* and *memory*: two cognitive processes, functioning of which can potentially be indicative of whether an art exhibition layout is allowing the visitors to comfortably view and process the exhibited art. As it has been indicated, the majority of studies originally investigating these aspects of human cognition bear some considerable limitations. They primarily used laboratory-based paradigms (such as 2-D viewing or virtual reality set-ups), and often included procedural goals related to memorisation, or navigation. Each of these aspects has been shown to affect the applicability of the results to our understanding of human cognition during a real-world situations, such as an art gallery visit.

This section will look at studies which often sacrificed experimental control to pursue higher ecological validity. We will review the *Visitor Studies* movement, which studied the viewers' engagement by observing everyday museum visitors during their actual visits. Despite the fact that only a small proportion of this research bears any relation to the spatial configuration of exhibitions, their relevance to this thesis lies in developing techniques of estimating the visitors' attention to, and engagement with an exhibit. We then consider the drawbacks of these research approaches and turn the reader's attention to the emerging field of more rigours cognitive studies inside real-life museums. Only the most recent technological developments have made it possible to assess the visitor's engagement with the exhibition in a relatively unobtrusive, yet more reliable manner.

Following the scope of the thesis, this chapter will focus on the cognitive and behavioural aspects of the museum visit; for a broader overview, including the social and the aesthetic aspects, see the work by Kirchberg and Tröndle (2012).

4.1 OBSERVATIONAL VISITOR STUDIES

Systematic investigations of the visitor behaviour started as early as in 1928. Robinson (1928) conducted a set of comparative museum-based observations in accordance to a list of predefined, standardised rules (such as when to assume that a stopping person is attending to the nearby picture). Tracking and observations were conducted secretly on randomly chosen visitors, timed with stopwatches, and recorded in notebooks (Robinson, 1928). Timings recorded in museums (such as total time inside, time of engagement with each picture, number of pictures skipped) were supported with laboratory-based experiments on separate groups of students (Robinson, 1928). In relation to artwork locations, Robinson (1928) emphasised the importance of a picture's isolation, especially for smaller pictures which are not formally predisposed to attract more attention. By varying the number of pictures presented to participants for simultaneous viewing in a laboratory, he found that the time spent on observing a single picture does not decrease proportionally to the number of presented artworks. He suggested, that isolation is only effective in lengthening the viewing time, if it is 'complete'; that is, if only one picture is present in the viewing field.

Rules of recording and engagement measures developed by Robinson for many decades formed the basis of the most popular research approach to studying human behaviour inside museums. Melton (1935) for instance, used very similar observational techniques and used the term 'competition' to suggest that paintings fight for visitors' attention in a museum setting. This followed an observation that people stop in front of smaller number of paintings, as the number of artworks in a gallery space is increased (Melton, 1935).

Many of these early findings were revisited by Bitgood (1991) when observational visitor studies gained significant popularity in the 1990s, mainly in the United States (for a historical review see Tröndle, Greenwood, Kirchberg & Tschacher, 2014).

In one the most impactful works from that period, Falk (1993) tracked over 300 visitors of the same science exhibition, in two versions of spatial arrangement: 'structured' and 'unstructured'. Visitor's path, use of exhibit elements, time spent inside, and approximated (by the investigator) demographic details were recorded. Participants showed better understanding of the material seen in the 'unstructured' version and expressed less negative comments than those who experienced the same material arranged in a 'structured' way. The significant disadvantage of the study is the method of interpretation of the open-ended questionnaires which assessed visitors' level of understanding. It did however link the notion of a flexible spatial setting with some earlier theoretical suggestions on the importance of empowering the visitors in their explorations and interpretations.

Beverly Serrell (Serrell, 1997, 2011) took a broader view on multiple empirical studies of museum visitors. She identified the most repeatable patterns in visitor behaviour, observed by multiple authors (e.g. by Bitgood, 1991). These include the visitors' tendency to turn right, more attention given to the beginning than the end of the exhibition, or ineffectiveness of centred 'island' exhibits. Her general observation is that visitors often leave museums at the very first opportunity they happen to be near the exit, and that the time they spent attending to anything is extremely limited (Serrell, 1997). As the author noticed however, the majority of this empirical data suffers from low transferability and generalisability, as it tends to derive from separate case studies. For this reason, Serrell (1997) made an attempt to synthesise the variables relating to visitor attention inside museums, with the particular focus being on science museums. She developed two measures of museum performance:

- *Percentage of Diligent Visitors* (%DV), which is the proportion of users who pause in front of more than half of the exhibits, and,
- *Sweep-Rate Index* (SRI), which is the ratio of mean time spent by visitors inside the museum divided by the exhibition area, resulting in a square-feet-per-minute ratio allowing to compare buildings of various sizes.

A meta-analysis was conducted and a large database consisting of over 100 cases constructed on the basis of the relation between %DV with SRI (Serrell, 1997, 2011). This allowed to investigate the current standards in the museum field and seek for the common characteristics of the exhibitions which place exceptionally bad, or exceptionally well on this scale. For instance, larger exhibitions were shown to be walked through faster, on average (Serrell, 1997, 2011).

Most recent technological advancements provided new opportunities to confirm these findings. Passively switched on Bluetooth devices were used to track visitors in Louvre by Yoshimura, Girardin, Carrascal, Ratti and Blat (2012). It was shown, that those participants, who travelled through less areas, spent more time in the museum than those who 'rush through' many spaces (Yoshimura et al., 2012). Despite the fact that the authors do not provide an explanation for this finding, it does stay in line with the summaries of earlier observational studies (Serrell, 1997, 2011).

The main reason for the lack of explanation is that observational studies which formed the basis of Serrell's dataset cannot account for the specific factors causing the overall effect. In observational studies, attention is only assumed to be given to an object when the visitor pauses in front of it. This discredits all attentional interactions taking place while walking and cannot reliably account for co-visibility of multiple objects. Serrell's measures neither take into account that SRI can vary for separate sub-parts of the exhibition, depending on its attractiveness, or layout. As these variables fail to explain the local characteristics of successful exhibitions, they might serve well as a comparative tool differentiating individual museums, but it is unclear how they can contribute to the design of new successful exhibitions.

In an effort to expand Serrell's ideas, Bollo and Dal Pozzolo (2005) observed 357 visitors in three different museums in Italy and asked the observed individuals to provide questionnaire answers. The analysis of those suffer from subjective interpretation and remain speculative. However, an important contribution of this work is that it attempts to visualise different behavioural patterns on the layout of the exhibition as 'hot' and 'cold' exhibition areas. This tackles the need for assessing visitors' engagement on the level of individual exhibits, although the potential influence of space is here strictly connected with the influence of the individual artworks and therefore the results remain case-specific. Moreover, the method is not replicable due to the subjective character of the estimations (Bollo & Dal Pozzolo, 2005).

Another approach oriented on local properties of individual exhibits was taken by J. K. Smith and Smith (2001) who timed visitors viewing paintings, trying to relate these patterns to their age, gender, and group size. Despite some flaws common to this type of observations (e.g. inability to distinguish reading labels from viewing pictures), the authors found a significant difference between viewing times of individual paintings (J. K. Smith & Smith, 2001). Also, it was observed that the viewing behaviour demonstrated a repeatable pattern consisting of two phases:

- the initial period of viewing shorter than 10 seconds, when (supposedly) the decision is being made about stopping or progressing forward,
- and the period of stopping and diligent viewing, averaging to about 30 seconds per picture.

However, the inability to control for label reading times makes these exact numbers questionable. Even the entire relationship between two viewing phases could be untrue if a consistent bias in human preference for reading labels were shown. This is especially concerning considering that labels tend to attract significant proportion of visitor's attention (Bitgood, 1991; Bitgood & Patterson, 1993; Bourdeau & Chebat, 2003). On the contrary, the potential influence of the gallery's layout could adopt distinct patterns for each of those qualitatively distinct viewing phases.

The two viewing phases distinguished by J. K. Smith and Smith (2001) should not be confused with the ones identified in laboratory-based studies. J. K. Smith and Smith (2001) had only the chance to observe their participants' whole-body locomotion. Thus, it is difficult to make assumptions of the attentional mode exhibited by the viewer at any given moment. However, the fact that participants tended to stop to examine pictures for longer than 10 seconds is an observation which has not been reliably measured before.

4.2 towards user-centric exhibition design

The need for consolidation of often similar case study findings has prompt researchers to propose more generic models, accounting for multiple factors affecting the visitor's engagement with an exhibit. These models look at the museum experience in the context of learning. This is often very different from an art gallery experience where the curator's focus rarely lies on teaching per se. In fact, even in science museums visitors have no obligations to learn anything (Serrell, 1997). What does connect these approaches however, is the fact that learning is a consequence of diligent cognitive engagement with the exhibition's elements. If such engagement is desired by art gallery curators, factors which contribute to it should be carefully considered. Even if the final visitor experience is much less predictable compared to learning environments.

Falk and Dierking (2004) suggested to divide these factors into three contexts:

- Personal Context is affected by visitor's motivation, prior knowledge, and the ability to utilise one's preferred choice while structuring own learning experience.
- Sociocultural Context is influenced by the relations within the group visiting a museum together and can be facilitated by external moderators (such as museum staff).
- Physical Context can affect the learning experience by facilitating whether people feel oriented (instead of disoriented) in the novel museum setting, whether exhibits are well-designed, and by providing contextual references that will reinforce the new knowledge after the visit.

Another widely acknowledged model for informal science learning environments, is the *Attention-Value model* proposed by Bitgood (2010). He suggests that attention inside a museum setting is a continuum consisting of three stages: *capture*, *focus*, *engage*. According to the author, various factors have an impact on each phase. For instance, it is hypothesised that attention is captured in a different way (either 'sequential' or 'simultaneous'), depending on the layout and organisation of the exhibition. According to Bitgood (2010), clearly organised sequence of the exhibition must result in longer visual engagement and better understanding of the exhibits. When the organisation of the exhibition leaves more freedom of choice to the viewer, the author suggests this will result in 'spotty focused' attention. He suggests that stimuli should be placed in a way in which they do not draw attention from one another and result in equal viewing chances of capturing the viewer's attention. Once attention is captured, Bitgood (2010) writes, it will be focused for a few seconds more on an object, if it is isolated from other artworks. The last stage of engagement is less dependent on spatial factors, and involves subjective factors of the feeling of immersion and making meaning. The model assumes this is based on *perceived value* of each element of the exhibition. This value would be implicitly calculated by the visitor as a ratio of potential satisfaction divided by its costs (such as time and effort required to engage). Such a perspective is a continuation of the 80-years old interpretations made by Melton (1935). Bitgood, McKerchar and Dukes (2013) revisited Melton's results suggesting that it is not only the perceptual competition of the pictures that decreases allocated attention, but the engagement's potential value being perceived as higher when one becomes more selective. The mechanisms responsible for this selection are suggested to be *perceptual distraction*, *object satiation*, *fatigue*, and *selective choice*. As Bitgood et al. (2013) admit however, the empirical data confirming their reinterpretation is scarce, and despite their attempt to unify multiple research findings under one explanatory model, both the explanations and the proposed guidelines for exhibit designers remain speculative.

Bitgood's work, similarly to Falk and Dierking (2004), was also based on science museums, but in its particular case the difference between this context and art galleries becomes more problematic. The idea of 'value' and 'cost' presented by Bitgood might not be relevant in relation to art, as both of these factors would be highly depended on the personal preferences and difficult to interpret or quantify in any manner.

Another problematic aspect of the model is the assumption that every person coming into a museum does so with very limited cognitive resources and the default goal of leaving it as soon as possible in order to preserve those resources (similar views were represented by Serrell, 1997). Designing highly engaging exhibits would then be barely the means of slowing the visitor down on his or her way to the exit, for which reason also sequential structuring of the exhibitions was recommended (Bitgood, 2010). Such understanding of the visitor experience is speculative and there are views contrary to it. Museums were, for instance, interpreted as *restorative environments* (Packer & Bond, 2010) which would bear quite the opposite claim - that visitors' mental resources restore, and not exhaust, throughout the museum visit.

A broader Model of the Dynamic Museum Environment (Thompson, 1990) refers to the interaction between the museum, the visitor, and the exhibits. Thompson (1990) reviewed many of the same visitor studies as the authors listed above and suggested that the *desired outcome* of a museum visit leaves many possibilities, depending on the museum's scope and goals. This fact makes Thompson's model easily applicable to all types of museums. Out of all factors which influence the visitor experience, only some, as he mentions, actually remain in the museum's control. Thompson (1990) divides those factors into two categories: *Physical Environment* and *Social/Organisational Environment*. The former is related to architectural properties of the exhibition space, and the latter to the museum as an institution - its goals, educational programmes, reputation, and regulations (Thompson, 1990). According to this classification, the focus of the current thesis is purely on *Physical Environment*. In this respect, Thompson (1990) names the critical architectural entities important for visitor-centred design of exhibition spaces:

1. Accessibility;
2. The ambient environment (including distinct acoustic, thermal, and visual elements);
3. The circulation route;
4. The relationship between architecture and exhibits;
5. Spatial relationships;
6. Symbolic variables (including building configuration, spatial configurations, materials, the nature of illumination, color and the nonvisual environment; Thompson, 1990, p. 79).

The specific scope of this thesis is on points 4 and 5: i.e., on the way spatial layout affects spatial relationships between artworks, and on the spatial properties of individual artworks.

4.3 LIMITATIONS OF OBSERVATIONAL VISITOR STUDIES

Observational techniques underling these theories have became a golden standard of visitor studies in the recent decades. In many countries they are a common part of the museum evaluation process—often required by the policymakers—with up to 10% of museum's budget allocated towards this goal (Arnsdorf, 2010). Methodological handbooks were published (e.g. Diamond, 1999; Yalowitz & Bronnenkant, 2009), also by public bodies (Westat, 2010). The methods contained in the guidelines involve observations, pen-and-pencil tracking of traveled paths (with annotations being made of 'events', such as stops), timing of viewing/walking periods, surveys and self-reports (B. L. Taylor, 2010). With the tablet technology becoming more accessible, the process of data collection and analysis can be computerised and simplified with an iPad application (Dalton, Conroy Dalton, Hölscher & Kuhnmünch, 2012). Yet, the fundamental methods of observation used to assess the visitor's engagement with, and attention to an exhibitions remain largely unchanged since the era of Robinson (1928).

These investigations have significantly contributed to the understanding of the process of exhibition viewing, but at the same time remain very limited. The reason being, they can only measure occurrence of an observable behaviour, or participant's much biased self-evaluation. Very little can be derived this way about the quality or intensity of the cognitive processes underlying this behaviour, and resulting from it. It can be observed, for instance, that a person looked in the direction of an artwork for a number of seconds, but no third person observer can be sure if the visitor's eye-sight was actually fixated on a painting nor how the dynamic changes in this process were facilitated by potentially co-visible objects. This fact is threatening to the external validity of each such a study - the correlation between *observable* viewing patterns of a visitor and the *actual* viewing patterns is not perfect. The amount of error introduced to the measure (as the majority of the reviewed papers seem to miss) potentially detaches these findings from the basic theoretical concepts which are supposed to guide the design of these studies and to which the findings ought to contribute.

Using questionnaires, as in some studies reviewed earlier (e.g. McManus, 1993; Bollo & Dal Pozzolo, 2005) provides no remedy, as those typically are designed with some pre-defined outcome of the visit in mind, in order to assess it. Open-ended questions, on the contrary, usually relate to high-level aspects of human cognition, while in reality the benefit of visiting a museum might be different for each visitor and therefore be only 'tangible' at earlier stages of processing, before interpretation (or the lack of it) takes place (Kirchberg & Tröndle, 2012).

Also, because of the case-study character of the majority of this research, they are not able to establish a causal relationship between artistic or curatorial actions and their impact on the visitor experience (Tröndle & Tschacher, 2012). Combining empirical results of multiple case studies can only be made on the common ground of more generic variables (e.g Serrell, 1997), and the theoretical models remain speculative as no data proving their claims in a generalisable manner is available (Bitgood et al., 2013).

As we have shown in Section 3 (p. 41), *visual attention* and *memory* are processes highly relevant to the evaluation of the museum experience. The former process allows visitors to interact with stimuli present in an art gallery in a non-random, selective, and directed manner. The latter process can be indicative, post-visit, of the fact that such an interaction took place and whether it bore any meaning to the person. As the methods widely used for the measurement of these processes origin from psychological laboratories, assessing them reliably in a real-life setting is a challenge. As the above critique of observational studies shows, this challenge is well-known to researchers interested in the behaviour and cognition of museum visitors. Only recently, the interdisciplinary linkage between Cognitive Sciences, Experimental Psychology, and the field of Aesthetics made it possible to study in depth how humans view and cognitively process art. The most recent technological advancements made it possible to tackle the problems of observational visitor studies.

4.4 EXPERIMENTAL AESTHETICS: LABORATORY-BASED APPROACH

Term 'Experimental Aesthetics' emphasises the experimental approach adopted from psychological methodology to study the aesthetic phenomena. Questions guiding these investigations derived from the belief that experiencing art—even if unique to each of us—might be studied and described in accordance to its universal cognitive components.

These investigations were started as early as in the 19th Century (Fechner, 1871; Witmer, 1893) in a psychological laboratory. They continued to develop and incorporate new techniques when Buswell (1935) used his 'aparathus' to record scan paths of his participants' eye-movement on analogue film while they examined a number of individual artworks. Some of the research questions stated by Buswell still remain open today. He wondered, for instance, if there is any common viewing pattern used 'by default' by all of his participants while they view new pictures. He was intrigued, if such viewing patterns would vary across different levels of art expertise, cultural backgrounds, or instructions given at the beginning of the task. As we have seen in Section 3.1.2 (p. 43) describing eye-tracking studies in non-art context, those factors constitute a challenging element of even the most recent studies (e.g. Land & Tatler, 2009). Buswell (1935) also touched upon the influence of more visually salient picture areas - 70 years prior to the *Saliency Map* concept proposed by Itti and Koch (2000), which has only recently been disregarded (B. W. Tatler, 2008). Buswell's research, despite technological challenges of his era, until very recently have still constituted the most comprehensive analysis of human eye-movement with respect to distinct cognitive and social factors influencing the way people look at art.

We will not focus on later, individual studies, usually smaller in scope, as their direct relevance to this thesis can be questioned by their laboratory-based paradigm (see Section 3.1). However, the resulting, more general models trying to describe the nature of the aesthetic experience can possibly help to explain human behaviour inside art galleries.

A model developed by Leder, Belke, Oeberst and Augustin (2004) took into account the fact that the art itself has changed substantially since Buswell's times. Figurative paintings are no longer the sole representation of 'art'. In fact, abstract forms are what we expect to encounter more in contemporary art centres. If there is anything universal in the emergence of the aesthetic experience (what can be observed through the oculomotor behaviour), it should remain constant across different art genres, also on the figurative-abstract spectrum. Leder et al. (2004) provide the framework of cognitive processes contributing to the rise of such an experience with the main distinction being made between automatic and deliberate processes. During the first—automatic—stage, the authors emphasise the role of the artwork's perceptual features, as well as the early activation of implicit memory processes which attempt to integrate the viewed painting into the existing knowledge structure (e.g. by identifying its similarity, or prototypicality). The latter, deliberate phase, involves domain-specific expertise and contextual interpretations. By making this distinction, the model accounts for both bottom-up, and top-down processes involved in the functioning of our visual attention and applies it to the aesthetic context.

In another model, largely in line with the one described above, Locher et al. (2007), building on their own eye-tracking experiments also emphasise the two-stage nature of the aesthetic visual experience. The authors differentiate between a quick, automated decision (a 'gist'), and longer, diligent viewing following later. In one of their laboratory experiments, participants were first asked to write down their reactions to 100 millisecond-long glance at artworks. In another study, art viewers' oculomotor behaviour was recorded alongside Think Aloud Protocols[1] while they were rating artworks for pleasingness. The results demonstrated that the 'gist' participants got from a 100 ms-long glance contained not only perceptual features of the artwork, but can also be related to its semantic meaning. When asked to assign a pleasingness rating, participants spent 32.5 seconds on average before doing so. This stays in line with the suggestion described earlier by J. K. Smith and Smith (2001) who observed that half a minute is an average period people spent in front of an artwork at The Metropolitan Museum of Art before they proceed forward. Locher et al. (2007) treat this as the indication of external validity of their laboratory study. The authors also relate their finding directly to the model suggested by Leder et al. (2004). This, when aligned with museum-based observations also seems to suggest that art viewing experience consists of an automatic, implicit 'gist' (when a decision is being made of whether to deploy diligent attention) and a scrutinised interpretation of an art piece. This would also explain, according to the authors, how the decision about 'skipping' an artwork is being made (Locher et al., 2007). The second phase of the process would be therefore optional, although in laboratory-based studies—where the 'cognitive aim' is to analyse pictures in grater detail—this optionality might be difficult to observe. In a real-world museum setting the viewers' behavioural freedom and a plethora of alternative 'visibility targets' might prompt more often 'skips' following the gist phase.

1 Think Aloud Protocols is a research method, in which participants are asked to verbally pronounce their thoughts as they conduct the initial task.

This model (Locher et al., 2007) also explains how the aesthetic value judgment is being made. Laboratory-based studies showed that fluency of processing has a significant impact on this judgement (e.g. Belke, Leder, Strobach & Carbon, 2010). According to Leder et al. (2004), two types of processing fluency play role in the process: *perceptual* and *cognitive fluency* of processing. In this model *perceptual fluency* makes its impact at the initial stage, after which *cognitive fluency* is assessed, judgment of which can override the initial, *perceptual* one. The model therefore emphasises the dynamic process-like character of the aesthetic experience.

According to this statement, perceptual fluency's influence can relatively easily be measured in a laboratory-based experiments, by limiting the time a picture is presented. However, it is unlikely to be of major importance in a real-life museum situation, where each person can view artworks for long enough to allow cognitive fluency to dominate one's judgment.

What the laboratory-based studies showed with confidence, is that the process of viewing art is significantly different from viewing everyday scenes, both perceptually and cognitively. It is because art provides 'a particular processing challenge' (Belke et al., 2010, p. 221) which emphasises the influence of processing fluency. It is this reason, why the role of presentation quality in the museum context should be so pronounced. Appreciating art is a resource-demanding activity sensitive to numerous factors that can enhance, or inhibit it. Spatial factors are, as we here argue, one of the most important of them.

It is important to note at this point, that viewing times, despite some previous assumptions, seem not to be correlated with aesthetic judgments (Isham & Geng, 2013). The above described characteristics of the 'aesthetic experience' do not therefore explain which objects are 'liked' more based on the oculomotor behaviour of the viewer. As Isham and Geng (2013) demonstrated, longer viewing instead predicts which object will be chosen in response to the experimental task.

4.5 EXPERIMENTAL AESTHETICS: MUSEUM-BASED APPROACH

Leder et al. (2004) emphasised how important it is to consider that the aesthetic experience might differ in and out of a psychological laboratory. It is an inseparable element of this phenomena that it occurs in a 'safe' environment of the art gallery. This argument adds up to the previously reviewed importance of considering 'cognitive goals' while studying human visual attention and memory. Any experiment aimed at explaining this experience must ensure that the participants' cognitive goals are as close to a real-life gallery visit as possible. Possibly the most reliably mean of achieving this is to conduct such studies in the actual art galleries.

With the progress of technology, aesthetic phenomena could only just now be measured reliably outside the laboratories. In a recently completed 5-year long project titled 'eMobility - mapping museum experience' (Kirchberg & Tröndle, 2012; Tröndle, Greenwood, Bitterli & van den Berg, 2014; Tröndle, Greenwood, Kirchberg & Tschacher, 2014; Tröndle & Tschacher, 2012; Tröndle, Wintzerith, Wäspe & Tschacher, 2012; Tschacher et al., 2012), an interdisciplinary team of psychologists, artists, curators, and computer scientists aimed at identifying physiological changes which occur in the human body while viewing art in a real-life museum setting. The underlying idea being, that such physiological bodily reactions can be indicative (at least partially) of the emotional reaction to the aesthetic experience. In this study, 576 participants entered an exhibition '11:1 (—3) = Eleven Collections for a Museum' in the Kunstmuseum St. Gallen in Switzerland. At the entrance, potential participants were offered to wear a 'data glove' during their gallery visit. The device recorded the wearer's bodily reactions to art, such as *Heart Rate* and *Skin Conductance Levels* (Tröndle, Greenwood, Kirchberg & Tschacher, 2014).

These variables are known physiological correlates of the experience of beauty and cognitive implications of arousal. *Heart Rate* was previously shown to increase while expressive aesthetic situations are experienced, and decrease when novel information is presented. *Skin Conductance* can be indicative of mental and emotional activation, either positive or negative. Fluctuations in *Skin Conductance* correlate with arousal, which was shown to underlie information processing and decision making mechanisms (for a more comprehensive review see Tröndle, Greenwood, Kirchberg & Tschacher, 2014; Tschacher et al., 2012).

Additionally to having their bodily reactions recorded via the 'data glove', participants were also asked to complete an Entrance Survey, providing demographic details, and an Exit Survey, which asked about their thoughts on a subset of artworks from the museum. This included three artworks pre-selected by the researchers and three which caused significant physiological reactions to a given participant. By comparing results of the Exit Survey with a group of participants who was not offered to wear the glove, the authors demonstrated that wearing the device had no significant impact on the declared experience (Tröndle, Greenwood, Kirchberg & Tschacher, 2014). Thanks to a wireless tracking technology it was also possible to track participants' paths throughout the gallery and to assess their time spent in the proximity of each artwork. This, however, was done without considering the person's head orientation or the eye gaze direction.

Connecting significant fluctuations in physiological measures to the specific artwork was done based on the position tracking data. For each artwork, a rectangular 'affective region' in front of it was specified on the museum's floor plan. These regions were drawn arbitrary by a team of scientists and museum staff based on their opinion on 'how close a viewer must approach the work to observe it' (Tröndle, Greenwood, Kirchberg & Tschacher, 2014, p. 12). The physiological influence of each artwork was defined as the deviation of *Skin Conductance Level* or *Heart Rate* by more than 2% from the participant's global mean within a window of 2 sec. (Tröndle, Greenwood, Kirchberg & Tschacher, 2014, p. 15). Each time such a deviation occurred simultaneously to the tracking system reporting the participant's position within a particular picture's 'affective region', the physiological deviation was associated with the influence of that artwork.

Such an approach bears a major methodological flaw, which must be considered from the viewpoint of a thesis focusing on the influence of spatial layout on human responses to art. Firstly, being positioned in front of an artwork is not equivalent to viewing it, especially in a space densely filled with pictures and sculptures, many of which are likely to remain present in one's viewing field simultaneously. In the research method not involving mobile eye-tracking, this could have still been partially tackled by making path-based assumptions on the visitor's viewing fields (Lu & Peponis, 2014). Secondly, the size of 'affective regions' was different for each artwork, based on the expert opinion about its influence area (Tröndle, Greenwood, Kirchberg & Tschacher, 2014). Even if such judgment is accurate, different size of the regions means that a visitor taking a random path throughout the environment at a constant speed is more likely to stay within the 'affective region' of some artworks, compared to others. This would be registered by the biased system as a significantly different time spent in front of various works, even if in reality the amount of attention given to them was equal—e.g. up to 10 seconds of 'gist' followed by 30 sec. of diligent viewing, as it was suggested by J. K. Smith and Smith (2001).

Despite this methodological flaw, authors did report multiple statistically significant results. This might be explained by the fact that each 'affective region' specified on the museum's layout did include the area located directly in front of the artwork. Informal observations of visitor behaviour suggest that this area is likely to be the place of the most diligent interactions with the painting.

Principal component analysis was used to analyse the answers of the post-visit questionnaires consisting of 19 items[2]. Five factors were identified: *aesthetic quality*, *surprise/humor*, *negative emotion*, *dominance*, and *curative quality*. Physiological reactions to a specific artwork explained up to 25% of variance in scores on those factors. This means, that it could have been partially predicted how a given person will

2 Principal component analysis is a statistical method which transforms correlating items of a questionnaire into non-correlated 'principal components', or 'factors'. These are likely to reflect more general variables, less dependent on the biased understanding of each specific questionnaire item.

assess each painting on a 19-item long questionnaire of descriptive adjectives (Tröndle & Tschacher, 2012) based on the physiological reactions alone. It could be argued to what extent such a limited set of words can actually represent the 'aesthetic experience' with respect to all its complexity. The authors do, however, provide a satisfying conclusion in relation to art theories which typically argue agains attempts of quantifying the artistic phenomenon. The aesthetic experience, according to Tschacher et al. (2012) cannot be entirely reduced to known physiological measures - which is a finding opposite to neurobiological reductionism. Their results show, however, that to some extent, qualitative art theory discourse has empirical grounding and is a research direction requiring further investigations (Tschacher et al., 2012).

Other analyses, among other findings, revealed significant negative influence of conversations and walking in groups on the picture's impact (Tröndle et al., 2012), the influence of artwork's *dominance* factor on *Skin Conductance Levels* (Tröndle, Greenwood, Bitterli & van den Berg, 2014), and the lack of influence of picture's fame on its physiological impact (Tröndle & Tschacher, 2012). Yet, they cannot be generalised without caution, as they suffer from the methodological flaw described above. This did not allow the researchers to reliably assess the actual physiological impact of each individual artwork without separating it from the potential influence of other objects in space.

4.6 EXPERIMENTAL AESTHETICS AND THE CURATORIAL ARRANGEMENT

The authors of eMobility study also attempted to predict aesthetic judgment of individual artworks based on their location in space, which could be of direct relevance to this thesis' research question. Tröndle and Tschacher (2012) demonstrated that the rating of *aesthetic quality* decreased as participants moved further along their path through the linearly organised museum space. Instead of the intuitively assumed 'museum fatigue effect'[3] the authors suggest an alternative explanation: the curatorial organisation of the museum space only allowed for linear movement through it. At the same time, placement of artworks in subsequent subspaces was based on chronology. Tröndle and Tschacher (2012) suggest it is therefore not the 'museum fatigue effect' that caused works seen later to be judged lower, but the style of those artworks. The authors admit that modern art might be inadequately described by the questionnaire items contributing to the *aesthetic quality* factor (Tröndle & Tschacher, 2012). This aspect of the research setting remaining out of the experimental control makes it

3 'Museum fatigue effect' predicts decreasing interest in artworks encountered further along museum visitors' walking paths.

unachievable to draw any causal relation between the spatial location of a painting and its potential physiological impact.

Another approach administered by the authors of eMobility project involved making 4 curatorial interventions to the museum space, keeping the rest of the setting unchanged (Tröndle, Greenwood, Bitterli & van den Berg, 2014). Each situation was kept in place until at least 100 participants had the opportunity to experience it. This allowed for each group to grow in number in order to compare them statistically and indicate the effect of small spatial modification on the visitor experience across those groups. One such intervention involved hanging a single picture in the foyer, just above the entrance door to the 'official' exhibition. The title of the exhibition and part of the exhibition space could be visible through the doorway above which the picture was experimentally hung, but the difference in environmental context was immediately noticeable. Due to the lack of eye-tracking technology it is impossible to assess whether the picture was noticed by the visitors or not, yet the analysis of physiological data clearly showed that the distribution of arousal immediately changed *after* crossing the doorway. No significant pattern of physiological reactions was identified in front of it, where the experimental picture was hung. This situation was repeated with a different, very distinct, picture and it did not show a significant difference to a situation where no picture was placed above the doorway. This finding Tröndle, Greenwood, Bitterli and van den Berg (2014) interpret as an empirical evidence of the importance of the museum context. It is suggested that this context enables an 'aesthetic viewing mode' - a state of increased attention triggered by the awareness of being inside an art exhibition, where things are displayed to be viewed and experienced (Tröndle, Greenwood, Bitterli & van den Berg, 2014; O'Doherty, 1986).

Ignorance towards the *museum context* has been a major point of critique towards many laboratory-based studies in experimental aesthetics (Tröndle & Tschacher, 2012). Many other aspects, equally important to the aesthetic experience are ignored not less often. This includes the 'aura' of originality (Benjamin, 1968), the impact of originally intended curatorial arrangement (as opposed to a set-up developed by the researchers), or material and non-visual properties of the setting (Levent & Pascual-Leone, 2013b). This might explain why scientific findings in the field of experimental aesthetics are often ignored by the artistic community (Tröndle & Tschacher, 2012). Close collaboration with a working art museum and its staff allowed eMobility researchers to partially account for these factors.

Other curatorial interventions facilitated by Tröndle, Greenwood, Bitterli and van den Berg (2014) involved rehanging individual pictures on the walls in order to enhance their visual salience in relation to the surrounding paintings. For instance, a nude portrait was introduced in-between landscape paintings sequentially hung on a single wall (Tröndle, Greenwood, Bitterli & van den Berg, 2014). It was expected that—compared to the same wall containing the original landscape sequence—–the introduction of nude will create 'a centre of attention', indicated by intensified physiological reactions. Surprisingly, the opposite happened. The version of the wall with the nude, created a pattern of diffused arousal along the whole wall, whereas the original landscape-only sequence generated distinguishable 'centres of interest' in front of each picture. Referring to the art theory Tröndle, Greenwood, Bitterli and van den Berg (2014) suggest this situation might have been caused by the fact that hanging similar pictures next to each other gives them comparable, non-hierarchical prominence. This could 'suggest' the visitors to treat them as individual pieces, and therefore induce discrete reactions to each of them separately. Conversely, introducing a highly distinguishable piece might have created a hierarchical structure and 'suggest' the visitors to view the sequence of individual works as a connected one (in this case, like a triptych). This would potentially encourage viewing it as a whole, from the distance, and therefore from a more diverse set of possible locations, perhaps without even stopping in front of any of its parts for longer (Tröndle, Greenwood, Bitterli & van den Berg, 2014). As noted in historical reviews, such an effect can be further enhanced by long, narrow exhibition areas and rhythmically dispersed hanging locations (Newhouse, 2005). Due to the methodological limitations described earlier, this explanation could not be proven empirically by Tröndle, Greenwood, Bitterli and van den Berg (2014).

Based on a number of similar interventions, Tröndle, Greenwood, Bitterli and van den Berg (2014) conclude that highly salient works do not create individual 'force fields' (what authors call areas potentially prompting increased physiological reactions). Yet, they have the potential of doing it within a particular spatial arrangement. It remains unclear from this study, what would be the distinguishing element of such arrangements.

One generalisable finding resulting from these quasi-experimental spatial modifications was the confirmation of the *primacy effect* in the museum context (Tröndle, Greenwood, Bitterli & van den Berg, 2014). We reviewed this effect earlier in Section 3.2.1 in respect to picture memory studies. As eMobility researchers noticed, no matter which artworks were hung, and how the sequence of their hanging was shaped, it was usually the first painting of the sequence that attracted most distinctive physiological reactions in its vicinity (Tröndle, Greenwood, Bitterli & van den Berg, 2014). Yet again, without eye-tracking data it can be questioned whether it was the impact of the individual artwork, or if participants' level of arousal increased every time they approached a large number of new stimuli. In a linearly-organised museum space, the appearance of yet-unseen paintings in one's viewing field can be highly correlated with entering into a 'force field' of the first picture in a new sequence. A corridor-like structure of the Kunstmuseum St. Gallen with multiple perpendicularly aligned narrow subspaces creates high possibility of such encounters, while the visitor's head orientation after turning a corner is in fact unlikely to remain directed towards the nearest hanging picture. The need to compare this potentially biased finding with eye-tracking data was expressed by the authors themselves (Tröndle, Greenwood, Bitterli & van den Berg, 2014).

4.7 MOBILE EYE-TRACKING INSIDE MUSEUMS

Visual information is a predominant stimuli we direct our attention to in an art gallery setting. It can be accessed and interpreted at a distance and therefore remains partially dependent on, but not fully predictable from the participants walking path. This fact makes mobile eye-trackers a tool with unique potential for studying human attention inside museums. As we mentioned earlier, it is not entirely obvious to what extent our oculomotor behaviour reflects the underlying cognitive processes. However, the existing theoretical and empirical background constitutes the use of this method in a plethora of studies, in diverse contexts (Holmqvist et al., 2011).

Only recently, the availability of mobile eye-trackers has made it possible to utilise this technique outside psychological laboratories. In the museum context, for the first time this has been most likely done by Wessel, Mayr and Knipfer (2007). In this exploratory study the authors tracked eye movement of three students in a mini-exhibition on nanotechnology. The recorded scanpaths have been classified according to the 'category' they belonged to, where a single block of text or a separated graphic constituted a discrete category. Participants' eye movement was analysed in a predominantly qualitative manner, but some recurring patterns were identified. Firstly, participants seemed to analyse the related parts of the exhibition in a successive way, altering between thematically linked 'categories'. Secondly, two modes of looking at the exhibits were identified by the researchers: participants tended to visually skim larger area (e.g. an entire wall), to then analyse particular exhibits in this area in detail (Wessel et al., 2007). No details are provided on how these measures were operationalised, although such oculomotor behaviour would be consistent with the idea of a 'gist' followed by diligent viewing. This was earlier investigated in observational visitor studies (Section 4.1) and laboratory-based art viewing research (Section 4.4).

In a later paper further analysing this study, E. Mayr, Knipfer and Wessel (2009) go into details of some identified limitations of mobile eye-tracking in the context of museum-based research. *Obtrusiveness* *of measurement* is one of such concerns, as the fact that the person is wearing an eye tracker might affect how he or she behaves in space. This might not need to be so crucial in terms of the visual restrictions (the equipment used does not limit the viewing angles in a way any different from a large pair of glasses), as of the participants' awareness of their reactions being recorded. In a similar context, the eMobility project (Tröndle & Tschacher, 2012) has already shown that a highly sophisticated 'data glove' did not affect the way participants responded to art. Not only were the bodily responses gathered by the glove much more detailed, but eye movement is also a more controllable (and therefore less exposing) reaction than *Heart Rate*, or *Skin Conductance Levels*. Therefore the need for further control of this potentially confounding variable has been rejected in the current thesis.

Another concern related to the obtrusive presence of a mobile eye-tracker are the potential reactions from other gallery visitors (E. Mayr et al., 2009). In the context of this thesis, this is a viable concern in one out of three studies conducted - the one during which other visitors were present in the gallery. However, the nature of the potential experimental confound resulting from the dynamic visual, and non-visual, social interactions in a museum space lies beyond the scope of this work. Since the focus was on viewer-centred measures, it was controlled how often participants looked at other people. Yet, without a specially designed experimental procedure it is impossible to establish how other people's reaction to seeing the participant wearing a mobile eye-tracking device might have affected this participant's behaviour.

Other concerns regarding the utility of mobile eye-tracking for museum-based studies are related to its technological imperfectness (temporal and spatial inaccuracy), ethical concerns, and laborious data analysis (E. Mayr et al., 2009). In the current thesis, ethical considerations were tackled with ethical clearance from the departmental committee (Appendix C), and technological limitations were taken into consideration during the data analysis (Appendix A). The precision of the device was fully satisfactory for the required purpose, as we will see in Section 8.

In another early museum-based eye-tracking experiment, Heidenreich and Turano (2011, 2010) asked four participants to view fourteen paintings in Baltimore Museum of Art. The paintings were pre-selected by the researchers (in a randomised order for each participant), and the starting position for viewing each artwork was predefined in front of it. After opening their eyes at this starting position, participants were free to walk around the painting to adjust the viewing angle. Researchers were particularly interested in the specific painting areas which were investigated for longer. No significant correlations were found between eye fixations and the painting's 'saliency map' (Itti & Koch, 2000), nor between viewing times and the subsequent aesthetic judgments of the artworks (Heidenreich & Turano, 2011). It is not within the scope of this thesis to consider which elements of artworks are looked at more often, although one relevant finding was that those patterns changed over time. This is consistent with the idea of the aesthetic experience emerging in a dynamic, process-like fashion, as described in Section 4.4.

In a recent work, Eghbal-Azar and Widlok (2013) conducted a mobile eye-tracking study on two groups of 8 participants in two separate science exhibitions. Half of all participants were experts in the exhibition subject. Following the visit, participants were interviewed, and an opportunity to discuss the eye-tracking recordings was provided (a method called 'Retrospective Thing Aloud Protocol'). This allowed the visitors to contribute additional insights to the patterns of eye movement observed by the researchers. Despite not reporting any specific results and only a preliminary character of the provided analysis, (Eghbal-Azar & Widlok, 2013) notice the potential of mobile eye-tracking for discovering 'exhibition visit scripts'. The term is used to describe a sequence of viewing, repeating across visitors on the local or global level of the exhibition. The current thesis argues, that such visit scripts could be partially controlled by the spatial layout of the gallery.

Most recently, Brieber, Nadal, Leder and Rosenberg (2014) used mobile eye-tracking to investigate the influence of 'the museum context' on viewing artworks. The authors asked two groups of 21 participants to view the same art exhibition either in an art gallery, or on a computer screen in a psychological laboratory. In both conditions participants were allowed to spend as much time as they felt necessary looking at each picture. They were also free to approach pictures closer (or zoom on them in the laboratory-based condition), although the sequence of viewing was predefined. The authors hypothesised that more focused art viewing mode in the museum setting will cause longer viewing times per picture. As the results showed, even though time spent reading labels did not differ across the two conditions, the difference in time spent on viewing the actual pictures was significantly different, with participants inside the museum viewing them for longer on average. Viewing time was also associated with questionnaire answers regarding appreciation, ambiguity and understanding of the viewed art.

From the viewpoint of the current thesis, this works suffers from a significant limitation. The authors do not account for the spatial aspect of the museum experience in the computer-based viewing task. Two experimental conditions differ not only on the level of their 'context' but also in terms of the actions required to view each artwork. Thus, the causal mechanism between 'the museum context' (i.e. the awareness of being in the real art museum) and viewing pictures for longer cannot be drawn from this study. Alternative explanations associated with the spatial aspect of the experience are likely to account for the difference. For instance, the spatial setting most likely involves viewing pictures from multiple distances and angles, which can result in longer time necessary to process them. The speed and effort of locomotion are likely to further modulate this effect compared to a situation where all it takes to instantly zoom on a chosen picture is the press of a mouse button. The role of the current thesis in relation to the above study is to account for these spatial factors.

The relationship between appraisal of artworks and their viewing times demonstrated by Brieber et al. (2014) is also relevant as it indicates the top-down nature of visual attention during art viewing. In reality, however, this process is most likely reciprocal (Neisser, 1976): as much as higher appraisal of art causes longer viewing times (Brieber et al., 2014), longer viewing times might foster the development of the aesthetic experience and cause higher appraisal (Leder et al., 2004). Thus, the influence of spatial layout can take both 'bottom-up' and 'top-down' forms. The former occurs each time a human viewer, due to his or hers biological constrains, looks longer in a specific direction compared to any other direction in space. This is likely to occur, for instance, at all areas where longer navigation perpendicular to the wall's surface is required to progress through space. Due to the location of our eyes in front of the head, the 'default viewing mode' during human locomotion is to fixate on elements aligned with the plane of progression (Hollands et al., 2002). This fact could possibly be exploited by the curator to attract more attention to a specific artwork. The 'top-down' influence is likely to occur in two ways. Firstly, if the 'default cognitive task' of viewing pictures, which is associated with the gallery visit overcomes the default oculomotor behaviour during locomotion and modifies it. Secondly, when the prominence of a picture's spatial location is sensed by the viewer, thus arousing his or her interest and resulting in longer viewing times for that particular artwork.

5

THE INFLUENCE OF SPATIAL LAYOUT

Previous sections have looked at the importance of engaging the visitor into a meaningful, cognitive interaction with the exhibits and at how this can be assessed empirically through studying visitor's attention and memory. This section will review studies investigating how this engagement can be facilitated by the exhibition's spatial layout. This is especially important in the context of modern 'white cube' art spaces, deprived of any potentially distracting visual stimuli. In such an environment, the walls' arrangement is likely to be the most impactful tool available to the curator.

Firstly, we will briefly review the current practices and guidelines for art exhibition layout design and point out their non-generalisability. We will then consider methods which can describe space on the abstract level and help to generalise spatial properties of museum spaces. The most relevant methods will be reviewed in parallel with chosen architectural case studies of the existing museum spaces. The section will conclude with an evaluation of studies aiming at establishing a link between those abstract spatial properties and the human cognition.

It is this thesis' primary goal to establish such a connection in the context of the art gallery experience. Knowing how the relevant spatial properties influence human cognition in these spaces, can be a tool potentially beneficial to any curator interested in designing exhibitions with the goal of encouraging meaningful cognitive interactions between the visitor and the art.

5.1 CURRENT PRACTICES

There appear to be two primarily recognised functions for the art gallery's spatial layout considered in the curatorial theories (Vergo, 1989b):

1. to facilitate the curatorial narrative of the exhibition, or communicate a predefined message (e.g. Newhouse, 2005; Psarra, 2009);

2. to strengthen the viewers' focus on the artworks, or his/her 'aesthetic experience' (e.g. McLean, 1993; Pallasmaa, 2013; Tröndle et al., 2012; Psarra, 2009).

This impact of spatial layout is possible as the layout defines physical boundaries imposed by the walls on the visitors' visual and locomotive ability to explore the otherwise internally unrestricted, open space (Hillier & Hanson, 1984; Dalton & Conroy Dalton, 2010). In terms of navigational behaviour, this impact might for example modify the time required to travel through space (as the shortest path possible from entrance to exit is prolonged), or the number of possible alternative paths (e.g. Choi, 1999). In terms of the visual exploration, controlling the visual access to various subparts of the exhibition from individual points in space can influence, among other factors, the most likely sequence of viewing, the size of the surface area from which viewing is possible, or the potential co-visibility of multiple objects from a single point. Each of these potential influences can contribute to fulfilling one of the curatorial functions listed above.

Throughout the recent decades, the overwhelming majority of art galleries has utilised the idea of the 'white cube' outlined by O'Doherty (1986), and summarised so well in this paragraph:

> 'The work is isolated from everything that would detract from its own evaluation of itself. This gives the space a presence possessed by other spaces where conventions are preserved through the repetition of a closed system of values. Some of the sanctity of the church, the formality of the courtroom, the mystique of the experimental laboratory joins with chic design to produce a unique chamber of esthetics. So powerful are the perceptual fields of force within this chamber that, once outside it, art can lapse into secular status. Conversely, things become art in a space where powerful ideas about art focus on them.' (O'Doherty, 1986, p. 14).

Even in these visually ascetic environments the influence of the curator on the final visitor experience is substantial. It can be noticed that despite a very specific atmosphere achieved in individual 'white cube' art galleries, the layout imposed by their planners varies. Spatial layout provides room for curatorial interventions described (but not limited to) the two functions listed at the beginning of this section.

This layout's influence is perceived intuitively, despite the lack of initial empirical confirmation (and has been since the ancient times: Newhouse, 2005). For instance, many curators pay exceptional attention to the doorways and areas of transition between sections of the exhibitions (Dorsett, 2013, Personal Communication). This intuitive practice has only recently been reflected by empirical findings from neurological (Zisch, Gage & Spiers, 2013), psychological (Radvansky, Tamplin & Krawietz, 2010, 2011; Radvansky & Copeland, 2006), and architectural studies (Franz & Wiener, 2008).

Another carefully considered aspect in the curatorial practice is the degree of imposed linearity in the viewing sequence: while some artists see it as an unnecessary restriction, to others it is a necessity (Scheibitz, 2013, Public Lecture). Yet, the role of space is critical for each of these viewpoints. It can serve curators wishing to communicate a specific message by decreasing the alternative exploration options. Conversely, empowering visitors to incorporate their own interpretations on the equal basis to those proposed initially, requires even more considerate spatial planning. Otherwise, an unintended spatial hierarchy might be introduced to the relationships between individual artworks decreasing the number of possible interpretations.

As museums vary their exhibitions, they are often forced to plan their narratives long before the exhibits are physically built. Funding system of some countries requires every single metre of a new wall set-up to be planned even before the content of the exhibition is fully finalised (Saciuk, 2014, Personal Communication). Architects are therefore often asked not only to design a new museum, or convert an old building into a museum, but also to actively collaborate during the preparation of spatial layouts for larger exhibitions (Saciuk, 2014, Personal Communication). This has been occurring to the extent at which the boundary between the architect and the curator can vanish (Pallasmaa, 2013). Yet, whether it is the artist, the curator, or the architect who decides about the exhibition layout, the need to consider it long before specific artworks can be hung on the walls is evident. Hence the necessity to visualise or simulate design alternatives, and the need for guidelines which could aid this process.

A joint publication of four British Art Councils (Sixsmith, 1999) was aimed as a comprehensive guide to designing temporary art exhibitions. It lists and describes many elements of the exhibition design process. The complexity of this is understandable, as it has to include elements as distinct as transporting artworks to the gallery, ventilation facilities, and fire safety regulations. Among these, it also includes recommendations regarding the human flow through the building (Sixsmith, 1999). However, it remains largely technical and focuses on the accessibility requirements, which are an important, but not the only aspect the museum needs to spatially facilitate. The book also recommends an exercise of 'imagining' the flow of visitors through the building, and of the ways it might be facilitated. It mentions that the visitor experience should be enhanced by circulation routes which provide resting areas, as well as places of 'excitement and delight'. Ease of wayfinding is also emphasised (Sixsmith, 1999). However, the guide fails to recommend specific methods of simulation other than a rather non-scientific approach of imagining traffic flows. The question of how the delight, excitement, or rest can be facilitated by the architectural characteristics of a museum, remains unanswered in the guidelines (Sixsmith, 1999).

Another publication aimed at aiding the curatorial practice in user-centred exhibition design is that by McLean (1993). In the chapter dedicated to user-centred spatial planning she recognises three 'spatial conditions' critical in the context of the visitor experience.

First of them: *harmony*, is related to the orderly arrangement of the exhibition's elements. This is equivalent to—as the author compares it—the feeling of 'rightness' McLean (1993, p. 117) when things such as proportion, scale, and rhythm are meaningfully designed in agreement with each other. For instance, the role of scale rises when different objects are placed close to each other. As McLean (1993) notes, this is perceived by visitors as grouping (which is consistent with the spatial cognition studies; e.g.: McNamara, 1986). Therefore placing small objects next to each other supposedly enhances their relative importance as a unit of exhibition. According to McLean (1993), this affects the 'emphasis' which can be placed through design on chosen units more over others. According to the author, it is crucial to understand the balance between too many emphasised elements (which will distract and create chaos), and too little high-points (resulting in uninteresting experience). One of the most potentially impactful means of creating 'emphasis' is the position of the object in the exhibition layout (McLean, 1993).

Second 'condition of space' described by McLean (1993) is *atmosphere*, defined as multisensoral experience blending one's feeling of presence with the world of the exhibition (for the neurocognitive perspective on this phenomenon see: J. Ward, 2013). As mentioned earlier, the current thesis focuses on the visual aspect of the museum experience.

Third 'condition of space' distinguished by McLean (1993) is *pacing*. The author defines this element as the combination of the visitor's movement through the exhibition, his or her interaction with it, and the total time spent inside (we have reviewed studies focused on this particular aspect in Section 4.1). McLean (1993) stresses the importance of spatial decision points, and considering visitor's fatigue during the planning process. She relates to earlier research mentioning the general tendency to turning right, as well as the observed preference to take directs paths towards the exit. In respect to considering traffic flows during exhibition planning, she writes that 'the only fixed rule is that people will surprise exhibit planners with their actions' (McLean, 1993, p. 123). As we will show later in this section, this statement is untrue. Human movement and viewing patterns can be partially predicted and planned for, at least on the aggregate level.

Contrary to her own statement she distinguishes between linearly sequenced exhibition layouts and those allowing for freedom of exploration (McLean, 1993). She provides examples of how each of them can guide visitors' movement, but does not suggest any single solution, mentioning that usually each exhibition is the combination of those different approaches.

McLean (1993) also lists a number of practical suggestions in relation to planning the exhibition layout. For instance, she suggests special awareness towards 'hot spots' or 'focal points' (also named by her as 'landmarks') inside a museum. These places should be used by curators to 'guide' visitors through space. She does not however provide a definition of such a 'landmark' and does not give any specific advices on how to use them (McLean, 1993). When it comes to the number of objects within an exhibition, emphasising a single object is suggested to enhance its visual importance (McLean, 1993). On the contrary, McLean (1993) reviews studies which showed that people prefer to look at more complex presentations. A balance is recommended, but no clear answer given (McLean, 1993). Another practical suggestion is to carefully design vistas and highlighted areas. Long lines of sight have the property of drawing people's attention (and consequently, their movement) towards new places (McLean, 1993). Noting the fact that space can guide visitor's experience is the most important contribution of McLean's textbook from the viewpoint of this thesis. As she writes, these spatial aspects can often prove more powerful than explicit signage (which, although useful in many situations, often remains unnoticed McLean, 1993).

Despite exemplifying the interest of Visitor Studies movement in the spatial layout of the exhibitions and making a number of points which will be empirically shown valid later in this Section, McLean (1993) is not able to propose generalisable conclusions. She neither can offer any analytical methods which could be repeatedly used to simulate and design new exhibition spaces. Her important work rose the awareness of the problem but did not offer a solution.

5.2 a formal description of the museum space

The eMobility Project described in Section 4.5 (Tröndle, Greenwood, Bitterli & van den Berg, 2014; Tröndle & Tschacher, 2012; Tröndle et al., 2012; Tschacher et al., 2012; Tröndle, Greenwood, Kirchberg & Tschacher, 2014) was perhaps the most technologically advanced and the largest in scope attempt to advance methodologically-restricted observational visitor studies. It established a linkage between viewing art in the real-world museum setting and human physiological reactions to this experience. This linkage is indicative of the dynamically changing emotional reactions as the visitors freely progress through the exhibitions.

Among multiple valuable results, and despite the explicit effort, the project failed however to establish any generalisable connection between the spatial arrangement of the exhibits and its influence on the visitor experience. Similarly disappointing can be the fact that, despite their almost a century-long tradition, observational visitor studies (Section 4.1) have not led to the emergence of layout-related guidelines which could be repeatably applied in the design of new exhibitions (despite suggestions that this is an important area for curators' consideration, e.g. McLean, 1993). One of the reasons for this limitation remaining yet to be tackled is the operationalisation of spatial influence at a rather generic level, such as total exhibition area, or the number of artworks displayed (Robinson, 1928; Melton, 1935; Serrell, 1997); and in a holistic manner, where all properties of the considered spatial setting contribute to its 'atmosphere' (McLean, 1993). The underlying theoretical view of this thesis is that the 'atmosphere' of any space can be described by a set of abstract properties. The extreme view on this would assume that any space can be described by a set of variables together with their values, and subsequently replicated anywhere else. The 'atmosphere' of such a replication, together with the emotional reactions and behavioural patterns of its visitors would be predicted to be equal to the 'original'. Metric dimensions are the example of such variable probably the most familiar to anyone. However, the previously described research on spatial cognition has put in doubt whether this is the most psychologically relevant aspect of space we are sensitive to (Section 3.2.2). Space can be described by a set of less obvious variables, potentially much more relevant to how humans perceive and comprehend it.

One of the methods allowing for formalising the aspects of space which are relevant to the visually rich experience of an art gallery exploration, is *Space Syntax*. Conceived by Hillier and Hanson (1984), it initially linked the formally described built environment with the social interrelations of its dwellers. Methodologically, Hillier and Hanson (1984) proposed to represent the layout of the environment in the form of a graph. Such a form, independent of metric parameters, allowed the researchers to calculate mathematical and logical relations between parts of a larger layout. It also allowed them to compare multiple layouts against each other on the dimensions represented by these formal calculations. Many of the original Space Syntax measures have been refined, or modified for new purposes. Most importantly however, this formal way of thinking about abstract aspects of the built environment constituted a platform for the development of new measures. Let us now take a closer look at the underlying methodological contributions being of the highest relevance to this project.

We will start with the concept of *isovist*, which is older than the theory of Space Syntax itself (for a comprehensive historical reference see: Montello, 2007) and has been popularised under this name and

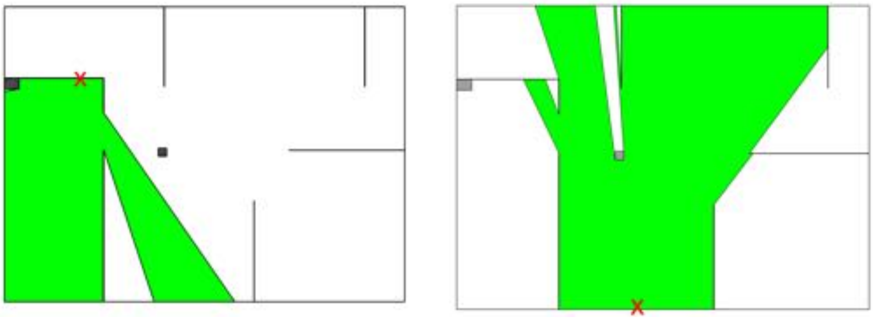


Figure 5.1.: Two sample isovists derived from separate locations; differing by area and shape.

usage by Benedikt (1979). Isovist is a 2-D or 3-D polygon bounding the area potentially visible from the reference point located in that space. This thesis will consider 2-D isovists only, as the majority of works contributing to its theoretical and operational development. Since modern art galleries tend to be uniformly differentiated on the vertical axis, it is unlikely that 3-D analyses would yield different results for the cases considered in this thesis.

The primary measure describing any isovist is the size of its *surface area*, which is indicative of the amount of space visible from the reference point. The fact that some defined area can be seen from the reference point, corresponds to the fact that this reference point can be seen from all the surface area bounded by the isovist. Calculating *isovist size* for a picture hanging on a gallery wall will therefore describe the metric size of the gallery space from which that picture may be viewed. This value can then be compared to isovist area values generated from other points in space (for example, for other pictures hanging in the gallery).

Similarly to the area size, multiple other mathematical properties of isovists can be calculated, just as they could be for any geometrical figure. For instance, shape can be mathematically described with the measures of *area-to-perimeter ratio* and *point 2nd moment* (which is the variance of line lengths drawn from the reference point to the polygon's boundary). The former variable will increase, and the latter will decrease as the isovist approaches a circular shape and therefore they describe how 'spiky' the isovist is. A 'spiky' isovist would result in a less stable (but potentially more varied) visual experience, where the connection between the viewer's eyes and the object of reference is potentially often interrupted, for instance by walls or columns. We will consider these and other properties in detail later.

Figure 5.1 presents two sample isovists derived from the same layout.

Benedikt (1979) also presented a preliminary assumptions for the Information-Field Theory, which aims to describe the cognitive exper-

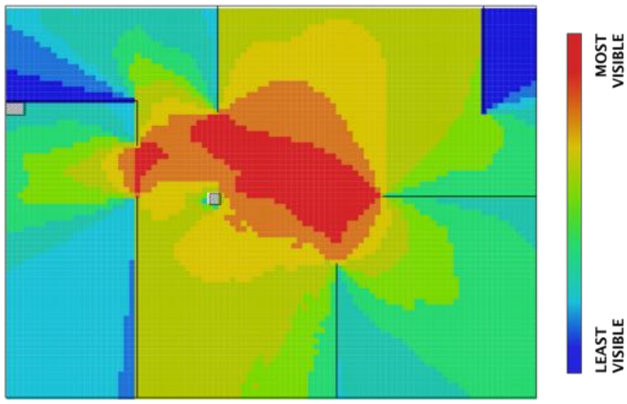


Figure 5.2.: Visibility Graph Analysis derived for the same layout as above, showing Isovist areas of all points in the graph.

ience of moving through space by the change of isovists generated from the points on the traversed route. The author gives the example of Frank Lloyd Wright's Guggenheim Museum in New York. He notices how the famous ramp linking multiple floors throughout the interior removes any sudden changes in the isovists' area sizes. This, as the author states, puts emphasis on the art presented inside rather than the building itself, although carries the danger that the spatial experience will become 'vanishingly boring' (Benedikt, 1979, p. 63).

This particular notion—of dynamically changing experience being described by the amount of information visible from the points along the traversed path—has been developed further. For instance, e-spaces and s-spaces representing visually stable parts of the environment have been proposed by Peponis, Wineman, Rashid, Kim and Bafna (1997) but do not seem to have gained much following. Up to date, the most significant expansion of the concept of isovist and the Information-Field Theory has been proposed by Turner and colleagues (Turner & Penn, 1999; Turner, Doxa, O'Sullivan & Penn, 2001). They proposed to calculate multiple isovists generated from equally spaced locations and describe them relatively to all other isovists in that space (a process known as *Visibility Graph Analysis*, or *VGA*). Visibility Graph is calculated by superimposing a grid on the analysed spatial layout. For each grid cell, the software 'counts' how many grid cells can be visible from this point, given the walls' restrictions. This results both in a mathematical measure of *Connectivity*, as well as its visualisation, where areas visible from (and connected to) the largest proportion of space are coloured red, and those which 'see' the least (and concurrently can be 'seen' from the smallest proportion of space) are coloured blue (Figure 5.2).

For the same reference point, its *Isovist Area* is therefore correlated with its VGA's *Connectivity* value, the only differences being the resolution, the method of calculation, and units used. When derived from a reference point lying on the layout's boundary (e.g. a paint-

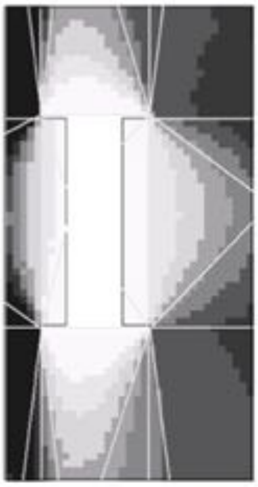


Figure 5.3.: 'Clustering coefficient' values in Visibility Graph Analysis overlaid with 'e-spaces'. Original black and white figure adapted from Turner, Doxa, O'Sullivan and Penn (2001). with permission from Pion Ltd, London (www.pion.co.uk / www.envplan.com).

ing hanging on a gallery wall), both of these measures become a 180°isovist directed perpendicularly to the wall surface. A similar, subsequently developed analysis is *Boundary Visibility Graph (BVG)*. The difference between VGA and BVG comes from the fact that BVG calculation only considers grid cells lying along layout boundaries and therefore its *Connectivity* value is *not* equivalent to the location's *Isovist Area*, but instead reflects the proportion of other *wall surfaces* visible from the reference point. Turner et al. (2001) also demonstrated how VGA corresponds to Peponis' e-spaces in a gradient-like, rather than discrete step-wise manner (e.g. Figure 5.3).

Visibility Graph Analysis and the accessible software package used for its calculations (Turner, 2001) opened new avenues for the development of the concept. Novel abstract descriptions of space, based on its visibility properties can be proposed in respect to the particular research aim. For instance, Lu and Zimring (2012) developed *Targeted Co-Visibility* - a measure, which described *not* the total area visible from the reference point, but the number of visible 'visual targets' relevant to the hospital's intensive care unit personnel (e.g. patient beds). So calculated Visibility Graph would then contain grid cells differed on that dimension, providing also the visualisation of the most strategically important parts of the intensive care unit's layout. In the art gallery context, this measure can describe how many pictures (which constitute natural 'visual targets' for our visual attention in this environment) are potentially visible from any point in space (see Figure 5.4).

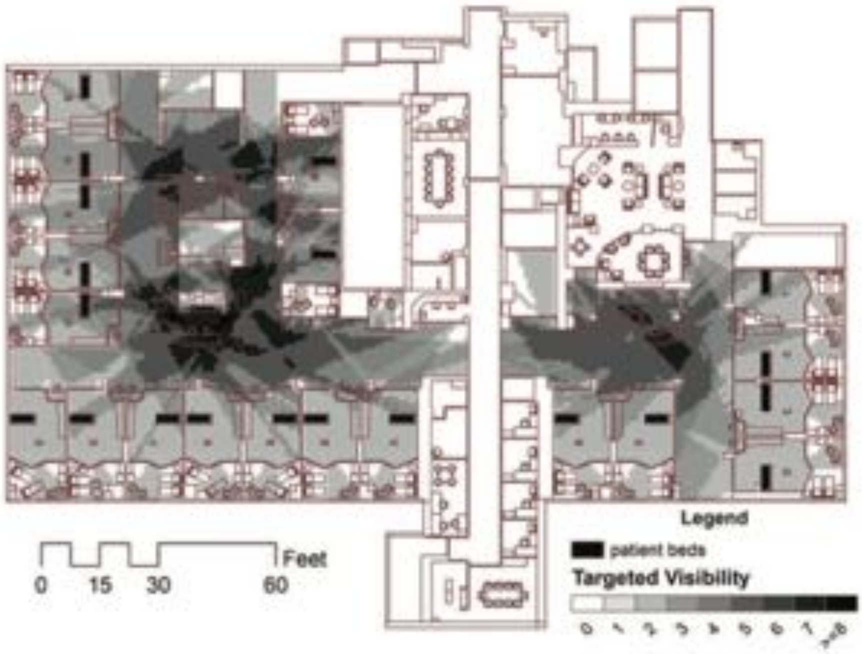


Figure 5.4.: 'Targeted visual connectivity' (in this paper referred to as 'Targeted Co-Visibility'.). Darker colours show higher number of visible patient beds in a hospital. Adapted from Lu and Zimring (2012).

The often discussed disadvantage of this approach is that isovists are usually calculated for a 360°viewing field, and therefore describe *potentially*, but never *actually* visible space/targets (as no person can simultaneously explore the entire visibility field surrounding him or her). An alternative methodology has been applied by Lu and Peponis (2014) to tackle this problem. In this study, they designed a number of virtual reality based art galleries differentiated by the pattern of co-visibility of pictures belonging to the same theme. The authors empirically showed that participants who explored those virtual galleries were sensitive to the co-visibility patterns when answering to the subsequent questionnaires about the clarity and understanding of the exhibits. As the experiment was conducted in the virtual set-up, participants' paths could be recorded and taken into account during the calculation of co-visibility measures. It was demonstrated that those measures, which took the path and orientation of the visitors into account, did better at predicting their responses compared to the path- and orientation-independent measures based on the assumption of 360°viewing field (Fig. 5.5).

This finding shows that co-visibility measures are of direct relevance to the curatorial design needs. It must be noted however, that for the purpose of predicting new layout's influence, their potential precision is limited. Lu and Peponis' study used path-dependent

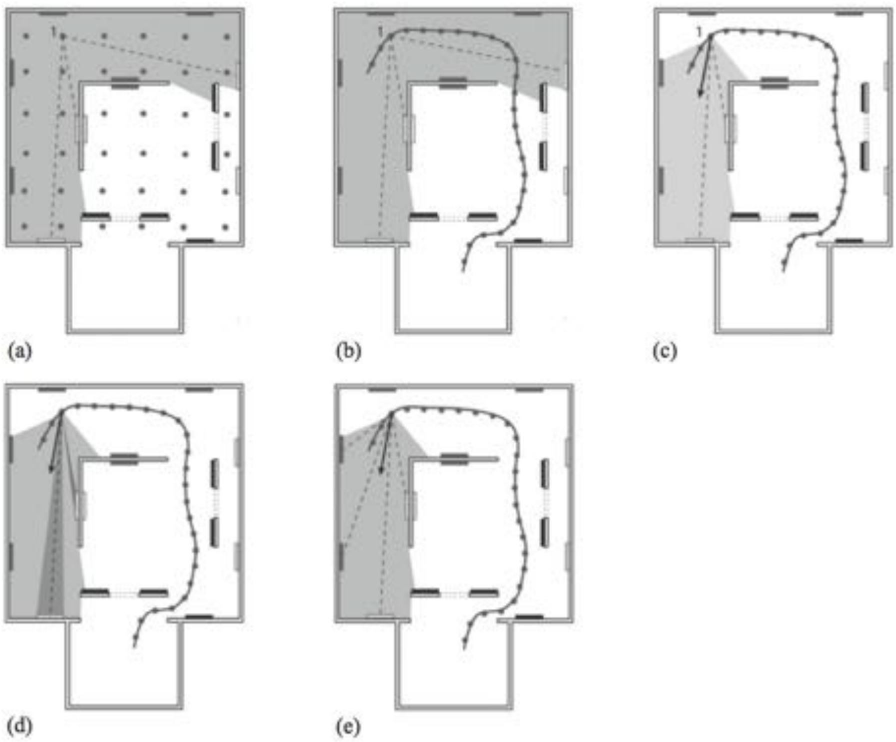


Figure 5.5.: Five types of 'co-visibility' calculated (from a-e respectively) with more detailed input describing participant's movement throughout the virtual gallery. Adapted from Lu and Peponis (2014) with permission from Pion Ltd, London (www.pion.co.uk / www.envplan.com). For a more detailed description refer to the source material.

measures, i.e. they required the exact data on the viewer's trajectory in order to make assumptions about the most likely 'visual targets' falling into one's viewing field. With the use of mobile eye-tracking device, such factual co-visibility could be established with even greater level of precision, post-visit. The aim of this thesis, however, is to develop measures which will allow curators to make partial simulations of the yet-unbuilt exhibitions, prior to the visit. For such a purpose, the exact trajectories can only be estimated, and even if this can be done with some satisfactory level of precision (Hillier et al., 1996), further assumptions of the head orientation would likely make those predictions rather unreliable. For this reason, the 360°-based assumption considering *potential*, but not *actual* co-visibility patterns can simply reflect the theoretical limitation, beyond which no prediction of highly variable human visual behaviour can be made without overfitting the model [1]. Despite the fact one could argue that some co-viewing events are less likely to occur (e.g. the viewer instantly turning around to compare two paintings hanging in front of each other from the location between them), they are still possible. Estimating their exact probabilities would require a specially designed empirical studies, results of which are likely to highly vary across contextual circumstances, and therefore lies beyond the scope of this thesis. For this reason, despite the fact that the current work will consider individual viewing sequences of gallery visitors, the environmental measures used will be path-independent and based on 360°viewing fields. While the observed and recorded viewing behaviour can be used to speculate *why* a phenomena occurred, the potential real-world application for design predictions must be based on *potential* and not *actual* viewing situations.

5.3 SPATIAL RELATIONS INSIDE MUSEUM BUILDINGS

This way of thinking about museum space has been used in a variety of architectural case study analyses based in museum buildings. Understanding how distinct museum buildings organise exhibitions in relation to the available space has always gathered considerable interest from Space Syntax researchers. This might be the result of the specificity of this building type indicated by Psarra (2009). Following the ideas developed by Markus (1993), she emphasises the evolution of museums as sites of 'visible knowledge'. She makes the distinction between the 'well lit warehouse' (referring to neutral galleries, structures of which could serve any function), and the idea of integrating the building with the displays. As she notes, the latter trend charac-

1 In Statistics, this term is used to describe a situation in which a random component of the predicted phenomena (or 'error') is being mistaken for its generalisable component. The resulting 'overfitted' model therefore predicts a highly specific variation of the generic situation. Even though the model 'fits' the available data better, it will be less successful at predicting other occurrences of the same mechanism.

terises the museums of the late twentieth century. This illustrates the changes in architectural and curatorial ideologies (Psarra, 2009), and shows the need for grater consideration of museum building design, if they are ought to be fit-for-purpose.

In this section we will review architectural case studies of the existing museum buildings conducted with the analytical methods of Space Syntax. We will focus on those studies, where the main scope was on revealing the non-trivial architectural relations between the museum building and the exhibitions contained in it. These works were important for the development of relevant techniques and served as a source of inspiration to other researchers. Later, empirical attempts have been made to link some of these techniques with their impact on human behaviour and cognition. Those attempts will be reviewed in the last section.

The pioneering work in this respect has been conducted by Peponis and Hedin (1982). The authors compared the spatial organisation of two distinct exhibition spaces at the British National History Museum. One of them was designed in the late XIX century, and the other has been a subject of major reorganisation in the late XX century. Following their analyses, the authors concluded that the layout of those spaces reflected the beliefs about enunciation, social organisation, and transmission of knowledge from those two epochs. Spatial layout has been shown to be closely linked to the beliefs popular in the given period. This finding has directed the attention of other researchers to the potential cognitive impact of museum spatial layout. A plethora of museum-based case studies followed.

For example, Tzortzi (2003) conducted isovist analysis on the extension to the National Gallery London - the Sainsbury Wing. One suggestion made by the author was that the thematic and aesthetic relations between the artworks (as opposed to their uniqueness) were spatially promoted by the distribution of their visual fields. Paintings also tended to be located at the ends of long lines of sight (Tzortzi, 2003), which often overlooked over a number of rooms. These analyses qualitatively reflect the later quantitative operationalisation of *co-visibility*, described in Section 5.1.

In a further extension of this study, Tzortzi (2007) included more buildings to demonstrate how these spatial practices differentiate individual museums. The distinction is made between those, which impose some pre-defined meaning upon the visitor, and those allowing for a rather unrestricted exploration and interpretation (Tzortzi, 2007). This particular problem has been similarly noticed in the context of science museums: Sue Allen (2004) noted the dilemma between allowing visitors to freely explore the space in an informal manner while making sure they still learn anything. As Sue Allen (2004) suggested, this should be done without enforcing a strict sequence of educational material and producing cognitive overload during interaction with the exhibits. The studies conducted by Tzortzi (2007) show that the patterns of visibility and co-visibility might be of particular importance to this aspect of exhibition design.

Psarra (2006, 2009) modelled visibility patterns in multiple museums to investigate how curatorial narrative is linked to the exploration patterns. The distinction she makes based on those comparison is between communicating a specific message and emphasising the aesthetic experience.

Another link has been suggested by Zamani (2009) in the analysis of 3 layout alternatives implemented over time on the second floor of the High Museum of Art in Atlanta. It is demonstrated, that as the curatorial philosophies changed, so did the spatial and visual relations in the galleries. The major differentiating factors are suggested to be the visual hierarchy and cross visibility. Their role was to encourage, or discourage the visitors to focus their attention on a specific artwork (or theme), as opposed to making connections across discrete sub-categories.

These three comparative studies (Psarra, 2009; Zamani, 2009; Tzortzi, 2007) with the use of formal spatial descriptors of Space Syntax analyses confirmed the distinction drawn in Section 5.1 from the curatorial literature (Vergo, 1989b). There seem to be at least two major curatorial goals that spatial layout can facilitate: strengthening the visitors' focus on individual exhibits, or encouraging comparisons and cross-categorial interpretations. The contribution of the above described studies lies in demonstrating that those two goals tend to polarise across the entire museum building, or at least its separate floors. Two main reasons might determine that. Firstly, a museum building is a 'physical container' for the museum institution. And these institutions tend to represent a particular set of beliefs and understanding of their public mission. Consequently, they might intend to organise the building's spatial layout in a manner complying with these beliefs. Secondly, as the above studies showed, visibility and co-visibility patterns are the most evident means used to achieve these curatorial goals. The interplay of these visual properties might be difficult to moderate on the local level if two contrasting approaches are adopted within a single building. For instance, limiting or maximising co-visibility can only occur in respect to multiple other objects in space.

An important study touching directly on the impact of visibility and co-visibility of specific artworks in respect to the entire environment was conducted by Stavroulaki and Peponis (2003). In it, the authors took a closer look at the spatial organisation of Castelvecchio Art Museum in Verona, Italy, redesigned by Carlo Scarpa. The authors draw attention to the pedagogical aspect of art exhibitions, which can be understood as a function of maintaining the boundaries between the categories of knowledge. Through this mean, exhibitions structure and guide a viewer through the meaning of exhibited objects using various spatial, organisational, and segregational methods. The operationalisation of this concept in the paper is the notion that the way artworks are positioned influences the path of the visitors. An example of sculptures facing different directions is given to illustrate this point (Stavroulaki & Peponis, 2003). Since museums are places where we come to look at art, it is reasonable to assume the visitor's willingness to see the exhibits from comfortable viewing locations. Therefore, positioning sculptures in certain locations and orientations increases the chance of attracting the visitors to specific areas; even in an open-plan exhibition setting where the number of potential movement paths is infinite. This method, however, never fully determines the visitor's experience, as the artworks are still perceived in various sequences, combinations, and from multiple angles, depending on the path chosen. Novel interpretations can still be discovered.

Stavroulaki and Peponis (2003) also notice that the positioning of artworks can be used to guide building users to the 'key' locations in space, important for comprehending the spatial organisation of the exhibition. Such a 'magnet' can be even more meaningful and powerful than the one provided by spatial characteristics of the physical layout (such as the relative accessibility of a specific subspace from other subspaces Stavroulaki & Peponis, 2003). This is a significant claim, as it suggests that moderating visual relationships by the spatial layout of an art gallery is a function more important from moderating movement through its physical boundaries.

Stavroulaki and Peponis (2003) describe their exemplary case study of Castelvecchio to justify their suggestions. In this gallery, multiple artworks are left freestanding in the middle of a room. A visitor desiring to view all artworks must complete a very specific movement pattern. This results in constant engagement with sequential juxtapositions of paintings, purposefully placed so by the curator, and yet still allowing for the freedom of exploration and variability of the experience.

Stavroulaki and Peponis (2003) also notice that there is a certain distance and angle range at which paintings of different sizes are most comfortably viewed, and therefore their placing can encourage the visitor to walk closer or further away from the wall. The authors propose the analysis of an isovist restricted to a 60°cone generated from the centre of each painting. They suggest this to be the optimum viewing location, in which every viewer might wish to position him- or herself in order to look at the picture comfortably. The importance of viewing each picture from the distance prior to approaching it is also highlighted, although the lack of cognitive or behavioural measures make this assumptions only intuitional.

As the works reviewed above showed, in the museum context the patterns of visibility and co-visibility can have a predominant influence on the resulting experience. The fact that world famous museums reviewed in the architectural case studies differ on this dimension shows the importance of the visual factor in the curatorial theory. Visitor movement can only be treated as a predictor of the amount and type of information one encounters in the surrounding and potentially cognitively processes. It therefore increases the probability of a cognitive engagement with the exhibit but is not equivalent to it. Drawing a direct connection between the spatial layout and human cognition is a distinct problem, and therefore this thesis will only review those visitor movement studies where attempts were made to link it with the cognitive experience.

One such a study was as conducted by Wineman, Peponis and Conroy Dalton (2006). Although only observational techniques were used (together with all their drawbacks: see Section 4.1), the observations were conducted for two separate travelling science exhibitions, each in 2 dissimilar spatial set-ups. This quasi-experimental design allowed for some limited generalisation of the findings in relation to the abstract aspects of the spatial layout (contrary to typical observational studies described in Section 4.1, or Space Syntax analyses of distinct museum cases reviewed above). Additionally to recording movement, the authors made an attempt to link its patterns with visitor engagement by counting stops (see Section 4.1 for the potential problems with this technique). Not surprisingly, spatial accessibility of exhibit elements was linked with the likelihood of the visitor's engagement.

More importantly however, the authors noticed that sequence-driven curatorial narrative is not the only aspect of exhibition design which contributes to the users' engagement. Assuming higher variability of the visitor movement in these open-plan exhibitions, more flexible spatial set-ups still are likely to facilitate the final experience with the means such as thematic grouping of the exhibits visible from other exhibits. Similarly to Stavroulaki and Peponis (2003), this finding shows the potentially superior influence of visibility and co-visibility properties moderated by the spatial layout, from the movement restrictions it imposes.

Importantly, Wineman et al. (2006) showed that the patterns of the layout's influence were similar for those visitors, who stayed inside for shorter and longer periods of time, but the consistency and size of these effects was higher for those, who spent more time inside. This finding was explored further in an overview paper by Wineman and Peponis (2010). Their view on the spatial organisation of museum spaces suggests that the visitors' experience is not barely the sum of individual interactions between the visitor and the exhibits, sequenced in a particular way by the spatial layout. The influence of the layout reaches much further through the creation of multiple possibilities for the understanding of exhibition hierarchy. This is most often achieved through the configuration of viewing fields that allows for seeing the same exhibitions in multiple ways and in juxtaposition with other displays (Wineman & Peponis, 2010). While free choice is still given to the visitor as to the specific pattern of his or her viewing sequence, the space can *suggest* some patterns over others and therefore not be a subject of a completely random exploration (Wineman & Peponis, 2010). The authors make the distinction between *spatially determined movement* and *spatially probabilistic movement*. In their opinion, human exploration of the museum space is somewhere between these two sides of the spectrum, and therefore can be referred to as *spatially guided movement*. This means, there is a large amount of randomness attached to the movement of humans inside museum spaces, but it is not completely random. Instead it is *guided* by the (primarily visual) properties of the exhibition layout (a property earlier noticed but not deeply analysed by McLean, 1993). Therefore, spatial layout can support the curatorial narrative, but not as a mere framework for sequencing the viewing experience. This idea has been identified in the curatorial theory as potentially more valuable, because it empowers the visitors (L. C. Roberts, 2004). Wineman and Peponis (2010) show the importance of researching the impact of visual measures which are reflecting relations between discrete exhibits and their relative standing in the entire gallery space. This is especially important when the cognitive outcome is of interest, as—contrary to the assumptions of the Visitor Studies movement—it is likely to follow patterns often not equivalent to the path taken through the museum space.

5.4 LINKING SPACE SYNTAX WITH SPATIAL COGNITION

Each of the aforementioned research took place in a separate gallery, containing exhibitions varying significantly from each other in highly uncontrolled conditions. This fact makes the results incomparable and gives those case studies barely an exploratory status, which prevents the researchers from deriving generic conclusions applicable to the museum architecture as a whole. Despite these limitations, they identified main spatial strategies differentiating the existing museum buildings (Psarra, 2009; Zamani, 2009; Tzortzi, 2007) and recognised visibility and co-visibility patterns as the crucial aspect moderated by the exhibition layout, potentially affecting the visitor's movement through and interpretation of the exhibition (Stavroulaki & Peponis, 2003; Wineman et al., 2006; Wineman & Peponis, 2010). Further importance of these studies lies in the fact that the abstract visual and spatial measures thereby used can serve cognitive science as independent variables describing experimental environments. Such methodology has the potential of establishing direct and generalisable linkage between definable, quantifiable, and predictable aspects of the built environment with their impact on the human cognition (Conroy Dalton, Hölscher & Turner, 2012).

The number of known studies joining the Space Syntax-inspired independent variables with psychological experimental design (as opposed to case studies) is limited. Not coincidently, the majority of those reviewed below used virtual art galleries as their research setting. One potential reason for it is the fact that barely any other building type similarly exemplifies the connection between the spatial layout, the spatial distribution of visual stimuli, and the cognitive aim of processing these stimuli (Lu & Peponis, 2014). We will now review these known studies, although the reader should note that some of them used 2-D Virtual Reality environments. The potential limitations of these approaches have been signified in Section 4.4.

The already mentioned study on picture co-visibility by Lu and Peponis (2014) (see Section 5.2) has been an example of such an attempt. Despite interesting design, it suffered from a number of methodological flaws, which put in doubt its psychological findings other than the fact that humans are sensitive to exhibit co-visibility. For example, all participants were architecture students, who are likely to perceive space differently and distinguish greater variation in the environment compared to laypersons (Montello, 2007; Gifford, Hine, Muller-Clemm & Shaw, 2002). Moreover, the cognitive outcome of the experimental conditions was measured with an arbitrary chosen set of questionnaire questions, the internal and external validity of which has not been assessed. Therefore, the specific nature of how this cognitive outcome is influenced by picture co-visibility and the casual relationship of these variables remain unexplained.

A small number of laboratory-based studies were specifically designed to establish this causal relationship between the visual properties of space (mainly isovist and its derivatives) with human spatial and oculomotor behaviour. As the focus of these works was on the cognitive processes and their linkage with quantifiable aspects of space, their main aim was not to aid curatorial theories; despite some of them being set inside museum settings.

Wiener et al. (2007) conducted a series of exploratory experiments in such a virtual art gallery. Sixteen spatial variations of the setting's layout were prepared. They were described by such isovist properties which could be potentially related to some aspects of the human perception of space (like the feeling of spaciousness, openness, and complexity). The method of calculating these properties required some modifications compared to the original concept (Benedikt, 1979), as isovist originally describes a local, rather than global spatial property; a property which is relevant only for a single point of reference in space. This study required a more generic measure in order to describe the architectural shape of the layout. To achieve this, the authors used a method of calculation similar to VGA (Turner et al., 2001) and averaged the resulting isovist metrics for the entire layout. In so designed virtual environments, participants were asked to explore them on a computer and to perform a navigational task. The task was to find the best 'hiding' place and the best 'overview' place in the gallery—both of which had previously been identified mathematically by the researchers using local isovist properties such as size. Following this phase, participants were also asked to provide questionnaire responses rating the subjective qualities of the environment, such as its pleasingness and interestingness. A number of significant correlations were found between the isovist properties, behavioural performance, and environmental ratings of the participants. Wiener et al. (2007) showed that (1) people are sensitive to local isovist properties and can correctly recognise places in the layout with the largest and smallest isovist area sizes; and (2) that human judgment of the generic 'look and feel' of space can be predicted by isovist-derived descriptions of its shape.

This finding provides direct empirical support for the concept of *spatially guided movement* proposed by Wineman and Peponis (2010). Since humans are indeed capable of detecting visibility properties of space as they are described by the isovist theory, these relationships might really be able to 'guide' our behaviour in and cognition of space.

While the above study investigated human behaviour and environmental judgments, visual attention was the subject of a comprehensive series of experiments by Wiener, Hölscher, Büchner and Konieczny (2012) (also: Wiener, Hölscher, Buechner & Konieczny, 2009). In it, participants were presented with a series of screenshots from a virtual environment consisting of two spatial choices (left or right, with a separating wall in the middle of the screen). Their task was to make navigational decisions in search of a target object hidden somewhere in the environment while their eye-movement was being recorded. Sequential presentation of static images simulated progressing through space. As the results showed, participants' behaviour presented strong preference for this navigational alternative, which could be described by a longer line of sight. The possible explanation——representing the 'top-down' view on human behavioural strategy—is the promise of higher informational gain favouring such a decision. Moreover, some consistent patterns in the participants' eye movement were shown to be dependent on the geometry of the visual stimuli (Wiener et al., 2012). Contrary to some studies described in Section 3.1.3 investigating the 'bottom-up' approaches to visual attention (e.g. via 'Saliency Maps'), here the authors fully controlled for the default viewing mode of 2-D images and were able to reliably demonstrate that the oculomotor behaviour observed in their experiments was task-specific.

In the context of the navigational search task, Wiener et al. (2012) therefore showed the impact of geometrical properties of space on human visual attention. They demonstrated the perceived shape of the environment—which can be described with the concept of isovist—can guide how we look at, and move through space. Emo (2014) further expanded this work with real-life pictures of street corners and linked it to the Space Syntax-derived concept of an 'axial map'. 'Axial map' describes which city streets are, on average, more accessible from the entire network of streets in the city. Despite the different contextual background of these experiments to the current thesis, they remain highly relevant. They contribute to the finding that human spatial behaviour and cognition are sensitive to—and can be predicted by—some quantifiable visual properties of space (Lu & Peponis, 2014; Wiener et al., 2007, 2012; Emo, 2014), but the nature of this relationship is likely to be determined by the context of the participant's task (Land & Tatler, 2009; A. Johnson, 2011).

Controllability of the virtual 2-D environment in the above series of experiments comes at a cost of decreasing its ecological validity. However, they do provide a reliable theoretical basis for testing the influence of isovists in the real-life environment.

One such attempt was made by Dalton, Marshall and Conroy Dalton (2010) who carried out an exploratory study to show that isovist properties can impact human memorability for the publicly displayed content. The authors placed four projection screens in a large hall of a public building and used them to present a standardised set of highly memorable words and images for the span of three days. The locations of those projectors varied in terms of their isovists properties (spiky/round and small/large isovist areas). Following the period of experimental presentation, building users received an invitation to an online survey and were asked to answer questions about words and images presented throughout the study period. The results showed that more pictures were remembered from words if presented from large isovist areas, and words were remembered better if presented from small isovist areas. Additionally, pictures proved to be more memorisable if presented from spiky, and not round, isovists. Despite the preliminary character of the study acknowledged by the authors, this suggests that the isovist properties of the stimuli's location might influence the memory of humans spontaneously traversing the area around them. However, as the authors note themselves (Dalton et al., 2010), forming more general claims would require a strictly controlled experimental space than a public building with fluctuating user numbers moving through it with different spatial knowledge and behavioural aims.

Building on the methods described so far, this thesis will employ a set of tools and workflows previously applied in separation in disciplines described in the previous sections. Next chapter will provide an overview of the most relevant ones, selected for the research conducted within this thesis.

Part III

METHOD OVERVIEW

This thesis presents three studies conducted inside three distinct environments. Two of them were experimental settings where the researcher had the possibility to alter pictures' spatial relation based on the research question—those will be referred to as 'Experiment 1' and 'Experiment 2'. The third study took place in a working art gallery, namely the BALTIC Centre for Contemporary Art in Gateshead, UK. We will refer to the latter case as 'BALTIC Case Study' since it did not involve any experimental manipulation and only a single group of participants was invited to explore it—all within exactly the same spatial set-up.

As the data-gathering and analytical methods used in all three studies were very similar, this section will present their overview and provide theoretical background for their adaptation to the current work. Those details which are distinct between the studies will be described in further, relevant sections of the thesis.

6

PROCEDURE

The procedure employed in each of three studies consisted of the following phases:

1. Participant arrives to a scheduled meeting outside the gallery space.

2. Participant signs informed consent form (Appendix C.1), a demographic information form (Appendix C.2) and completes an excerpt from Ishihara's test for colour blindness (Ishihara, 1917, Appendix C.3).

3. Participant is briefed about his or her task (Appendix D.2). They are asked to explore the gallery 'just as you would explore any other art gallery' whilst wearing a mobile eye-tracking device. It is indicated that they should look at least once at each picture, but there is no instruction regarding walking trajectory, speed, or revisiting the already seen artworks.

4. The eye-tracker is calibrated (this procedure takes about 30—60 seconds). Exit and meeting points are indicated. The participants walks inside (see Fig. 6.1).

5. After exiting the gallery the eye-tracker is demounted, the visitor is asked to take a sit in front of laptop computer, which was not open prior to the visit.

6. First of the two unanticipated memory tests follows (Recognition Memory Test; Section 9). Instructions are presented on the screen.

7. After completion, another test is presented to the participant (Spatial Memory Test; Section 10).

8. Visitor is debriefed, paid £6 for participation and provided an opportunity to leave an email address in order to receive further updates on the study's results.

Standard ethical procedures were employed throughout the study (Appendix C).

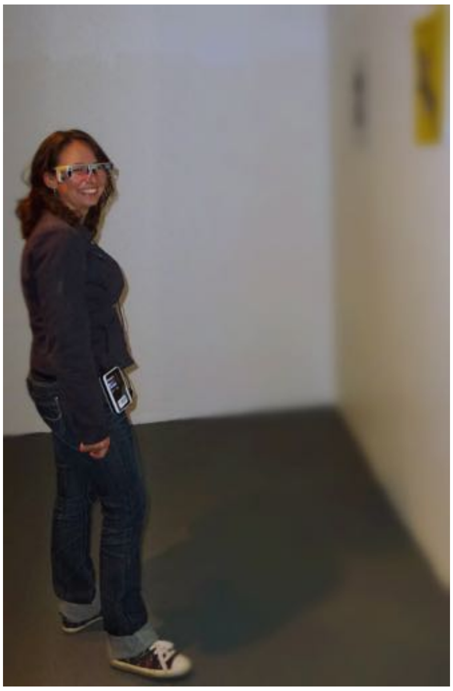


Figure 6.1.: Sample participant wearing an eye-tracking device inside the gallery.

None of the two memory tests was previously mentioned in the research invitation or the briefing phase. Informal conversations following the study confirmed that participants were convinced that collecting eye gaze recordings was the only purpose of the experiment. This fact is important as the thesis' focus is on *spontaneous* memory and not deliberate memorisation of pictorial stimuli. Primary reason being the fact that knowing one will be tested enhances memorisation (McDermott & Arnold, 2013) and is not representative of the cognitive goals typically associated with an art gallery visit. Participants were also unaware of the existence of alternative experimental conditions[1].

For the purpose of the study, it was equally important to preserve the 'restorative' (Packer & Bond, 2010), relaxing atmosphere of a museum environment (Kirchberg & Tröndle, 2012) by avoiding additional pressure potentially associated with participating in a scientific experiment. For this reason, time limits given to the participants were as liberal as possible and depended primarily on the battery capacity of the eye-tracking device. To avoid the opposite effect of rising expectations of unnaturally long viewing times, the briefing instruction

1 From the ethical viewpoint, it is important to note that the procedure did not involve providing misleading information but only delayed the disclosure of some of its part. Participants were able to withdraw from the study at any moment, including the memory tests. None of the participants reported any concerns when asked to participate in the additional memory tests following the visit.

emphasised that participants can spend 'as little or as much time as they wish' inside the gallery. To further facilitate the lack of pressure, the researcher did not follow the participants into the experimental space.

This methodology, by going beyond observational studies of the visitor behaviour, attempts to identify the quantifiable cognitive outcome of a gallery visit. The accessibility of mobile eye-tracking is what seems to have been of particular desire to many researchers in that field the past (Section 4.1). Additionally to collecting more detailed and more reliable data than majority of previous visitor studies, this methodology provides one more advantage: altering spatial set-up across experimental conditions within Experiments 1 and 2 reaches beyond the typical case-study approach and provides the possibility to generalise the results more reliably. The use of quantifiable descriptors of space derived from Space Syntax and Isovist theories made it possible to conduct those spatial modification based on the specific research questions. Such methodology answers an explicit, often repeated call for more rigorous studies of the art gallery experience and its dependence on generalisable factors of the physical environment (Bollo & Dal Pozzolo, 2005; Bourdeau & Chebat, 2003; Lu & Peponis, 2014; Bitgood, 2003; Wineman & Peponis, 2010; Locher, 2011; Tröndle, Greenwood, Bitterli & van den Berg, 2014).

Table 6.1 presents an overview of the methodology employed in the thesis.

Variables	Method of Data Collection and Analysis
Independent Variables	
Quantitative descriptors of space, separately calculated for: Experiment 1 (2 distinct spatial configurations), Experiment 2 (2 distinct spatial configurations), BALTIC Case Study (1 spatial configuration)	Space Syntax and Isovist analysis (Depthmap, CAD software); Section 11
Demographic information (e.g. age, gender)	Demographic information form
Oculomotor behaviour	Mobile eye-tracking; Section 8
Dependent Variables	
Oculomotor behaviour. Note that oculomotor behaviour can be treated as both dependent and independent variable: it can be the aim of analysis to explain where and how people looked at pictures depending on their spatial location, but also to explain participants' memory of pictures based on how they looked at them.	Mobile eye-tracking
Memory of objects	Recognition Memory Test; Section 9
Memory of objects' location	Spatial Memory Test; Section 10

Table 6.1.: Methodology overview.

7

PARTICIPANTS

To further ensure ecological validity of the method, special care was given to population sampling. It has been widely acknowledged (e.g. most recently by Brieber et al., 2014) that controlled art gallery experiments should ideally be conducted on spontaneous museum visitors, and not on homogenous groups, e.g. of psychology students. This has been an infamous issue of many psychological disciplines (Peterson, 2001), but it might bear a particular importance in the context of art gallery studies. The reason being, art is often created with the aspiration of producing the widest spectrum of experiences possible. Thus, the variance of behavioural and cognitive patterns during the exploration of art space is likely to be larger than in other common everyday situations. Homogenous study groups consistently underrepresent this variance (Peterson, 2001).

This challenge has been tackled in the current thesis by recruiting participants from multiple sources, thus ensuring that the variability of the actual population visiting art galleries is represented in a more adequate manner. However, even though participation was voluntary, it cannot be fully ensured that all of the participants of the studies here described would equally likely visit such an exhibition if no experiment was taking place. It must be acknowledged that some proportion of visitors was attracted by paid reward, and some by the novel possibility of using an uncommon mobile eye-tracking device.

Another issue with population sampling for art studies is the fact that the most eager potential participants are artists themselves. Artists, curators, and people with educational background in art share higher interest in the subject in general. In his early studies, Robinson (1928) already noted that artists' way of exploring an exhibition might significantly differ from the rest of population—a tenet which has been later confirmed with more advanced measurement methods (Pihko et al., 2011). Presuming that the goal of public museums is mainly to serve *the public* (as opposed to a situation where artists would exhibit art for other artists), this thesis will focus on non-expert art viewers. For this reason, no artists or art graduates were allowed to participate—this statement was included in the research invitation and verified again during the briefing phase.

As the research invitation mentioned that the study investigates 'associations between spatial layout of art galleries and human perception of an exhibition' (Appendix D.1), it was equally possible that a group who might exhibit over-representative eagerness to participate would be architects. If indeed the fact that the study focused on 'spatial layout' would be a strong attractor to a group professionally interested in the subject, it would be possible that they pay special attention to the layout of the gallery, unlike they would when visiting a museum in order to view art. Thus, in a similar manner to artists, architects were not allowed to participate.

Normal, or corrected-to-normal vision was required. Contact lenses and correcting glasses were allowed, as long their frames were not very large (due to the restrictions of the eye-tracker). Normal colour vision was required and tested prior to the gallery visit.

As previous knowledge of space might affect spatial memory (e.g. Caduff & Timpf, 2008), participants who visited the experimental space prior to the study were also excluded from the analysis. In spaces where layout was modified for the purpose of the exhibition (Experiment 2 and BALTIC Case Study), this was only relevant if space was visited under the same spatial configuration.

Advertising invitations in sources external to the university ensured large spread of participants' age—another factor shown to have a significant, negative impact on memory (Wai, Lubinski & Benbow, 2009).

8

MOBILE EYE-TRACKING

'Tobii[1] Glasses 1' monocular mobile eye-tracker was used in all three studies. The device consists of a large plastic glass frame, a wire and a recording unit saving data on a SD-type flash memory card. The recording unit is lightweight (200 grams) and could easily be attached to a belt or placed in the visitor's pocket, depending on the preference. The frames were large enough to fit on glasses of these participants who wear any. If this caused any discomfort, priority in positioning the device was given to the participant's ease of wearing, even if this compromised the quality of the recording. The glass frame contains two cameras—one points in the same direction as the user, and records the scene in 640x480 pixel resolution at 30 frames per second (Fig. 8.1). The other camera tracks the wearer's right eye using infrared light with the frequency rate of 30 Hz. The estimated eye position is then superimposed on the scene video in the post-processing phase within Tobii Studio software package. The superimposed coordinates are based on the calibration process proceeding each gallery visit. After the calibration phase participants were asked not to touch the device, as this could have introduce a consistent spatial shift in all fixation estimates. If the quality of calibration was compromised (e.g. by the user's discomfort, or unintentional movement of the glasses), the videos were excluded from further analysis. All participants entered the gallery wearing the eye-tracker and they were not informed of potential calibration problems, as this could be only established during the post-processing stage, via the desktop software. Consequently, all participants viewed the gallery through the device, and even if the eye-tracking data was compromised, the memory test results could be included in the analysis of the complete dataset, on the equal basis to participants whose calibration process was successful.

However, even with successful calibration the device was not robust in the presence of highly varied, often very dynamic eye-movement associated with the naturalistic exploration of a vast spatial environment. During the post-processing stage, Tobii Studio software indicates the proportion of each participant's fixations that it managed to successfully record. The producer's suggestion is to exclude those

1 Parts of this section were previously published in (Krukar & Conroy Dalton, 2013a)

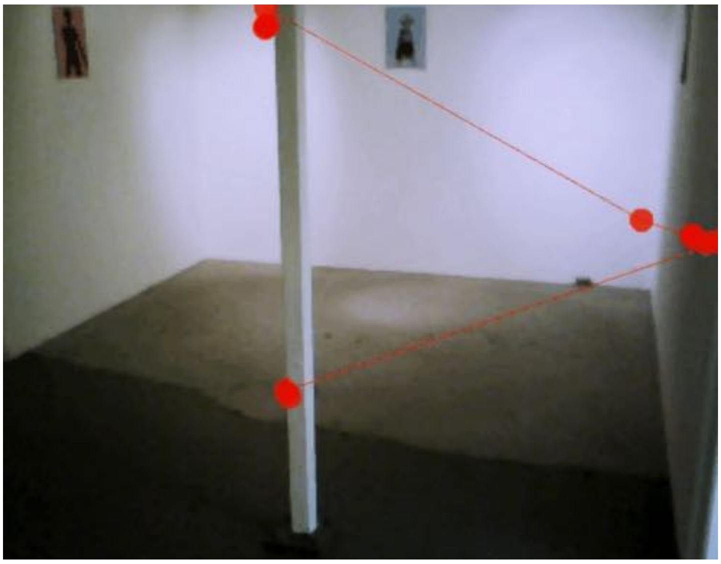


Figure 8.1.: A sample frame from the Tabii Glasses recording.

recordings from the analysis that did not meet the 60% threshold of recorded fixations. The factual patterns of oculomotor behaviour might be wrongly represented by the recorded subset of so low quality (Tobii webinar series, 2014). Such recordings were excluded from all eye-tracking analyses.

8.1 ESTIMATING EYE GAZE

It is important to note the technical limitations shared by many eye-trackers available on the market. The most common misconception is the understanding that eye-trackers *detect* fixations. This is untrue. Eye-trackers detect pupil's movement and its temporary suspensions. Based on this movement, and a built-in algorithm, the software *estimates* the occurrence of fixations. The estimation is primarily based on the velocity of eye movement and the length of its pauses. The software then matches the estimates of fixation locations with the scene video based on the pre-indicated distance between the eye and the viewed object (in case of monocular eye-trackers). Therefore an eye-tracker does not 'know' where a person has fixated—fixations are just estimations based on processed data (Oakes, 2010).

The real length of a single fixation can span from 50 to 500 ms, and its mean values oscillate around 200—275 ms, depending on the task (Rayner, 1998). The device used for the current experiments samples eye movement at 30 Hz, which is equivalent to taking a 'snapshot' of the current position of the pupil every 33 ms. Consequently, fixations can only be recorded by Tobii Glasses at multiplies of this time period. Hence, almost every fixation estimate made by the device will contain 'temporal sampling error' on top of its real length (Andersson, Nyström & Holmqvist, 2010). This thesis will refer to these biased estimates as 'fixations', yet the reader should be aware of the technological limitation standing behind the impreciseness of the concept.

8.2 EYE-TRACKING MEASURES

Temporal sampling errors accumulate throughout the study and in some cases may lead to highly unreliable total length estimates (Andersson et al., 2010). For this reason, using fixation-based measures recorded with sampling rates as low as 30 Hz would result in a misleading level of accuracy, not reliably describing the actual oculomotor behaviour. To decrease this distortion of results, fixations were reduced to the measure of *dwell*. A single dwell constitutes an event when eye gaze moves into the boundaries of a distinct object (in our case—a picture), explores it through numerous fixation-saccade sequences, and redirects outside the boundaries of the object (Fig. 3.1, p. 44). This operationalisation reduces the cumulative impact of temporal sampling error, as in such case the error only influences the timing of the first and last fixation of each dwell. It is also sufficient for guiding the answers to the research question stated in this thesis - its interest does not lie in the specific patterns of viewing within individual pictures, but in the influence of space on frequency and lengths of visual encounters with each picture as a whole.

Dwell-based approach to eye movement analysis has been previously explored in multiple studies and comprehensively reviewed by Holmqvist et al. (2011). Table 8.1 lists the variables obtained from the recordings in the current thesis. The selected measures aim to describe the oculomotor behaviour in terms of its quantity (e.g. how long/how often visitors looked at objects?), type (e.g. were the dwells long, or short?), and dynamics (e.g. were the dwells followed by immediate fixation on another object?). For procedural and methodological details of the coding procedure please refer to Appendix A.

Measure	Definition	Reason for Including	Previous Usage in the Field
Total Dwell Time	The sum of all dwell lengths within the boundaries of a single picture during the entire visit.	Previously linked to the meaning of the viewed area, its informativeness and interest (Holmqvist et al., 2011), as well as the memory of object's position (B. W. Tatler, Gilchrist & Land, 2005).	Established measure. Often defined as 'total fixation length'.
Number of Dwells	The count of all dwells falling onto a single picture throughout the entire visit. Equal to the number of events when participant's eye gaze, entered an object's boundaries and at least one fixation occurred.	Previously related to object's informativeness and the confidence levels in object recognition (Holmqvist et al., 2011). Encountering the same object on multiple occasions updates its representation in long-term memory and, cumulatively, gives rise to individual differences (Caduff & Timpf, 2008; B. W. Tatler, Gilchrist & Land, 2005) in its memorisation.	Established measure used previously (Holmqvist et al., 2011), but not in any of the eye-tracking studies reviewed in this thesis. Also used in observational visitor studies, although most likely with lower precision.

Measure	Definition	Reason for Including	Previous Usage in the Field
Normalised Total Dwell Time	The percentage of time a visitor spent looking at each picture relative to the whole time spent inside the gallery.	Compared to 'total dwell time', this operationalisation is insensitive to differences in time spent inside by each participants. Consequently, it might reflect the subjective significance of each picture relative to all other objects.	Previous usage unknown; measure introduced for the purpose of the thesis.
Normalised Number of Dwells	The percentage of all dwell events relative to cumulative number of dwells on all objects.	as above	Previous usage unknown; measure introduced for the purpose of the thesis.
Time to First Fixation	Difference between the time of the first dwell occurring on the given picture and the time passed from entering the gallery.	Indicative of the sequence of visual encounters with novel stimulus.	Popular measure in a different context (Human-Computer Interaction studies). Adapted for the purpose of the thesis.
Mean Dwell Length	Average length of a single dwell calculated for each participant and for each picture.	Potentially indicative of different behavioural patterns across varied spatial conditions.	Its estimates were previously used in observational studies.
Median Dwell Length	Median length of a single dwell calculated for each participant and for each picture.	Potentially more accurate measure of central tendency due to skewed distribution of the data.	Previous usage unknown; introduced for the purpose of the thesis.

Measure	Definition	Reason for Including	Previous Usage in the Field
Number of Short Dwells	Number of dwells shorter than 2 seconds.	The threshold of 2 seconds was suggested in the model describing two-stage processing of artworks (Locher et al., 2007) when holistic gist of the scene is extracted. Dwells shorter than that could be indicative of a mode of processing different to diligent picture viewing. A less clearly defined barrier of 'few seconds' was also suggested by Bitgood (2010) in his 'capture-focus-engage' model of museum-based attention. 'Focus' stage occurring below the period of few seconds does not, according to the model, involve deeper cognitive processing of the artwork.	Previous usage unknown; introduced for the purpose of the thesis.
Number of Long Dwells	Number of dwells longer than 2 seconds.	see above	Previous usage unknown; introduced for the purpose of the thesis.

Measure	Definition	Reason for Including	Previous Usage in the Field
Number of Very Long Dwells	Number of dwells longer than 10 seconds.	In observational studies of art gallery visitors, this period of viewing has been suggested to constitute a threshold between slow, 'walking-and-viewing' mode and stopping in front of artworks in order to examine them in detail (J. K. Smith & Smith, 2001).	A spacial-case subset of 'long dwells' category. Previous usage unknown; introduced for the purpose of the thesis.
Long-to-short dwell ratio	Number of long dwells divided by the sum of (short dwells and long dwells) for a given participant/stimuli. A value of 0 implies no long dwells at all. A value of 1 means that all dwells were long. A value of 0.5 suggests that the same number of long and short dwells was performed.	Given a heavily skewed distribution of dwell lengths, this is a better approximation of the dynamics of the oculomotor behaviour than mean dwell length. Higher value indicates a higher proportion of long dwells and smaller value is a result of eye movement consisting primarily of dwells shorter than 2 seconds.	Previous usage unknown; introduced for the purpose of the thesis.

Measure	Definition	Reason for Including	Previous Usage in the Field
Picture-Switching	The proportion of dwells which were followed by a quick, immediate fixation on another stimuli, without an intermediating dwell on a picture-free area (such as an empty wall surface).	Operationalisation of the concept of 'spotty-focused attention'. Differentiates between visual interactions which form a middle part of a continuous chain of engagements, and those which are disconnected from the next cognitively demanding dwell.	Previous usage unknown; introduced for the purpose of the thesis.
Engagement Ratio	Cumulative dwell time on all pictures divided by the time spent inside the gallery by a given participant.	Describes the proportion of time each person remained engaged with the stimuli while being inside the gallery. Can potentially describe the individual difference in the level of interest each visitor had in looking at artworks compared to engaging with alternative activities (such as examining constructional details of the space, one's own watch, or other people in the BALTIC Case Study).	Previous usage unknown; introduced for the purpose of the thesis.

Measure	Definition	Reason for Including	Previous Usage in the Field
ET Sequence	Rank-order of appearance in the eye-tracking dwell records.	Describes in what order participants saw each picture, i.e. in what order they performed the first dwell on each artwork.	Various classifications of fixation order are common in scan-path analyses.

Table 8.1.: Description of the eye-tracking measures.

8.3 VIEWING SEQUENCE SIMILARITY

On the most general level, eye tracking recordings make it possible to establish what sequence participants viewed images in. This raises a question whether viewing sequences of individual visitors were similar. Irrespectively of what image was hung at each location, it can be investigated whether participants were more likely to visit some parts of the exhibition in a similar, sequential manner. For instance, it can be expected that artworks located closer to the gallery entrance will generally be viewed at the beginning of the visit, while those located near the exist shall be engaged with closer to the end of the visit. However, this simple analysis can be expanded to establish more complex patterns of sequential viewing. In this thesis this will be based on the analysis of *sequence similarity*. The question asked is whether viewing sequences are similar across participants and, if so, are they more similar within separate experimental conditions than they are across the conditions.

Only *long dwells* were taken into account to define viewing sequences, as those—from the theoretical point of view—are likely to contribute to a qualitatively distinct aspect of the gallery visit. They are likely to describe engagements which are not the result of orienteering, but of diligent investigation of an artwork leading to the formulation of aesthetic judgment. After removing *short dwells* and dwells on objects other than the analysed artworks, multiple situations were identified when the same artwork has been the subject of engagement on more than one occasion in a sequence. These instances were considered as a single engagement. Short interruptions and the dynamics of the viewing behaviour will be a subject of separate analyses, while here the attempt is made to identify 'exhibition visit scripts' defined as the sequence of artworks visitors were likely to diligently explore in a similar manner. It is of little interest at this moment to inquire what was the 'quality' or the dynamics of each particular engagement. The question asked is whether the sequence of likely meaningful viewing can be 'pre-programmed' by a curator, and how various spatial arrangements affect the stability of this 'curatorial program' across the individuals.

In order to assess the similarity of viewing sequences a method developed by Conroy Dalton (2001) will be employed. It is based on a string-matching technique from mathematics and computer science - specifically, the 'Levenshtein Edit distance' (Levenshtein, 1966). This technique takes two ASCII strings as input and returns a value describing how many operations (the possible operations being: inserting a character, removing a character, or substituting a character) is required at minimum to transform one string into another. Low 'Levenshtein edit distance' between two strings means that two strings are quite similar to each other - for example transforming 'ABC' into

'ABD' would only result in value of 1. The higher 'Levenshtein edit distance' is, the more dissimilar the strings (e.g. consider operations require to transform 'ABC' into 'AXQYZ' resulting in the value of 4).

Each location in the sequence has been given a unique ASCII character (either a capital or small letter, as these are considered distinct symbols in ASCII). Based on that, a unique string was constructed for each visitor, representing this person's viewing sequence - a set of interactions with individual locations, each longer than 2 seconds, ignoring any potential interruptions in these engagements[2]. Following the method developed by Conroy Dalton (2001), the resulting strings were compared with each other (within the same experimental condition) using a two-step procedure. Firstly, each string was compared to each other string using Levenshtein distance algorithm from `stringdist` R package (van der Loo, 2014). As viewing sequences had different lengths, it was important to consider that the same 'Levenshtein edit distance' could indicate a radically different level of similarity for two strings which are short, and for two strings which are long[3]. For this reason, each so extracted 'Levenshtein edit distance' was divided by the average of the length of two strings, resulting in *normalised Levenshtein distance* (Conroy Dalton, 2001). The same operation was conducted for all pairs of analysed strings. Secondly, *normalised Levenshtein distance* scores associated with each viewing sequence (and its combinations with all other sequences) were summarised and divided by the number of (all - 1) strings in the sample. This created *mean normalised Levenshtein distance* (*MNLD*) describing how similar a giving string is to all other strings. Low value would indicate that a string is very similar to all other considered strings. Higher *MNLD* indicates that the string is rather unique. A custom R script was written to perform these operations based on a C program provided by the original author (Conroy Dalton, personal communication).

8.4 task-specific oculomotor behaviour

When considering the nature of human oculomotor behaviour, we stressed the importance of cognitive goals in guiding eye movement (Section 3.1.3). An important assumption underpinning the validity of the method employed in this thesis is that these goals are uniformly understood by all participants. In the situation created, participants

2 Note that for this analysis we do not consider a path defined by the interactions with pictures, but defined by the interactions with artwork locations. In other words, the strings describe with some level of precision how the individual was exploring the space of the gallery, ignoring what specific pictures were hanging at the explored places at moment of engagement.

3 The reader might wish to consider the value 4 from the example above, which resulted from two strings which had only one symbol in common. The same value of 4 would indicate much higher similarity if both strings were about 10 characters long.

were expected to 'explore the gallery as they would explore any other art gallery' and to 'make sure they look at each picture'. This carries the assumption that a deliberate emphasis is placed by the visitors on visual engagement with the artworks. Experimental procedure would not be valid if other cognitive goals are predominant instead. For instance, if the space was too confusing, the need of finding the exit might dominate over the idea of relaxed art gallery visit and result in a qualitatively different pattern of oculomotor behaviour acquired for navigation (Wiener et al., 2012; Emo, 2014).

Some of the earlier reviewed laboratory-based experiments had the opportunity of accounting for the 'default viewing mode'. Wiener et al. (2012) for instance controlled the experimental condition against an inverted version of the presented images. This provided a differentiation between the experimental manipulation (in this case: spatial search), and a situation were no such task was given, or assumed. For obvious reasons, potential experimental modifications to the way the environment is viewed are limited in the real-life setting. For example, viewing perspective cannot be inverted on the vertical axis.

In the real-life setting, introducing a group controlling for the 'default viewing mode' is therefore much more problematic. The primary reason for this is the fact that giving participants any other tasks in a gallery-like environment does not ensure they will not attend to pictures while performing it, thus potentially switching between two viewing modes. In this case, introducing a non-obvious goal for an art gallery visit (e.g. find your way out) would most likely disturb, but not fully replace the pattern of visitors' oculomotor behaviour. Since viewing artworks is the 'default' mode of behaviour in any gallery-like setting, it is very probable that humans—once inside it—naturally pay exceptionally high attention to all discrete, rectangular, visual stimulus hanging on white, otherwise empty walls. Introducing non-art place-holders instead of actual art is unlikely to change this behaviour on the gallery-wise level, as the design of 'white cube' spaces is ideologically (O'Doherty, 1986) and historically (Camfield, 1990) linked with the artistic character of any object seen inside it—no matter how trivial it would be—and successfully aimed at heightening visitor's attention (Alpers, 1991; Tröndle, Greenwood, Bitterli & van den Berg, 2014).

Inviting participants to an 'art gallery study' and asking them to explore space completely deprived of any visual stimuli is equally unlikely to provide the desired control situation. The awareness of being in a gallery-like space (with white, empty walls and rectangular geometry) and the knowledge of having one's eye-movement recorded is a strong contextual cue most likely prompting 'the visual search' mode of viewing—complexity of which constitutes its own distinct subject in perception studies.

How can we then ensure that participants employed a strategy focused on looking at paintings, and did not simply randomly walk around the environment for some time with their eye-movement *not* guided by the locations of artworks? In the real-life situation, in the presence of contextual cues as strong as they are in a white cube art gallery, this is a purely theoretical concept. However, a completely random eye movement during a random exploration of the environment would have to result in the gazes falling *on* pictures according to the proportion of total wall surface area occupied by those pictures. Given that some wall surfaces are more visible than others, such a calculation could be adjusted based on the visual accessibility of the pictures' positions. In three studies presented in this thesis participants spent overwhelming proportion of time fixating on pictures, thus allowing to make an assumption about their task-specific eye movement patterns. Informal observations of the eye-tracking recordings further confirm this point—none of the recorded participants seemed to exhibit 'non-gallery-like' behaviour. Sections 15.2, 19.2 and 23.2 contain detailed statistical analyses confirming this observation for each experimental condition analysed in the thesis.

9

RECOGNITION MEMORY TEST (REACTION TIMES)

The method involved two separate memory tests. The first of them was Recognition Memory Test, similarly to the procedure incorporated by Janzen (2006), Miller and Carlson (2011). All instructions and stimuli were presented on a laptop computer screen with OpenSesame software (Mathôt, Schreij & Theeuwes, 2012) - Appendix D.3. Participants were asked whether they have or have not seen the displayed image in the gallery they just visited. Answers were given by pressing one of two keys on the laptop keyboard labeled as 'YES' and 'NO'. The instruction emphasised accuracy and speed of responses. First three pictures were shown for the purpose of procedural training only, and responses to them were not recorded. Following the training phase, a set consisting of all pictures seen inside the gallery and of an equal number of new stimuli (previously unseen by participants) was presented sequentially, one picture at a time. The order of presentation was randomised and each stimuli was proceeded with a fixation cross lasting for 250 ms, located in the centre of a blank screen. This was included in order to cue the location of the appearing stimuli and ensure equal eye gaze position was assumed prior to the presentation of each picture. Figure 9.1 presents a sample participant performing the task

Such conducted Recognition Memory test is a classic procedure used in experimental psychology for over a century (Wundt, 1874). Two measures can be calculated from it: *Reaction Time* (RT) and yes-no *Accuracy* of response (in this thesis referred to as *RT Accuracy*). Shorter *Reaction Times* are typically indicative of better recognition memory. Accuracy tends to produce the ceiling effect[1] in simpler tasks, and can be less useful.

Reaction Times have been linked to the level of processing (Craik & Tulving, 1975) due to the fact that the deeper we process something, the better it will be remembered. In consequence, better memory will most likely prompt a shorter correct answer in the Recognition Memory Test. As it is unclear if depth of cognitive processing can be

1 In psychology, the 'ceiling effect' describes a situation when the administered task is so easy in the given context that most of participants reach scores close to the maximum, preventing the researcher from drawing any comparisons from the data. The opposite 'floor effect' occurs when the task is too difficult and most of the scores are close to the minimum possible value.

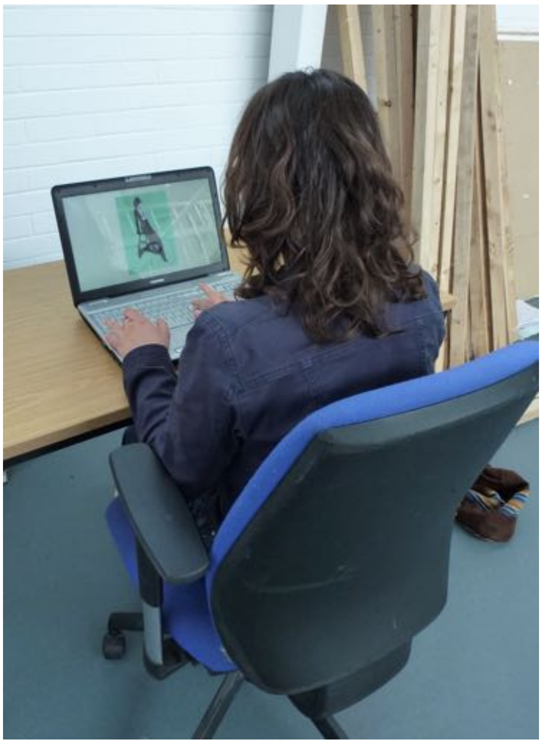


Figure 9.1.: Sample participant performing the Recognition Memory Test.

reliably measured in situ at all (see however: Tröndle, Greenwood, Kirchberg & Tschacher, 2014), such a post-visit assessment constitutes a well-theoretically-grounded operationalisation.

Moreover, implicit memory processes were linked to the aesthetic experience of art appreciation (Leder et al., 2004) suggesting a linkage between memorisation and liking. Yet, the scope of this thesis and the type of data collected does not allow us to make assumptions about the preference for separate artworks.

Table 9.1 below provides a brief summary of the measures used to analyse Recognition Memory Test results.

Measure	Definition	Reason for Including	Previous Usage in the Field
Reaction Times (RT)	Number of milliseconds it took a participant to correctly recognise a picture (i.e. press 'YES' on a keyboard when a non-distracting stimuli appeared on the screen). The higher the score, the 'worse' the response.	Indicative of the depth of processing and pictorial memory of the stimuli.	Established measure.
RT (inverted)	(1/RT) transformation of the Reaction Time measure. The higher the score, the 'better' the response.	This inverted measure is a recommended transformation of the original data for avoiding problems with non-normal distribution of RT variable.	Recommended by Baayen and Milin (2010)
RT Accuracy	Correct or incorrect answer to the stimuli appearing on the screen. 'Personal Mean RT Accuracy' would indicate the proportion of correct answers given by a single participant.	Correlated with Reaction Times, although often producing the 'ceiling effect'.	Established measure.

Table 9.1.: Recognition Memory Measures.

It must be acknowledged however, that recognition memory does not comprehensively account for all aspects indicative of the retrospective experience of an art gallery visit. If this was true, spatially-defined relations between artworks would bear no difference to the memory of the visitor. As we have shown in Section 3.2.2, they most likely do and therefore assessing the memory for these spatial relations is equally important. In order to account for both picture-based memory and location-based (spatial) memory, as previously done by Janzen (2006), Miller and Carlson (2011), another test was required.

10

SPATIAL MEMORY TEST (MINIATURE TASK)

Measuring Spatial Memory as one of the dependent variables derives from the assumption that the museum experience is more than just a set of individual interactions between the viewer and a number of separate artworks (Wineman & Peponis, 2010). As it has been shown in Section 3.2.2, individual viewing experiences are combined in human memory and form a hierarchical knowledge construct. This can be represented in the form of a cognitive map[1] of the environment.

A number of distinct methods of measurement were used in spatial memory studies. For example, Pointing Task assesses participants' memory for object's spatial location by asking them to point to the location of the target from the position where it is not visible from. As the outcome variable is the angular error between the pointing direction and the correct target location, this task is more viable for large scale environments, where potential targets are likely to be spread across wide range of angles relative to the participant's position. In the procedure followed by this thesis, all potential targets (i.e. pictures) were located in roughly the same direction when considered from the room adjacent to the gallery. Thus, angular error would most likely have very low variance and there was a high probability of providing an answer close to correct purely by chance.

An alternative method, often used to tackle this limitation, is to ask the participant to imagine standing at a specific location inside the measured environment and to perform the Pointing Task from this imaginary perspective (e.g. 'imagine you stand with your back to the entrance, object X is on your right, point to object Y'). This procedure, however, involves two possibilities of making a mistake: if participant's memory for the layout is poor, he or she is likely to wrongly set him-/herself in the imagined environment in the first place. The total angular error is therefore a sum of two possible mistakes, which makes this procedure not viable for a spatial setting of non-trivial geometry, lacking any distinctive landmarks serving as a universal point of reference.

A more flexible method of assessing spatial memory is a Sketch Map Drawing Task. In it, participants are simply asked to draw the

1 A cognitive map introduced by (Tolman, 1948) is a mental representation of survey, or 'bird's eye' knowledge of the environment acquired spontaneously from navigation (Nadel, 2013).

layout of the environment they have visited. Due to varied drawing skills and multiple ways of representing space, the variability of obtained answers is typically large. This makes making comparisons across larger number of participants problematic. For example, poorly remembered landmarks might be omitted in some maps and other objects, not initially meant to serve as such, might be included.

The main aim of measuring spatial memory in this thesis was to draw conclusions on the participants' understanding of the relationships between individual pictures. For this reason, the task employed must have encouraged participants to report assumed location of all of them and not only the most prominent ones. Concurrently, other objects potentially well remembered from the visit were not of interest to this thesis.

Those requirements could only be fulfilled with a purpose-built method. 'Miniature Task' was designed, inspired by various modifications of sketch map-based exercises previously reported in literature (von Stülpnagel & Steffens, 2012; Münzer, Zimmer, Schwalm, Baus & Aslan, 2006; Waller, Loomis & Steck, 2003). In the Miniature Task, participants were presented with miniature versions of all pictures previously seen inside the gallery and the floor plan of the space. No distractors (such as pictures not previously seen in the gallery) were presented, contrary to Recognition Memory Test. Visitors were asked to align picture miniatures with the walls on the floor plan according to where they think they saw the picture in the gallery. Duration of the task was not limited. During Experiment 2 and BALTIC Case Study the task was presented on a computer screen and participants used 'drag-and-drop' method to manipulate miniatures with a mouse. The Miniature Task can therefore be interpreted as a forced-choice adaptation of the sketch map concept (Fig. 10.1).

The reason participants were presented with this test after the Recognition Test is that it did not contain any distracting stimuli—all pictures presented as miniatures were previously seen in the gallery. Thus, administering the tests in the opposite order would most likely update participants' recognition memory and bias their reaction times in the Recognition Memory Test as testing increases the likelihood of future recall (McDermott & Arnold, 2013). Concurrently, since Miniature Task only measured spatial memory performance, it is unlikely that the Recognition Memory Test had enhanced the scores of the Miniature Task.

10.1 analysis of the miniature task results

Analysing Miniature Task, similarly to Sketch Map Drawings involves an interpretational challenge at the stage of transforming visual output produced by the participants into quantitative descriptors required for further statistical analyses. Any type of such numeric variable, in

Figure 10.1.: Sample participant performing the Miniature Task.

order to allow comparisons between individual cases, must reduce the variance of the exact pictorial output to a single descriptor of some limited aspect of the answer. This thesis considers three distinct methods of analysis in order to account for most relevant aspect of the obtained output. Depending on the exact spatial differences across the analysed conditions, distinct methods of analysis were used.

Bidimensional regression is a method noticeably gaining popularity for sketch map analysis in spatial cognition studies. This statistical concept was initially introduced to spatial sciences by a geographer Waldo Tobler (Tobler, 1965) and later it has been adapted to sketch map analysis by (Friedman & Kohler, 2003). Most recently, Carbon (2013) developed an R function, and Gardony, Brunyé, Taylor and Wolford (2013) a GUI software with the aim of popularising the method in spatial cognition research. From a statistical viewpoint, *bidimensional regression* is very similar to *regression analysis*, but it is performed on two-dimensional coordinates of each point. A single point can represent the memorised location of a landmark, or—in the context of this thesis—an artwork. The positions of all points on participant's sketch map are statistically correlated with their true positions on a reference map.

One clear advantage of *bidimensional regression* is the fact that it is robust to consistent distortions (such as underestimated distances) and represents whether the relations between the objects are kept. However, it remains insensitive to the presence of walls segregating the layout into a nonlinear arrangement of subspaces. In consequence, the measure does not differentiate between 'same-wall' and 'across-wall' types of errors. For example, the participant might make a relatively small mistake by placing a picture miniature 1 cm to the side of its original location on the true wall surface ('same-wall' error). Another participant might place the object on the wrong side of that wall, which is also often equal to 1 cm shift in its coordinates ('across-wall' error). Bidimensional regression is insensitive to this difference. In result, there exists a risk of inflating the final scores of those participants, who performed poorly by making across-wall types of errors more often. Another disadvantage of the measure is the fact that it only offers a generic score reflecting the participant's performance in the whole task. It therefore does not make it possible to make any performance comparisons between individual locations (i.e. if the content of some of them was remembered better from others).

In case *bidimensional regression* would indeed prove to inflate spatial memory results of poorly performing participants, an alternative method of scoring Miniature Task solutions was prepared. *Back-to-the-Wall* statistics was developed specifically for the purpose of this thesis. It calculates the result of Miniature Task based on the question whether the participant placed each picture back on the correct wall. For each picture, 1 point is given if the solution of the Miniature Task contains its representation back on the wall where it has really been seen inside the gallery, and 0 points if it is placed on a different wall. The sum of points divided by the total number of pictures creates Mean Personal Back-to-the-Wall score ranging from 0 to 1. Personal Mean equal to 0 indicates that a participant has placed all pictures on wrong walls while the value of 1 would reflect a situation where all pictures were placed anywhere along the true walls they have been seen inside. Similarly, mean values for each location can be calculated by pooling their average from all participant's results. This will make it possible to compare locations which were, on average, remembered better than others.

Despite these advantages, *Back-to-the-Wall* statistic bears a number of significant limitations. Firstly, it ignores many details about the actual relations between objects remembered by the participants. For example, it is possible to obtain the maximum possible mean score of 1 even in the presence of multiple mistakes, such as misplacing two pictures adjacent to each other on the same wall. It therefore does not support reasoning about the detailed nature of spatial memory, but only rough estimates of its general performance. This might particularly affect the scores of better performing participants by reducing the variability of results between them.

Secondly, despite the fact it is improbable to perform well in the Miniature Task by chance (due to the number of available wall sections), it is possible that a poorly performing participant makes a consistently biased, non-random decision about the placement of a specific picture. For example, choosing walls containing higher number of pictures increases the chance of success. This bias however seems to be inevitable. Even in layout-independent measures, such as pointing task, participant can always acquire a strategy of pointing to the angle range containing higher density of targets. However, such bias actually promotes having *some* spatial memories of the environment over having none, possibly distinguishing those poorly performing participants who at least remembered most densely occupied walls, from those who did not.

Lastly, *Back-to-the-Wall* score might be biased by the wall length. It is possible that a participant having no spatial memories of the environment bases his or her answers to the Miniature Task purely on the perceived length of the walls on the layout provided. If this was true, we can expect that longer walls will be linked to higher 'false positive' error rates than shorter walls. Such bias would not jeopardise the validity of the measure as long as it would only affect the incorrect answers. This would only suggest that those participants who had poor spatial memories of the environment used the cues provided by the task (i.e. by the miniature layout) to guide their responses. Given there is no linear relation between correct locations of pictures and wall lengths, this strategy would not help them to score higher. However, a situation in which correct answers are correlated with the wall length could mean that this cue also affected the responses of well-performing participants. We will control for these factors in the relevant part of the analysis.

Third measure of spatial memory employed in this thesis is based on String Matching techniques originally developed in mathematics and commonly used in modern computer science. The method has been previously employed for analysing spatial data by Conroy Dalton (2001), and already described in Section 8.3. In the context on analysing Miniature Task results, we can represent the sequence of pictures hanging on a single wall as a sequence of letters, where each letter is associated with a single picture. A one-dimensional representation of the sequence of pictures, as they are hung on the wall, constitutes a 'string'. For example, we could represent a set of three pictures hung next to each other as a string 'A-B-C'. In the Miniature Task, participants are asked to place miniature versions of the pictures on the walls they think they saw them on inside the gallery. For the purpose of this example let us assume that a sample participant of the Miniature Task has placed three pictures on a sample wall: 'A-C-F'. String Matching algorithm can be used to determine how much the response provided by the participant differs from the reference string originally present in the gallery, allowing us to assess the level of participant's spatial memory for the wall's content. Different String Matching algorithms consist of different rules used to quantify the difference. For a detailed review of various algorithms, the reader might wish to turn to the works by Wagner (1975), Kruskal (1983), and Stephen (1992). Appendix B provides a summary of this literature and the necessary context for understanding the custom-built algorithm used specifically for the purpose of Miniature Task analysis. This custom algorithm was designed and written in Python based on Weighted Damerau-Levenshtein Distance (Chey, 2013), but allowing for unlimited number of *swap* actions. Additionally, custom weights can be associated with each action type. The algorithm would determine whether to perform multiple *swaps* or a *substitution* based on their cost, so that the result is always the lower of two possibilities. Hence, in the Miniature Task analysis, the lower (better) score would always be given to those participants who placed picture miniatures closer to their original locations.

The advantage of this measure lies in the fact that—unlike Back-to-the-Wall score—it considers participants' performance at higher granularity of exact sequence ordering. This provides more accurate differentiation between individual participants, especially at the higher level of performance. However, the measure becomes problematic in a more complex layout settings, where the relations between individual walls are non-trivial. It is not always clear what constitutes a single string, and how the pictorial two-dimensional representations should be coded to a one-dimensional string of characters. For instance, deciding that the inside wall surface of a rectangular room can be considered as a single string, bears significant theoretical assumptions about the organisation of human spatial memory which go beyond the scope of this thesis. For these reasons, the measure will only be used in very simple layout settings of Experiment 2.

Table 10.1 below summarises all measures used to score Miniature Task. Further details will follow in study-specific sections, as we will determine which of the variables are most suitable for the given purpose.

Measure	Definition	Reason for Including	Previous Usage in the Field
Back-to-the-Wall	The proportion of pictures which were placed back to their correct, original wall in the Miniature Task.	Generic assessment of spatial memory performance which is robust to complex spatial settings.	Previous usage unknown; measure introduced for the purpose of the thesis.
Bidimensional Regression	Correlation of two-dimensional coordinates between picture miniatures' placements and their reference locations.	Contrary to Back-to-the-Wall, it preserves detailed spatial relations between individual objects.	Established measure.
Adrien-Kuba Distance (String Matching)	Quantified difference between two sequences of pictures (participant's solution and the true reference sequence), based on their one-dimensional character representations.	Contrary to Back-to-the-Wall, it preserves detailed spatial relations between individual objects.	Previously used in a similar context by Conroy Dalton (2001). Custom-built modifications were introduced to the algorithm.

Table 10.1.: Miniature Task Measures.

11

SPACE SYNTAX ANALYSIS

All galleries used in the study were measured by the researcher with a laser measuring device and sketched as CAD drawings. All Space Syntax and Isovist analyses were conducted using Depthmap 10.14.00b (Turner, 2001) and its newest version depthmapX. Each variable can be represented as a visualisation of the relations between points in space, with a numeric value associated with each of them.

The measures used in the analyses included standard variables describing size and shape of an isovist (previously explained in Section 5.2) which were calculated from the centre point of each picture's location in the gallery layout. Additionally, a number of factors which were suggested to play a significant role in the context of an art gallery visit were operationalised under the following spatial measures:

- *Visibility Catchment Area (VCA)* - as described in Section 5.2, it has been suggested (Stavroulaki & Peponis, 2003) that a 60°visibility cone in front of each artwork might bear a significant impact on the visitor experience. Depthmap software offers the functionality of calculating partial isovists but the options available do not include a 60°angle and the method of defining its direction is imprecise. For the purpose of the analysis, for each painting, a 60°cone was superimposed on the CAD drawing perpendicularly to the wall, with its apex at the centre of the picture. Subsequently, a standard isovist was generated from the inside point of the tip of the cone. As a result, the boundaries of the visibility cone served as outer boundaries of the isovist and its area could be derived.

- *Potential Co-Visibility* - As discussed in Section 5.2, making assumptions of the exact co-visibility of multiple objects based solely on their spatial location is impossible. Since participants' trajectories through space and head orientation cannot be known until they are observed, co-visibility can be only estimated. This thesis defines *Potential Co-Visibility* of a picture as the number of other pictures remaining within its isovist area. Higher number on this scale reflects higher *probability* of being co-visible with any other picture at any given moment. In other works, the measure operationalises the potential impact of a certain spatial location due to its co-visible positioning. Minimum value was 0.

- *Targeted Co-Visibility* - Assuming the special role of VCAs in guiding the visitor's experience, it would be theoretically plausible to expect that co-visibility occurring from within VCAs might bear more impact on participant's memory than a situation in which pictures are co-visible at oblique angles. One potential explanation of such phenomenon would be that co-visibility disturbs participants while they remain within the 'comfortable area of viewing' a picture—its VCA. To account for this possibility, the two concepts described above were combined to create another spatial measure. It is based on the concept of *targeted visibility* proposed by Lu and Zimring (2012), and already described in Section 5.2. A custom made Depthmap script was kindly provided by Yi Lu (personal communication, 2013) and modified for the purpose of the analysis. The script supplements Visibility Graph Analysis by calculating how many predefined 'visual targets' are visible from each grid cell in space. Centre points of picture locations were defined as such 'visual targets'. Following these calculations, 'targeted visibility' values of all grid cells lying within the VCA of a given picture were averaged and the resulting mean was associated with the location of the reference artwork. As a result, the measure of *Targeted Co-Visibility* reflected mean number of potentially visible pictures for a person moving within its Visibility Catchment Area with the minimum possible value of 1.

Table 11.1 below presents a combined summary of all Space Syntax and Isovist measures used in the study.

Measure	Definition	Reason for Including	Previous Usage in the Field
Isovist Area	Total floor area visible from the point of reference (in the case of this thesis: the centre of a picture).	A classical measure describing visibility.	Established measure.
Isovist Jaggedness	Area-to-Perimeter ratio describing the 'spikiness' of an isovist.	Previously linked with cognitive response to the spatial environment.	Established measure.
Visibility Catchment Area (VCA)	Isovist area constrained to 60 degrees.	Describes theoretically pleasing concept of 'area of comfortable viewing'.	Suggested by others researchers. Unused in empirical experiments.
Potential Co-Visibility	Number of pictures co-located within the isovist area of the reference artwork.	Describes the potential level of visual distraction.	Used in at least one other study.
Targeted Co-Visibility	Mean number of pictures visible from within the reference artwork's VCA.	A more refined measure of potential spatial distraction.	Used previously in a different context (hospital analysis).

Table 11.1.: Space Syntax Measures.

12

HYPOTHESES

Being grounded in empirical and theoretical work at the intersection of three distinct disciplines, this thesis aims to make original and confirmatory contributions at distinct levels of generalisability. This section will gather and describe stated hypotheses. Some of them are detailed confirmations of controversial findings from the existing studies, and others suggest novel, broader contributions at the intersection of multiple disciplines. Key original contributions are indicated with an asterisk next to the hypothesis' number. These are complemented and reviewed as the work progresses in accordance to the developing empirical interpretations (Section 17.6 on p. 227 and Section 21.6 on p. 278). Further in the thesis, Section 25.1 on p. 319 will look back at these hypotheses and review them in the context of empirical work undertaken in the current series of studies. Section 25.1 will also further discuss the weight of contribution deriving from verifying (or falsifying) each individual hypothesis.

12.1 WHAT ARE THE GENERALISABLE PATTERNS OF HUMAN OCULOMOTOR BEHAVIOUR INSIDE ART GALLERIES AND HOW THEY IMPACT MEMORY (A)

- *(A1)*: The process of viewing a work of art seems to be divided between (a) getting a quick 'gist', and (b) diligent examination of individual pictures. Thus, it can be expected that the distribution of dwell lengths will be heavy-tailed: with a large number of *Short Dwells* similar in length, and a more diverse set of longer ones. Similar distribution has already been observed (but not quantified) in a preliminary eye-tracking museum study by Wessel et al. (2007).

- *(A2)*: A finding from the study conducted by J. K. Smith and Smith (2001) suggested that *Total Dwell Time* per picture averages to about 30 seconds. Similarly, it took about 30 seconds before participants in the laboratory-based study of Locher et al. (2007) processed the artworks to a satisfactory degree to assign a pleasingness rating to the viewed artwork. This can be directly verified within the current thesis in context of multiple different art gallery set-ups, containing different numbers of artworks.

- *(A3*)*: Concurrently, high variance of average viewing times per picture has been found (J. K. Smith & Smith, 2001), presumably resulting from the difference in their attractiveness or complexity. Previous studies, however, were not able to disconnect the impact of spatial location from the impact of artwork's content. It remains to be verified whether a significant difference between average *Total Dwell Times* of individual pictures occurs in this thesis' experimental set-up, where pictures were visually similar and rotated their locations. It is expected that the impact of by-picture variance, partially explained by *Salience Rating* will remain significant, but to a lesser extent than the impact of spatial factors.

- *(A4)*: Eghbal-Azar and Widlok (2013) suggested that mobile eye-tracking might be useful in discovering 'exhibition visit scripts' - a local, or global viewing sequence repeated by multiple participants. Similarly, (Wineman et al., 2006) stated that museum visitors tend to engage separate elements of an exhibition within thematic grouping and the more visible spatial grouping is, the stronger this tendency. This work hypothesises that the presence and variability of such repeatable viewing sequences vary, depending on the layout of the exhibition. *Mean Normalised Levenshtein Distance (MNLD)* described in Section 8.3 can be used to compare the similarity of viewing sequences between individual participants. Significant cross-conditional differences are expected.

- *(A5)*: In terms of the eye movement's relation to memory performance, B. W. Tatler, Gilchrist and Land (2005) found that memory for object's *position* accumulated with number of fixations, while memory for object's *presence* and *colour* did not. This suggests that *Number of Dwells* and *Total Dwell Time* should predict participant's *Spatial Memory* of individual objects, but not their *Recognition Memory*. One limitation of this statement is the fact, that in the above study, participants were only allowed to look at the scene for 5 seconds. Since viewing mode fluctuates with viewing duration (Krejtz, Duchowski & Çöltekin, 2014), this hypothesis might only be true for participants who spent less time looking at pictures.

12.2 HOW SPACE IMPACTS OCULOMOTOR BEHAVIOUR (B)

- *(B1)*: Above all, this thesis assumes that in a museum setting the 'default cognitive task' of viewing pictures overrides regular oculomotor behaviour typical for locomotion outside art galleries (Hollands et al., 2002). Equivalently to this statement, it has already been suggested that the presence and location of artworks affect locomotive and viewing behaviour inside museums more than the physical boundaries of the building (Stavroulaki & Peponis, 2003). Since it is difficult to imagine a situation where a real-life 3-dimensional environment is explored with no 'cognitive task' whatsoever, it can only be expected that a pattern of eye movement with no 'default' cognitive task of viewing pictures would be highly dependent on the visibility of individual wall segments[1]. If pictures really become the default visual targets inside a gallery-like environment, it can be expected that the proportion of fixations falling on pictures will be higher, than what it could be expected from the share of those pictures' surface in the entire wall surface area or in the visibility of individual wall segments.

- *(B2*)*: As we have discussed earlier, the 'top-down' influence of high-level cognitive processes on visual attention for pictures can also become present in situations when visitors sense higher prominence of a particular location (Wiener et al., 2007). In such a situation, the prominence of location might be associated with the importance of the artwork which occupies it, and result in longer and more frequent viewing. Thus, the primary hypothesis this thesis states in relation to the influence of space on human attention is that more prominent locations (operationalised as those having larger *Isovist Area*, and larger *VCA* attract longer *Total Dwell Lengths* and higher *Number of Dwells*. An alternative—'bottom-up'—explanation of such effect would be the higher chance a more visible painting has of falling into the viewing field of a visitor freely exploring the gallery under the assumption of a purely random selection of the visual target.

- *(B3)*: In his early work, Robinson (1928) suggested that the average time of investigating individual pictures does not decrease proportionally to the number of pictures in the gallery. Therefore, it can be hypothesised that the average *Total Dwell Time* of 30 seconds per picture will remain unchanged regardless of the total number of pictures present in the gallery.

- *(B4)*: Robinson (1928) also emphasised the impact of low *Co-Visibility* in lengthening viewing time, but only if the 'isolation is complete', i.e. when only a single picture is visible on its own. He found no such effect when any other, 'non-complete'

1 A person randomly exploring a 3-D environment and performing random, task-free eye gazes, would fixate more often on wall stretches which are more visible relative to other wall stretches in that space.

isolation was arranged. Robinson's findings can therefore suggest that pictures placed at locations where no other pictures are *co-visible* will be viewed for longer than others.

- *(B5)*: In a similar manner, Bitgood (2010) predicts longer, undisturbed individual dwells and better 'understanding of the exhibition' when the exhibition is organised sequentially, with less freedom left to the viewer. He does not mention that the isolation must be 'complete', thus suggesting a more linear, negative relationship between *Co-Visibility* and the proportion of long dwells.

- *(B6)*: In his model, Bitgood (2010) also indicates that leaving more freedom of choice to the viewer and organising the exhibition in a non-sequential way (i.e. providing higher *Co-Visibility*) will result in—as he termed it—'spotty-focused' attention, as the stimuli will draw attention from each other. Such situation can be operationalised as higher *Picture-Switching* value in the current method. The resulting hypothesis is that *Co-Visibility* will be positively correlated with *Picture-Switching*, as 'spotty-focused' attention is more likely to occur under such circumstances.

- *(B7*)*: Bitgood (2010) additionally suggests that the last stage of viewer-artwork interaction is mostly dependent on subjective factors involved in making meaning of the viewed object. Since meaningful interpretation develops over time (Leder et al., 2004), Bitgood's suggestions can be operationalised as longer period of deliberate viewing. The author downplays the role of external, environmental factors at this stage of the interaction, indicating that the presence of *long dwells* is not correlated with spatial characteristics (such as *Co-Visibility*, or *Visibility Catchment Areas*). Considering research in Architecture and Spatial Cognition, this thesis however proposes an alternative hypothesis: that environmental factors will play a role even at this stage of the process and that the presence of *long dwells* will not affect the dependence of the human attention and memory on the external spatial factors.. This would demonstrate that making meaning of the exhibition is not a process guided by purely subjective decisions, but that space can—to some extent—facilitate even longer, deeper engagement with art.

- *(B8)*: One of the environmental factors affecting number of *long dwells* might be the presence of multiple pictures hanging *on the same wall*. As Tröndle, Greenwood, Bitterli and van den Berg (2014) demonstrated, multiple pictures adjacent to each other may be perceived as a holistic creative entity and, as a result, viewed from a larger distance, and from a larger set of diverse locations. Such situation is contrary to viewing individual works from uniform 'hotspots of interest' typically located directly in front of the paintings. Spatial configurations containing higher number of adjacent pictures at the same walls are—according to this hypothesis—likely to result in smaller number of *long dwells*.

12.3 HOW SPACE IMPACTS MEMORY (C)

- *(C1)*: The negative effect of divided attention on memory predicts that higher *Co-Visibility* of multiple pictures will decrease *Recognition Memory* (i.e. increase *Reaction Times*).
- *(C2)*: Landmark literature suggests, that in real-life 3-dimensional spaces object-oriented *Recognition Memory* is correlated with the objects' perceived salience while its *Spatial Memory* is not. Instead, *Spatial Memory* of its location is linked to the importance of object's location for the given spatial task. Such interpretation assumes a positive correlation between picture-oriented *Recognition Memory* results and the artwork's *Perceived Salience*.
- *(C3)*: This framework also provides a hypothesis regarding the effect of *Potential Co-Visibility* on Spatial Memory for object's location - the higher *Potential Co-Visibility*, the higher might be the perceived spatial importance of an object, and thus its location might be recalled better.
- *(C4)*: In their preliminary study, Dalton et al. (2010) have shown that pictures presented on a public display indoors were remembered better if the location of the display was defined by a spiky, and not round isovist. According to this finding, it can be hypothesised that higher *Isovist Jaggedness* (describing a more spiky isovist) will be correlated with better *Recognition Memory*.
- *(C5)*: One additional finding by Dalton et al. (2010) revealed that pictures were remembered better from words when presented from large *Isovist Areas*, and worse from them if presented from small *Isovist Areas*. Despite the fact that no direct comparison can be made to the method employed in this thesis (as no words were subject to testing), this finding allows to form a speculative hypothesis that in the real-life 3-D environment explored spontaneously, larger *Isovist Areas* will result in better *Recognition Memory*.
- *(C6*)*: A similar effect can be predicted for larger *Visibility Catchment Areas*, as—contrary to the entire isovist area—they describe the 'area of comfortable viewing'. It can be hypothesised that

maximising such an area reduces the cognitive load associated with processing fluency of the picture's content relative to other artwork locations and thus enhances visitors' *Recognition Memory*.

- *(C7)*: Based on the classic memory experiments and physiological studies (Tröndle, Greenwood, Bitterli & van den Berg, 2014) a 'serial position effect' can also be expected. It is predicted that the picture seen as the first and last will be remembered better, however the recency effect might be much smaller (even statistically non significant) in the presence of a temporal gap proceeding the recall phase. Importantly, it is predicted that strong primacy effect will prevail under distinct spatial configurations, despite different visibility properties of the first stimulus in the series.

12.4 OTHER FACTORS (D)

Additionally to the hypotheses directly linked to the main research question stated by the current thesis, a review of the existing literature indicates the importance of considering additional factors strongly influencing human memory and attention in our everyday surrounding. Even though it is not the aim of this work to investigate the nature of these factors, considering them during data analysis is important, as they might influence the effects this work is interested in.

- *(D1)*: For instance, classical findings in visual memory noted that older participants remember less pictures. This factor will be considered, as the empirical studies described in this thesis recruited a very diverse group of participants.
- *(D2)*: Also the importance of 'the museum context' has been exemplified in a number of recent museum-based studies (Brieber et al., 2014; Tröndle, Greenwood, Bitterli & van den Berg, 2014). Viewing times inside real art galleries have been shown to be longer compared to laboratory-based settings. Therefore, it is expected that *Total Dwell Lengths* will be higher in the BALTIC case study, compared to two other experiments conducted within the current thesis. It must be noted however, that this hypothesis cannot be reliably verified in a situation where each of the studies was conducted at a different floor area, and with a different number of artworks. It would, however, offer an alternative explanation to the assumption that the attention devoted to individual pictures shall remain stable despite their larger number (Robinson, 1928), or decrease under such condition (Serrell, 1997; Yoshimura et al., 2012; Bitgood, 2010). Consequently, in case of significantly higher *Total Dwell Lengths* in the larger and more densely arranged BALTIC gallery, such an effect can potentially be linked to its stronger 'museum context'.

- *(D3)*: Also, Wineman et al. (2006) found that the impact of spatial layout was higher for participants who spent more time inside the considered exhibitions. Since time is expected to be one of the main linear predictors of *Total Dwell Times* and *Number of Dwells* performed inside the gallery[2], this statement, even if verified, might have limited applicability.
- *(D4)*: Considerable number of studies showed a direct linkage between viewing time and memory (Loftus, 1972; Hollingworth & Henderson, 2002), although the exact reasons for this effect, especially in consideration to task-specific factors are a subject of an ongoing debate (Hollingworth, 2012; B. W. Tatler & Tatler, 2013).

2 It is unlikely that at some point of the visit participants simply stop attending to pictures, thus dwells will accumulate over time.

Part IV

EXPERIMENT 1

13

METHOD

In this exploratory experiment[1], participants were invited to visit a non-public art gallery arranged in the 'Squires Annexe' building otherwise used as a studio and exhibition space by Fine Arts students of Northumbria University in Newcastle, UK. Figures 13.1 and 13.2 demonstrate the overview of the environment. The study was conducted between the 16th and 31st of August 2012.

13.1 PROCEDURE AND PARTICIPANTS

The procedure was based on the one described in Section 6. Thirty-two participants were invited for the study through the university email system and local job-seeker internet discussion forums. One participant declared previous familiarity with the gallery space and was removed from all subsequent analyses due to the influence of previous knowledge on spatial memory performance (O'Neill, 1992; von Stülpnagel & Steffens, 2012). The remaining 31 visitors (13 female) were aged between 18 and 63 years (M = 31, SD = 11.84). No artists or architects were allowed to participate. One participant arrived without reading the briefing information provided in the invitation email, and therefore was briefed with the remaining details prior to starting the experiment.

Participants were randomly assigned to one of two experimental groups depending on the scheduled day of the meeting. Fourteen participants explored the gallery in experimental Condition 1, and 17 in Condition 2. Time limit of 30 minutes was imposed due to the battery capacity of the eye-tracking device. The instruction emphasised that the participant is not required to spend all this time inside the gallery and can leave on his/her own discretion. Following the calibration of the eye-tracker, each participant was brought in front of the gallery entrance and asked to use the other door at the end of the gallery to exit. Two participants have ignored this instruction and left the gallery through the same door as they entered. Their data was included in the analyses, as all statistical procedures considered the real oculomotor dwells, i.e. the sequence of looking—and

1 Parts of this section were previously published in (Krukar, 2014b)

Figure 13.1.: A view of the gallery.

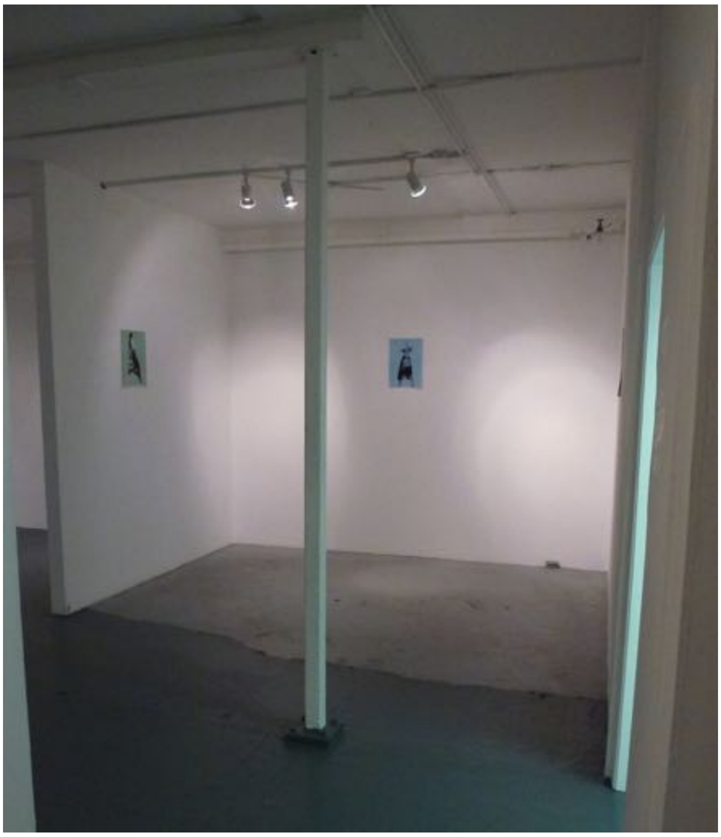


Figure 13.2.: A view of the gallery (same space, different angle).

Figure 13.3.: Pictures used in the study together with single-letter identifiers used throughout the data analysis.

not the spatial sequence—in order to investigate the relation between oculomotor behaviour and memory.

13.2 SPACE AND MATERIALS

Both conditions contained the same set of 14 artworks. Images used in the study were digital collages of equal dimensions (portrait-oriented A3), created by a local artist and an MA student at Northumbria University's Fine Art Department, Susi Bellamy (Bellamy, 2012, Figure 13.3). No labels or textual information about the artworks were provided.

The arrangement of the gallery walls was identical in each condition, but the placement of the pictures' locations differed (Figure 13.4). Note that (as it will be shown quantitatively in Section 14.1) this is sufficient to create the diversity in visual measures, which are the main scope of this work. Therefore, the effect of the modification of spatial and visual relations can be generalised to spatial layouts per se.

It is important to note that what people spontaneously memorise might result from the interaction of the location's spatial prominence and the stimulus's perceived salience. For that reason, pictures' perceived salience has been established in an independent study, and controlled (Section 13.6). Moreover, in order to further separate the effect of space from the effect of artworks, the sequence of the pictures seen inside the gallery was randomised for each participant. Therefore, each visitor had seen the same pool of pictures (Figure 13.3) in one of two sets of locations (Figure 13.4), but in a unique, random combination. This makes it possible to separately investigate the the influence of picture and the influence of space. It was hypothesised that significant impact of space will prevail, despite the influence of individual pictures' perceived salience.

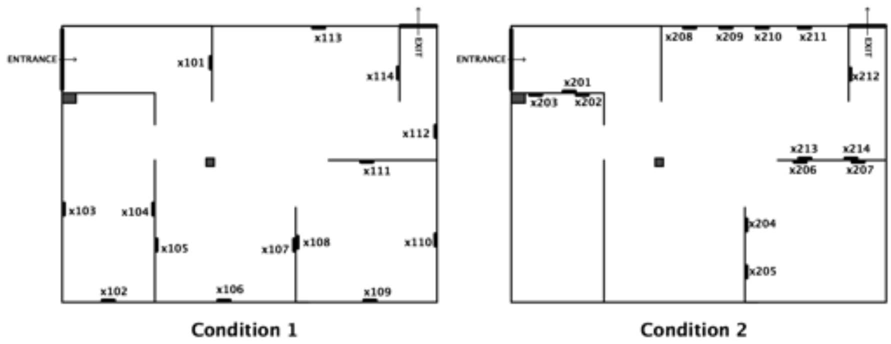


Figure 13.4.: Two spatial layouts arranged for Experiment 1 together with unique location-identifiers used throughout the data analysis. Note the identical arrangement of the walls but different arrangement of an equal number of pictures. The dimensions of the gallery's external walls were 9.8x7m.

13.3 EYE-TRACKING RECORDINGS

'Tobii Glasses 1' eye-tracker was used in the experiment. The calibration was successful for 30 participants, whose recordings were extracted and coded according to the procedure described in Appendix A.1. The coding procedure considered instances of each picture dwell, as well as (jointly) dwells on walls, floor and the ceiling.

13.4 RECOGNITION MEMORY TEST (REACTION TIMES)

Recognition Memory Test was presented via OpenSesame software (Mathôt et al., 2012) on a 17" laptop with two keys ('A' and 'L') labeled as 'YES' and 'NO' with black-and-white printed stickers. Subjects were instructed to answer whether they saw the displayed picture in the gallery or not. Accuracy and speed were emphasised in the instruction. Images were shown one at a time in a random sequence, each preceded by a fixation cross on a blank screen lasting for 250 milliseconds. Three additional, unrelated objects were shown at the beginning of the test for the purpose of procedural training. All 14 pictures presented in the gallery were included, with another 14 being new - either completely new or modified versions of the pictures seen in the gallery. Appendix E contains all distracting stimuli used in the Recognition Memory Test. Despite the fact that the eye-tracker's calibration was unsuccessful in two cases, the affected participants were unaware of the equipment failure and therefore still performed the task in the same context as all other participants. For this reason, the data obtained from their memory tests was used in all analyses which did not require matching with eye-tracking data.

Figure 13.5.: Sample solution of the Miniature Task.

13.5 spatial memory test (miniature task)

The Miniature Task was presented on a separate table as a printed A3 layout of the gallery. Picture miniatures were cut out separately and participants were free to move them around the layout. No time limit was imposed.

All participants produced valid solution. If ambiguities in miniature placement were present, the researcher asked to specify which exact wall location the participant considers correct and to adjust the position of the miniature. The solution was photographed together with a unique participant ID number in order to link the result to one's performance from other stages of the experimental procedure. Memory card failure caused data lost of one solution. Figure 13.5 presents a sample solution.

13.6 perceived picture salience study

In order to further control perceived salience of individual pictures contributing to the final gallery experience, a separate online experiment was conducted on an independent group of 54 participants recruited through social network portals. Qualtrics online platform (Qualtrics, 2013) was selected to host the study. The procedure was designed to imitate the one described by Miller and Carlson (2011). Participants were presented with a pool of 14 pictures used in the gallery study and asked to 'drag and drop' them on their screen according to 'how much they draw your attention'. Because this method would not be feasible for a large number of visual stimuli presented simultaneously on a small computer screen, pictures were displayed in two sets of three and two sets of four. The order of the pictures displayed on the screen was fully randomised, and the content (i.e., the neighbourhood of other pictures in which each picture appeared) was quasi-randomised in four experimental blocks. Out of 54 participants who took part in the independent salience study, 14 were removed due to not finishing the survey, outlying engagement time, or no 'drag and drop' action taken on at least one of four picture sets. Results of the remaining participants (N = 40; 22 female; *mean age* = 28.74, SD = 8.06; *mean time spent on the survey* = 189 sec., SD = 64 sec.) were calculated in the following way: for being dragged to the first position on the screen, a picture was given the score of 1. For being placed 2nd, 3rd and 4th the scores were 0.66, 0.33 and 0 respectively for four-picture sets. For the 2nd and 3rd position in a three-picture set 0.5 and 0 points were given. Mean score of each picture was pulled from the positions given to it by each participant and this constituted the *Salience Rating* falling between the range of 0 and 1. *Salience Rating* results were linked to individual pictures. In this case a score of 1 would indicate that every participant dragged the picture to the first position along its neighbours. Mean value equal to 0 indicated that all participants placed the object at the bottom of the set in which they saw it.

14

DATA ANALYSIS

All statistical analyses were conducted using R (R Core Team, 2013). Figures were generated using `ggplot2` package (Wickham, 2009).

14.1 CROSS-CONDITION DIFFERENCES: SPACE SYNTAX AND ISOVIST ANALYSIS

In order to quantify spatial differences between Condition 1 and Condition 2, mean values of spatial properties of separate picture locations were compared (Figure 14.1).

Firstly, mean *VCA* for locations in Condition 1 was larger (M = 117464, SD = 63268) than in Condition 2 (M = 99261, SD = 69807). Secondly, mean number of other objects present within single location's isovist was larger in Condition 2 (M = 3.86, SD = 2.35) compared to Condition 1 (M = 2.86, SD = 1.56), indicating that *potential co-visibility* of multiple pictures at any given moment was higher in Cond. 2. On the contrary, the size and 'spikiness' of isovist areas for locations in both conditions were similar. This confirms that both conditions allowed visitors to look at pictures from a similar proportion of the total gallery area and that the stability of these visual experiences was comparable. The difference lied rather in their quality: Condition 1, with higher *VCAs* and lower *potential co-visibility* was expected to facilitate more comfortable and less distracted viewing.

It also important to note, that the coverage area of picture isovists in both conditions was equal to 100% of the floor area. This means that no matter where the visitor would stand in the gallery, he or she would always be able to visually engage with at least one picture. There were no 'blank spots' of unused floor surface. This fact, together with similar mean size of individual isovist areas further demonstrates that both spatial arrangements were similar in terms of the 'quantity' of the afforded visual engagement.

14.2 WITHIN-CONDITION DIFFERENCES: SPACE SYNTAX AND ISOVIST ANALYSIS

Measures described in Section 11 were calculated for each picture location using Depthmap (Turner, 2001). Figure 14.2 below presents

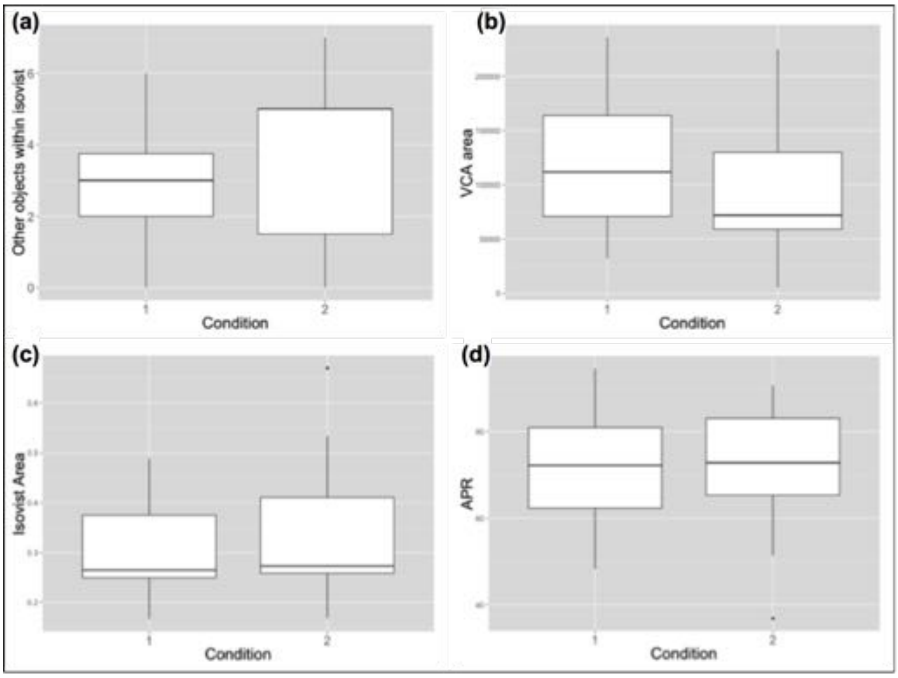


Figure 14.1.: Cross-Condition comparison of main spatio-visual properties.

sample visualisations. Each layout location was linked with a numeric measure describing its relation to all other grid cells of the layout. Table 14.1 presents an example of this numeric representation.

14.3 EYE-TRACKING MEASURES

Eye-tracking measures described in Section 8.2 were obtained from the coding logs, using a custom-built R script. Eye tracking data of two participants could not be extracted due to technical reasons (invalid calibration). Additionally, three participants have not complied to the instruction and did not look at all pictures, but only fixated at

Loc.	Isovist Area	VCA	Iso Jagged.	Targ. Co-Vis.	Pot. Co-Vis.
x101	169526.55	52121	72.78	2.40	1
x102	134220.58	110788	32.94	4.47	2
x103	182688.61	48408	47.81	4.85	3
x104	126229.85	32137	33.12	3.98	2
x105	334998.31	112856	57.14	4.99	6
x106	315590.81	235743	47.69	6.77	3

Table 14.1.: An extract from the numerical dataset describing spatio-visual properties of separate locations. Three isovist-derived measures are specified in arbitrary Depthmap units.

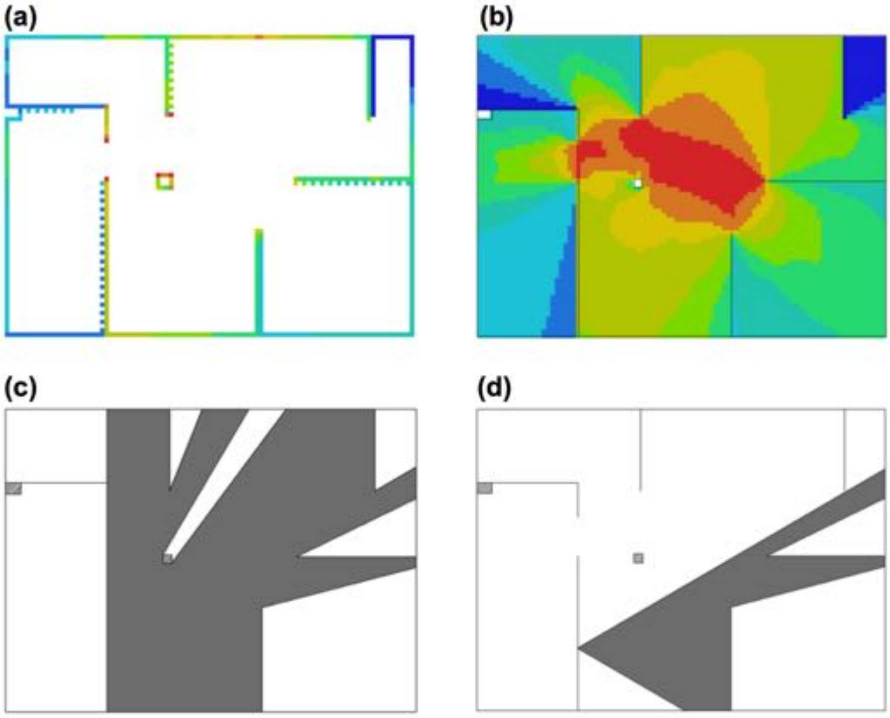


Figure 14.2.: (a) Boundary Graph Analysis; (b) full Visibility Graph Analysis; (c) sample Isovist, and (d) sample Visibility Catchment Area for the layout employed in Experiment 1.

13 of them. However, their results were kept in the database as informal observations of the eye-tracking recordings revealed that the missing pictures did appear in the periphery of their vision[1].

The recordings were also used to calculate total *time spent inside* the gallery by each participant (from the moment of the first dwell occurring inside the gallery to the moment of opening the exit door). In two cases where calibration was not successful, the videos still provided an estimated value of *time spent inside*.

In order to establish what affected the oculomotor behaviour of the gallery visitors, a statistical technique of linear mixed-effects modelling is used (Winter, 2013; Barr, Levy, Scheepers & Tily, 2013; Baayen, Davidson & Bates, 2008). The relevance of this method comes from the fact that the dataset contains multiple observations obtained from single participants, i.e. the experimental design is a mixed within-/between-subject design. In such circumstances, linear mixed-effect models have been shown to bare higher statistical power compared to by-item, or by-subject aggregated multiple regression, and to be robust in the presence of missing values (Baayen et al., 2008).

For the analysis, all visual engagements with individual stimuli were represented as 14 interactions associated with every single participant. Table 14.2 presents a sample from the dataset used for the model.

ids	pic	age	loc	Salience Rating	Iso. Area	VCA	No. of Dwells	RT
1	l	27	x101	0.19	169526.55	52121	6.00	2601
1	b	27	x102	0.40	134220.58	110788	6.00	886
1	j	27	x103	0.33	182688.61	48408	6.00	1878
1	i	27	x104	0.37	126229.85	32137	2.00	2586
1	e	27	x105	0.56	334998.31	112856	7.00	875
1	n	27	x106	0.46	315590.81	235743	6.00	2017

Table 14.2.: An extract from the dataset used for modelling. Each row describes the set of interactions between a single participant (id) and a single location (loc) while taking into account the picture (pic) which was there located for a given person. The interactions are also described with spatio-visual properties calculated by Space Syntax measures, and the responses provided via eye-tracker and in the memory tests. The example only contains a sample of all measures as their exact selection differed on a case-by-case basis.

1 The lack of a valid fixation in the dataset therefore did not result from misunderstood instruction but from technical reasons defining what a valid fixation is. The participants' understanding could be that they looked at all pictures even if no fixation was formally recorded within the employed data analysis procedure.

Individual participants have distinct viewing patterns and vary in the number and dynamics of the visual dwells they typically employ during an art gallery visit (or any other everyday experience). In result, their oculomotor behaviour in relation to 14 stimuli is linked within-subjects. For instance, it can be expected that some viewers will have higher average dwell numbers, and dwell lengths across the entire gallery visit purely because they tend to explore art in a more diligent manner in general. This violates the 'assumption of independence' associated with the typical Analysis of Co-Variance (ANCOVA) which could alternatively be used for the statistical analysis in this situation. Linear mixed-effect models account for these individual dependencies as they allow the base-line intercepts of all investigated effects to vary within separate subjects. This means that a linear mixed-effect model can estimate the influence of spatial predictors on oculomotor behaviour after correcting for distinct patterns in individual viewing tendencies. Such a model assumes that the potential 'base-line' behaviour is different for each individual and therefore the potential increase (or decrease) in eye-tracking variables caused by spatial factors will 'start' from a different level for each visitor.

Another violation of the 'independence assumption' comes from the expected effect of individual pictures and individual picture locations. Some pictures, and—as this thesis argues—some locations, are likely to generate more dwells or longer cumulative viewing times, than others. Similarly to individual differences, this by-picture and by-location variance can be accounted for in a mixed-effect model. Additionally to differences between individual gallery visitors, this makes the model 'expect' similarities within the same picture and/or the same location. Those expected by-subject, by-picture, and by-location effects are known as 'random' effects. On the contrary, generic predictors known as 'fixed-effect' predictors can be used by the model to explain the observed variance after accounting for the expected similarities within random effects. Most importantly for the research questions asked in this thesis, linear mixed-effect models allow the researcher to establish what proportion of variance *due to differences between individual locations* remains unexplained by the *known* ('fixed') spatial factors. In sum, this statistical technique makes it possible to establish what impact do spatial factors have on the oculomotor behaviour of gallery visitors, and how much of this impact remains present, but unexplained by the employed isovist measures. Additionally, the technique allows the researcher to consider all data, derived from both experimental conditions, in a single analytical process. Lastly, by avoiding prior aggregation of the data (e.g. by deriving by-picture, by-location, or by-participant means as it would be required for ANOVA-type analyses), the technique accounts for full variability of the original dataset, making the predictions more

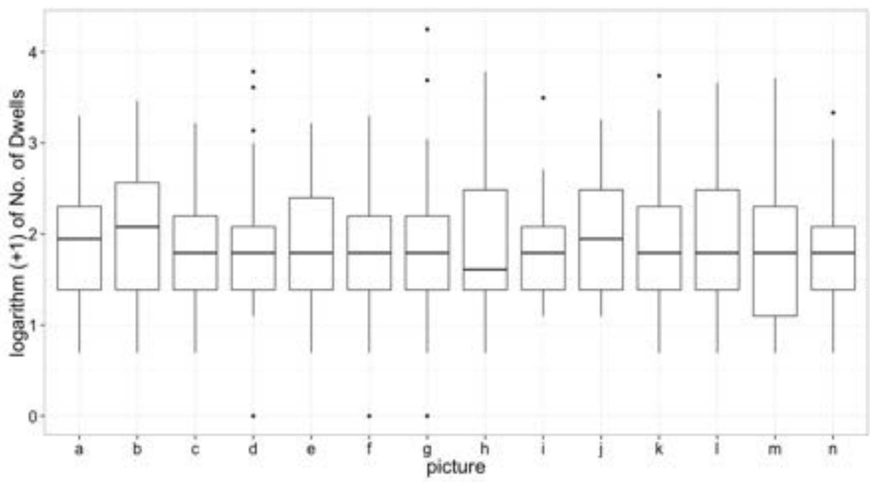


Figure 14.3.: Variability in mean (logarithm of) Number of Dwells classified by picture.

precise and increasing the chance of finding even smaller, but still significant, effects.

Four key Dependent Variables were selected for building 4 separate models. The variables were selected based on the fact that they describe distinct aspects of the 'quantity' and 'dynamics' of eye movement. Dependent variables describing its 'quantity' were *Number of Dwells*, and *Total Dwell Time* per picture/location. Dependent variables describing the 'dynamics' were *Picture-Switching*, and *Long-to-Short Dwell Ratio*. The predictors considered in the models were: *Experimental Condition*, *Isovist Area*, *VCA*, *Isovist Jaggedness*, *Potential Co-Visibility*, *Targeted Co-Visibility*, *Salience Rating*, *age*, *gender*, and cross-conditional interaction effects. As mentioned above, the model also included 'random effect' by-subject, by-picture, and by-location predictors. In other words, the model was 'aware' that the same *participants* looked at the same set of *pictures* hanging at the same set of *locations* and that the data points within each *participant*, within each *picture*, and within each *location* might be linked more, than observations across participants, across locations, and across pictures. Figures 14.3, 14.4, and 14.5 visualise this variability.

All numeric predictors were scaled and mean-centred (so that their mean equaled to 0 and standard deviation equaled 1) in order to make the intercepts easily interpretable. This transformation does not affect the model itself. Additionally, *Number of Dwells* and *Total Dwell Times* variables were log-transformed to approach normal distribution (Tufte, 1974). Because some entries in these variables contained 0, the transformation was conducted as log(x+1).

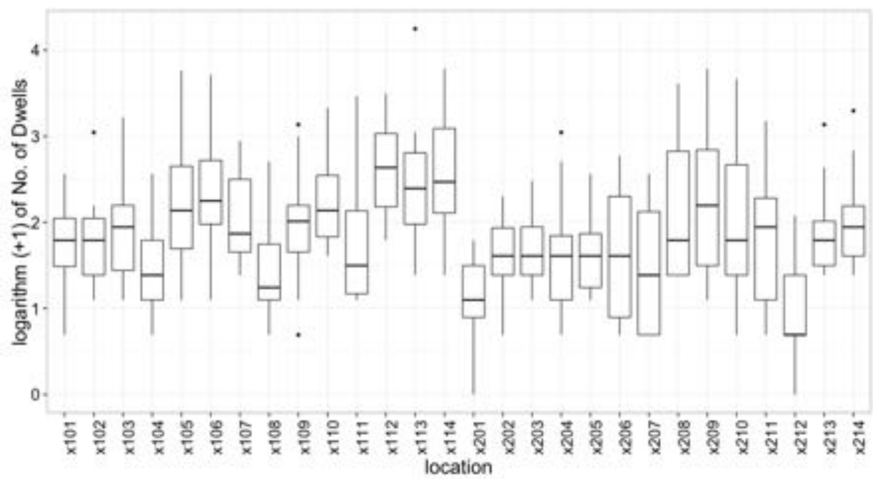


Figure 14.4.: Variability in mean (logarithm of) Number of Dwells classified by location.

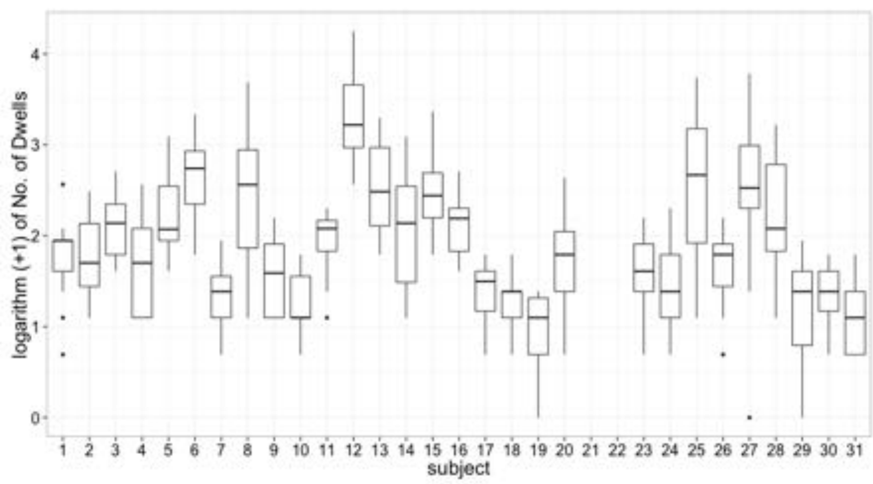


Figure 14.5.: Variability in mean (logarithm of) Number of Dwells classified by participant - each boxplot represents one person.

14.4 RECOGNITION MEMORY TEST (REACTION TIMES)

For Reaction Time data analysis, only the correct 'yes' answers were taken into consideration (78% of the whole dataset). *RT accuracy* was also recorded. Observations lying below the threshold of human physical ability (250 ms) or further than two standard deviations from the mean of the whole dataset were removed (Ratcliff, 1993; Whelan, 2008), leaving 71% of all responses valid (i.e. correct, above 250 ms, and within two standard deviations)[2]. Similarly to the eye-tracking dataset, linear mixed-effect models were used for statistical analysis. For this type of experimental design (mixed within-/between-participant design), this technique offers a considerable improvement in statistical power over traditional methods which require aggregating the reaction time data. In case of Reaction Time analysis, linear mixed-effect models allow the researcher to control for many confounding variables which are bounded to this specific type of task. For example, it has been shown that that the length of the proceeding RT is correlated with the next RT. Also, despite the fact that images were presented to participants in a random order, the order in the sequence of presentation might also affect the given Reaction Time (for a detailed discussion of the above effects as well as more arguments for the use of linear mixed-effect models for Reaction Time analysis please see Baayen & Milin, 2010). Since *RT accuracy* was also recorded, *Proceeding RT accuracy* will also be taken into account as a potential confound. Considering these factors helps to avoid the within-participant correlation between subsequent reaction times and to measure the main experimental effect of interest with reduced measurement error arising from the type of the employed task (Baayen & Milin, 2010). In result, r-squared values of the model increase (Baayen & Milin, 2010).

Memory performance of the experiment's participants can be modelled based on three sets of predictor variables: (a) spatio-visual predictors, (b) demographic predictors, and (c) the oculomotor behaviour (Fig. 6.1, p. 108). The first two sets are identical to the ones used to model the oculomotor behaviour. The process of including the oculomotor behaviour in the set of predictors would not be typically available to an actual gallery curator prior to physically building the exhibition. It does, however, provide the researcher with an opportunity to link human visual attention—as it naturally occurs in a three-dimensional, real-life context of an art gallery visit—to spontaneous memory, and thus build a wider understanding of the final

2 The reason why wrong answers and outlying reactions are excluded is that the cognitive processes responsible for very long or incorrect reactions might be potentially different from the ones being in the centre of the researcher's interest (Ratcliff, 1993). It is highly probable that those reaction times do not correlate with the depth of processing of the picture, but instead with some other, task-specific factors, such as a distraction during providing the answer.

gallery experience. Including all three sets of predictors in a single model also makes it possible to more reliably estimate effect sizes of the spatial factors under the presence of individual differences in eye movement. In other words, it puts the *influence of space* into context of individual viewing patterns which potentially might have a higher impact on memory than generic spatio-visual characteristics. One important issue with the data available, however, is the fact that the eye-tracking recordings are only available to a subset of all visitors[3].

For this reason, in order to utilise all data available, two types of models were built to explain human memory of the viewed artworks. One set of models included spatio-visual and demographic predictors, simply substituting the dependent variable compared to the models considered in the previous section. So instead of various eye-tracking variables being modelled as the outcome, the memory variables were the subject of the same analysis. This analysis was conducted for the entire dataset for which memory performance results are available. In the case of Reaction Time task, this is equivalent to all correct reactions within 2 standard deviations from the mean.

The second set of models consisted of spatio-visual and demographic predictors, as well as the oculomotor variables. The dataset analysed this time was restricted to those participants, for whom the complete eye-tracking recordings were available.

For models where eye-tracking variables were used as a predictor, additional data transformation was performed. *Number of Dwells* and *Total Dwell Time* variables highly correlated with each other (r = .59) and thus violated the collinearity assumption of linear models. To decorrelate these variables, their values for each set of visitor-artwork interactions were multiplied, creating a new measure: *ET-quantity*[4]. This new measure correlated well with original variables (r = .73 with *Number of Dwells* and r = .78 with *Total Dwell Time*) meaning that the decorrelation was justified and the resulting measure meaningful. However, at least one considerable theoretical issue prevails with the use of so transformed variable. A participant whose *Total Dwell Time* for a single artwork equalled x (e.g. 500 seconds), and who performed y dwells (e.g. 5 dwells) on that artwork would have a lower *ET-quantity* score than someone whose *Total Dwell Time* was the same, but *Number of Dwells* was higher. It means that a person who looked at a given picture for the same cumulative time, but did so in a more distracted manner, would score higher on *ET-quantity*. Since the variable correlates well with original variables, it can be used as a generic predictor for *Recognition Memory* in the subsequent models,

3 As the reader will see, the disproportion among those group sizes will be bigger in further studies.

4 Dividing the variables by each other would mean that a person who performed little short dwells had the same *ET-quantity* as someone who spent a lot of time on very numerous dwells. Multiplication inverts this effect: the more dwells OR the more cumulative fixation time, the higher the value.

but because—by being derived from two measures—it must make assumption of their theoretical relationship, it would not be reliable for models where specific relations between *Total Dwell Time* and *Number of Dwells* are of particular interest. In such cases, one of the two original measures will be used instead (in separation).

Lastly, For easier interpretability, all continuous predictors were scaled and mean-centred so that their $M = 0.0$ and $SD = 1.0$. The output variable *Reaction Time* has been inverted (1/RT), in order to increase its similarity to normal distribution and thus potentially increase statistical power as well as reduce the risk of non-normal residual distribution (Baayen & Milin, 2010).

14.5 SPATIAL MEMORY TEST (MINIATURE TASK)

To assess the results of Miniature Task, three measures previously presented in Section 10 were considered. String Matching technique was not suitable for this Experiment, due to possible ambiguities at the stage of transforming two-dimensional representations of angular wall relations into one-dimensional strings. The least controversial solution would be to calculate individual strings on a wall-by-wall basis, but in case of Condition 1 this would mean that all strings consist of a single character only. Two remaining alternatives were *Bidimensional Regression* and *Back-to-the-Wall* score.

Gardony Map Drawing Analyzer (Gardony et al., 2013) was used to conduct Bidimensional Regression analysis. The software allows its user to extract and correlate cartesian coordinates of correct landmark positions with their sketch map projections provided by the participant. For each Miniature Task solution, a picture of the results was taken, cropped to match the size of the reference image, and semi-transparently superimposed on the reference image of the gallery layout. Points representing centres of each picture were manually specified and the software conducted relevant statistical analyses. The resulting measure of *r* represents overall performance of a single participant in the task in accordance to the theoretical underpinnings of Bidimensional Regression.

In order to calculate *Back-to-the-Wall* score participants were given 1 point for each miniature placed on the correct wall. For each of 14 interactions, the measure thus provided a binary value indicating whether the answer was correct or not. Additionally, aggregated score of each participant ranging from 0 to 14 was divided by the total number of pictures (14) to create *Personal Mean Back-to-the-Wall* result. The minimum possible value was 0, which would indicate that the given participant misplaced all miniatures in relation to where they were located in the gallery during one's visit. Maximum possible value was 1, which would demonstrate that the person placed all picture miniatures back on the wall he or she saw them on.

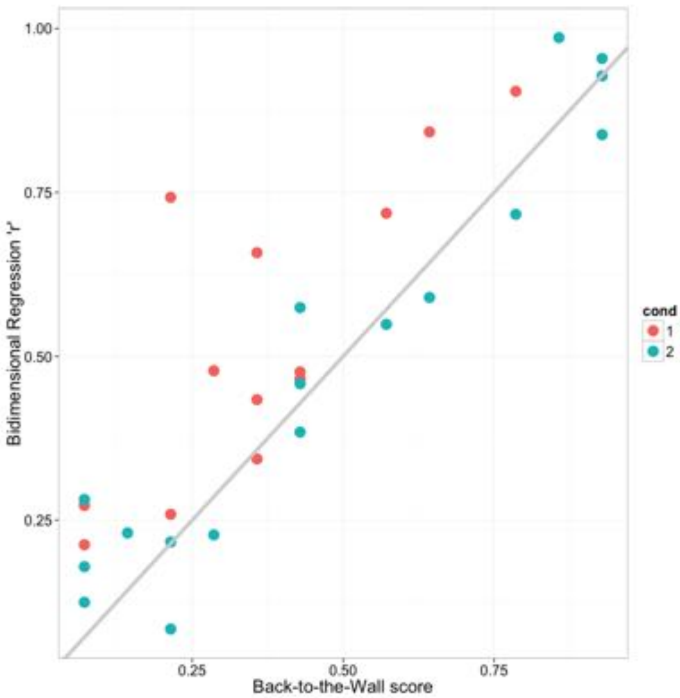


Figure 14.6.: Participants' spatial memory performance represented by Bidimensional Regression and Back-to-the-Wall results. Each data point represents the final results of a single participant calculated from their Miniature Task solution. The diagonal line represents a theoretical level of perfect correlation (r = 1.0).

As mentioned previously in Section 10, Bidimensional Regression applied to a layout where two sides of the same wall are equally possible solutions could inflate the results of poorer performing participants. In order to assess if this situation occurred in the current experiment, *Bidimensional Regression* and *Back-to-the-Wall* results were plotted together (Figure 14.6). If both measures reliably represented the same concept, the correlation between participants' scores on both variables should be equal (or close to) 1.

As visible on the figure, in a notable number of cases *Bidimensional Regression* scores lie above the theoretical line of perfect fit. This seems to be particularly visible in case of those participants, who achieved a relatively low score on *Back-to-the-Wall* measure. It thus reflects the prediction that *Bidimensional Regression* technique inflated the results of poorer performing participants. This seems to be the case especially in Condition 1, where pictures are located on a larger number of inner walls, which creates more opportunities for the type of error *Bidimensional Regression* does not control for (i.e. placement of a miniature on the wrong side of the correct wall). The technique was therefore not suitable for the analysis of this data set. For this reason, less detailed but more robust *Back-to-the-Wall* score was used to assess participants' spatial memory performance.

15

RESULTS

15.1 TIME SPENT INSIDE

Participants spent M = 537 seconds (or almost 9 minutes) inside the gallery (Cond. 1: M = 615 sec., SD = 536; Cond. 2: M = 474, SD = 413). The distribution of visit lengths showed a steep decrease at a very specific point, indicating that to the majority of participants, the time limit imposed was sufficient to explore the gallery. Concurrently, the observed bimodality of the distribution demonstrates that a significant subgroup of all participants stayed in the gallery for much longer than the average (Fig. 15.1).

15.2 VISUAL ATTENTION

Average participant fixated on the first picture within 1.92 sec. from entering the gallery and performed 108.2 dwells in total. Each picture was a subject of M = 7.73 dwells during a single gallery visit. Mean *Long-to-Short Dwell ratio* was 0.33, meaning that an average participant, for each *long dwell* (i.e. lasting more than 2 seconds), performed about two *short dwells* (< 2 sec.). As indicated by *Picture-Switching* metric equal to 0.49, almost half of all dwells were followed by an immediate fixation on another picture.

Mean *dwell length* was 4.1 sec. and the cumulative distribution of dwell lengths was heavy-tailed (as predicted by Hypothesis A1) in both conditions (Figure 15.2). The longest single dwell on a separate picture lasted for 157.54 seconds.

Looking at pictures constituted 76% of all eye movement activity inside the gallery (*Engagement Ratio* variable; 80.6% in Cond. 1 and 71.7% in Cond. 2) even though only 6.5% of all walls' length (64.44 m) was occupied by pictures (summed width = 4.2 m). Assuming that the majority of eye-movement in a gallery-like type of environment falls on the walls at an eye-sight level, we could expect that under the lack of cognitive goal of viewing pictures (i.e. assuming a completely random eye movement), visitors' eye gaze would fall at wall segments based on their visibility relative to all other wall segments in the gallery. Summing up the *isovist area size* values of all grid cells adjacent to any gallery wall made it possible to calculate what percentage of this

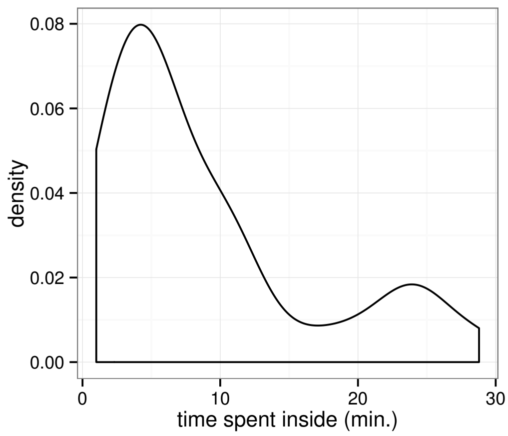


Figure 15.1.: Time spent inside by the experiment's participants.

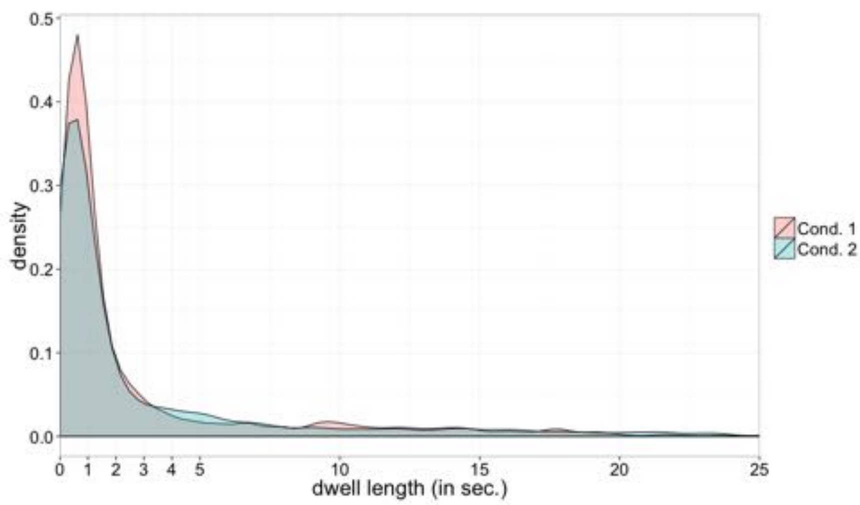


Figure 15.2.: Distribution of dwell lengths across experimental conditions. Please note that the graph has been trimmed at the value of 25 seconds. The 'long tail' of the data spreads to 157 seconds.

value is due to grid cells representing artworks' locations. This value (7.3% in Condition 1 and 7.0% in Cond. 2), can be interpreted as the proportion of eye movement activity which *should* fall on those wall segments, if this activity was dependent solely on the relative visibility of the given wall segment. Chi-Square Goodness-of-Fit test was used to compare the *expected* proportions to the *observed* proportions. The observed proportions were significantly inconsistent with the expected distribution of proportions both in Condition 1 ($\chi^2(1)$=7.94, p=0.005) and in Condition 2 ($\chi^2(1)$=6.43, p=0.011).

Analysis of *viewing sequence similarity* was performed as described earlier in Section 8.3 (p. 120) to account for Hypothesis (A4). The analysis resulted in *Mean Normalised Levenshtein Distance (MNLD)* values associated with each participant. *MNLD* describes how similar the viewing sequence performed by each participant is to all other viewing sequences within the given experimental condition. The value normalises for length and number of considered viewing sequences, and thus makes it possible to perform cross-conditional comparisons. As the distribution of those values was not normal, Wilcoxon rank sum test (also known as Mann-Whitney U test) was performed on *MNLD* values across the conditions. The result of the test was significant (W = 160, p = 0.016), indicating that *MNLDs* in Condition 2 were significantly lower from Cond. 1.

Consistently with Hypothesis (A2) mean *Total Dwell Time* per picture was not significantly different from 30 sec. (M=31.71 sec.; one sample t-test: $t(13)$=1.33, p=0.207), which is an average value observed in previous museum-based experiments.

For linear mixed-effect modelling, `lme4` R package was used (Bates, Maechler, Bolker & Walker, 2014). The following formula was used to describe the initial model:

```
log(No. of Dwells) ~ (Iso. Area + VCA +
Iso. Jaggedness + Potential CoVis. +
Targeted CoVis. + Salience Rating)*cond +
View. Sequence + age + gender + Total Time Inside +
(1|ids) + (1+cond|pic) + (1|loc)
```

Model 15.1: Logarithm of Number of Dwells (preliminary formula; multiplication symbol indicates interaction effects)

The notation can be read as: 'natural logarithm of *Number of Dwells* (+1) is predicted by the fixed effects of: experimental condition, isovist area, VCA, isovist jaggedness, potential co-visibility, targeted visibility, salience rating, age, gender, and time inside the gallery; as well as the the interaction of experimental condition with the above predictors; and also by the random effect of: participant (ids), location (loc), and a random-slope and random effect of picture (pic) within condition'. The random effect introducing adjustment to the intercept is marked as (1|). Additionally, (1+cond|pic) indicates that a random effect grouped by picture is expected to have not only a random intercept, but also a random slope in two individual conditions. This means, that the random effect of individual pictures is possibly expected to be stronger in one condition and weaker in the other and to increase/decrease at different rates.

One of the model's statistical assumption is lack of collinearity[1] between separate predictors. The above model has reached *kappa* score of 13.8, indicating moderate collinearity[2] which suggests that the assumption might be violated. *Isovist Area* was a measure highly correlated with other spatial predictors (*VCA*, *Potential Co-Visibility*, *Targeted Co-Visibility*, and to a lesser extent, with *Isovist Jaggedness*). Additionally, *Targeted Co-Visibility* was correlated with *Potential Co-Visibility* (unsurprisingly, as they aim to describe a very similar concept), and with *VCA* (as it is derived from the VCA). Despite the fact that collinearity does not inflate Type I error rates, it can notably reduce the statistical power of the model, and makes it impossible to reason which particular factors contribute to explaining the variance of the outcome variable. A typical solution for such situations are dimension-reduction techniques, which allow the researcher to combine a number of correlated predictors into a single variable (Winter, 2013). However, in the case here presented, there is a considerable theoretical difference between the individual predictors. *Isovist Area* and *VCA*, even though derived from the same measuring technique, bear two different assumptions about the nature of human exploration of space (Section 5.3). The measures therefore cannot be combined in a single model. *VCA* had lower correlation values with other predictors, meaning that it is a more distinct concept than the generic *Isovist Area*. For this reason, *Isovist Area* was excluded from the model, and a more theoretically-pleasing factor of *VCA* was preserved. Additionally, *Potential Co-Visibility* could be excluded, as it represents a concept similar to *Targeted Co-Visibility*, but is less justified by being derived from the exact location of the picture, and not from the possible location of the observer. However, by being derived from the spatial analysis of grid cells belonging to *Visibility Catchment Areas* of individual artworks (Section 11), *Targeted Co-Visibility* unexpectedly remained highly correlated with *VCA* in all analysed datasets (e.g. $r = 0.63$, $p<.001$ in Experiment 1). This indicates, that as *Visibility Catchment Area* of an individual artwork gets larger, its *Targeted Co-Visibility* (i.e. mean number of pictures visible from inside that

1 Collinearity describes a situation when two or more predictors inside a model are highly correlated with each other. For example, trying to explain the ripeness of harvested tomatoes from the record of rainy and sunny days would depict such as a situation: it rarely shines and rains at the same time. Thus, a researcher presented with a significant model would not be able to establish whether it is the increased sun or decreased rainfall that actually improved the taste of vegetables.

2 *Kappa* above 30 is associated with troubling collinearity, below 30 with moderate collinearity, and below 6 with no collinearity (Baayen, 2008). Alternative sources suggest that *kappa* below 10 is sufficient.

VCA) also increases. Despite the fact that concepts are dissimilar and were designed to describe different aspects of space, in the case here presented they do not. It can be explained by the fact that *VCAs* get wider as they get further away from the picture itself. As they extend into open space, the likelihood that they will overlap with isovists of other pictures gets higher. *Potential Co-Visibility* measure is free of that bias. Therefore, despite representing *Co-Visibility* as it would be seen by someone standing at the location of the analysed picture it still represents the potential of engaging with alternative pictures, especially for a person standing close to the work.

Following these steps, the following formula was constructed:

```
log(No. of Dwells) ~ (VCA + Iso. Jaggedness +
Targeted CoVis. + Salience Rating)*cond +
View. Sequence + age + gender + Total Time Inside +
(1|ids) + (1+cond|pic) + (1|loc)
```

Model 15.2: Logarithm of Number of Dwells (preliminary formula)

This revised model has reached *kappa* score of 6.79 and was a subject of *step* function from `lmerTest` R package (Kuznetsova, Bruun Brockhoff & Haubo Bojesen Christensen, 2014), which performs a backward elimination of all non-significant effects from a mixed-effect model [3]. The suggested, reduced model

```
log(No. of Dwells) ~ VCA + View. Sequence +
Total Time Inside + (1| ids) + (1 | loc)
```

Model 15.3: Logarithm of Number of Dwells

has reached *kappa* score of 4.3. This is the best-fit model, given the available data, meaning that all the remaining factors (*VCA*, *ET Sequence*, *Total Time Spent Inside*, random by-subject effect, and random by-location effect) significantly predicted normal logarithm of *Number of Dwells*. *Potential Co-Visibility* and *Isovist Jaggedness* effects were approaching significance. The intercept estimates of *VCA* and *Total Time Spent Inside* were positive, meaning that an increase in each of those predictor variables resulted in an increase in the outcome

3 The reader should note that automatic stepwise model selection procedures should be used with caution, as they tend to produce overfitting models that do not generalise well (Judd, McClelland & Ryan, 2011). Throughout this thesis, *step* function is treated as a suggestion, but not necessarily the best final solution. Where theoretical inconsistencies were returned by the function, the model was adjusted manually. Moreover, the 'best-fit' models here presented should not be directly used for making predictions in the future., as—for the purpose of clarity—the formulae were often cleaned of non-significant random effects. For future predictions however, random effects should always be included (if justified by the experimental design) to avoid violating the independence assumption of the model. As a result, the formulae here presented can be read as a set of predictors which were significant for the considered outcome variable, in the considered dataset. Their aim is exploratory, and should be further adjusted for predictive purposes.

	Model 15.3
(Intercept)	1.95 (0.08)***
VCA	0.29 (0.04)***
Total Time Inside	0.42 (0.06)***
View. Sequence	−0.01 (0.01)*
AIC	411.68
BIC	439.67
Log Likelihood	-198.84
Num. obs.	403
Num. groups: ids	29
Num. groups: loc	28
Variance: ids.(Intercept)	0.11
Variance: loc.(Intercept)	0.03
Variance: Residual	0.11

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.1.: Statistical description of the Model 15.3. It should be interpreted as follows: each predictor is associated with an estimate, its standard error (in brackets), and a significance symbol. The rule for interpreting similar models is that 1 increase in the unit of the predictor causes an increase in the output variable by the value of the estimate. Since in majority of the cases here described (as reported), predictor variables were standardised, their unit is 1 Standard Deviation. Moreover, the size of random effects can be interpreted according to their variance (bottom of the table) compared to the residual (error) variance. Residual variance indicates how much the data varies in a manner unexplained by the predictors. Random effect variance indicates how much variability there is in the pre-specified random effects.

variable[4]. Figure 15.3 presents these effects visually and Table 15.1 demonstrates the numerical output.

The R^2 approximation values (P. C. D. Johnson, 2014; Nakagawa & Schielzeth, 2013) were obtained using `MuMIn` R package (Bartoń, 2014). In case of mixed-effect models, two types of R^2 are calculated: *Marginal* R^2 indicates the proportion of variance in the dependent variable explained by the fixed effects alone (similarly to classic linear models), and *Conditional* R^2 by the fixed effects and random effects jointly. For the purpose of generalisation of the research results, *Marginal* R^2 is more relevant, as fixed effects are the ones that specify the

4 A negative intercept would mean that an increase in the predictor variable resulted in a decrease in the outcome variable - as it was the case with *ET Sequence*.

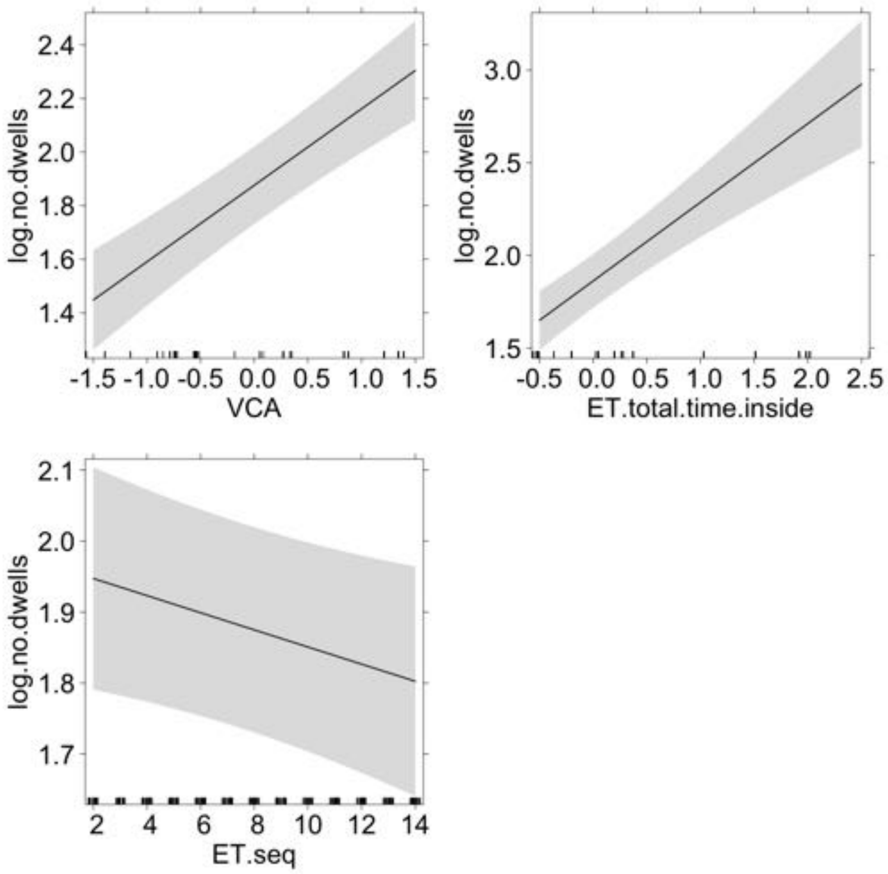


Figure 15.3.: Effect plots for Model 15.3. The graphics should be interpreted as follows: the lines indicate the predicted linear effect of the specific predictor (usually described on the horizontal axis) and the output variable (typically on the vertical axis). The grey area around the line shows 95% confidence limits based on Standard Error of the mean of the normal distribution. Larger grey area suggests that the prediction is relatively uncertain. A narrow grey area indicates that the prediction is relatively certain (i.e. there is a 95% chance that the actual predicted value falls somewhere within this grey area; compare the confidence limits for *VCA* and *ET.seq* with the numerical values of their effect sizes in the tabular description of the model). Lastly, the unevenly distributed ticks at the bottom horizontal axis indicate the spread of the factual data points in the analysed dataset. Denser areas of ticks mean that the model had more data available at this particular value.

particular factor affecting the outcome value[5]. However, *Conditional R^2* can indicate how well the model performs under the existing, unexplained variability. In the context of the current study, this is especially important, as there are elements affecting the experience of an art gallery visit which might be difficult to quantify and generalise with any external measures for the purpose of making predictions (such as personal preferences, previous memories related to the viewed exhibition, higher average density of visitors in front of particular artwork locations, and others). *Conditional R^2* approaching 1 would demonstrate that almost all variability in the data unexplained by the fixed effects does lie in the specified random effect factors (in case of Model 15.3 these are subjects and locations, as the by-picture random variability was negligible), but cannot be explained by the fixed effects specified in the model. Model 15.3 reached *Marginal R^2* of 0.52 and *Conditional R^2* of 0.79[6]. Homoscedasticity and normality of the model's residuals assumptions were not violated.

In order to establish whether the model explains the outcome variable better than its alternatives the model was compared with its alternative containing *Isovist Area* instead of *VCA* as the first predictor[7]. This modification decreased the overall performance of the model as indicated by higher AIC value (Akaike, 1992)[8].

Similar procedures have been employed to select models best predicting *Total Dwell Time*, *Picture-Switching*, and *Long-to-Short Dwell Ratio*.

Natural logarithm of *Total Dwell Time* (+1) was best predicted by the formula:

```
log(Total Dwell Time) ~ VCA + Iso. Jaggedness +
Total Time Inside + (1|ids)
```

Model 15.4: Logarithm of Total Dwell Time

Adding random by-location effect did not significantly increase the fit of the model, but the result of the likelihood ratio test was approaching significance ($\chi^2(1)=3.56$, p=0.593). The lack of significant random by-location effect in Model 15.4 would indicate that the fixed-effect spatial predictors included in the formula already sufficiently described the majority of the by-location variability. The current result is inconclusive.

5 In other words, fixed effects are those which are known to the researcher and have been quantified.

6 Perhaps the most popular 'rule of thumb' for interpreting R^2 effect sizes was given by Cohen (1988). He proposed three threshold values as an indicator of how big the effect is: small (R^2 = 0.02), medium (R^2 = 0.13), or large (R^2 = 0.26).

7 As it was discussed previously, the measures could not be used together in a single model because of their correlation.

8 Note that R^2 values alone should not be used for model comparison due to their approximate characteristic in mixed effect models (Bartoń, 2014) and model selection should be supported by AIC value judgement (Akaike, 1992) which also penalises for a higher number of included factors.

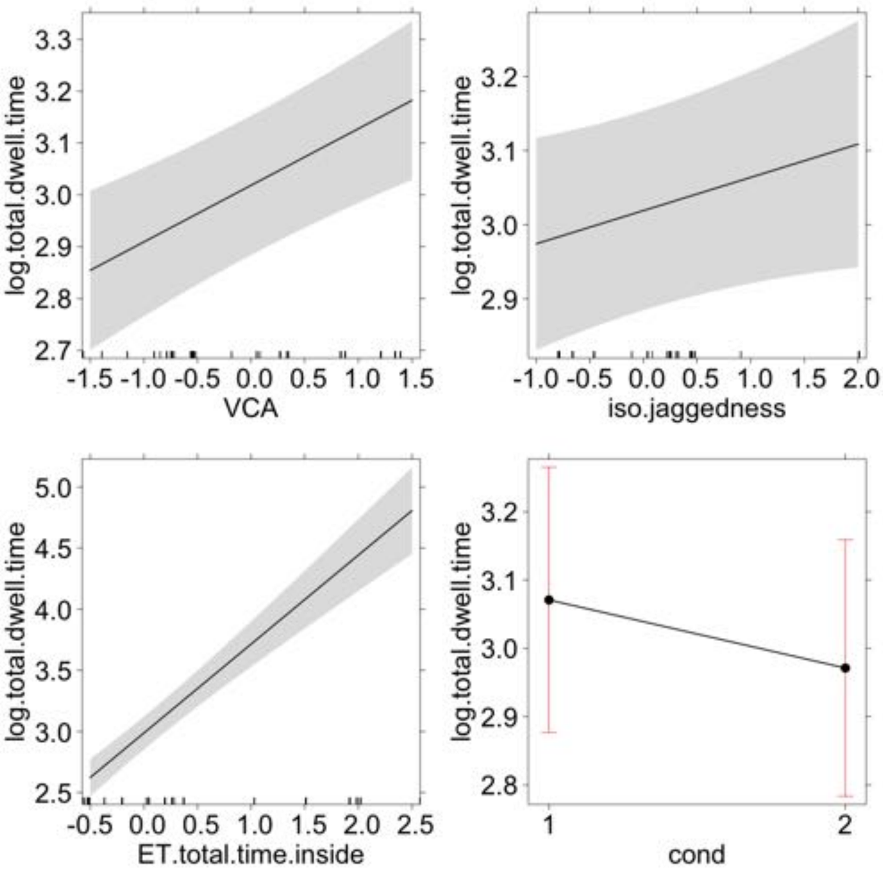


Figure 15.4.: Visualisation of the effects for Model 15.4.

Substituting the predictor of *VCA* by *Isovist Area* did not improve the overall performance of the model (as indicated by nearly identical AIC, and identical R^2 values). Homoscedasticity of residuals was not violated and their distribution was close to normal, except a single outlier. *Kappa* score of the final model was 1.4, *Marginal* R^2 was 0.62 and *Conditional* R^2 was 0.75. Table 15.2 describes the model in detail and Figure 15.4 visualises the effects.

To describe the dynamics of the oculomotor behaviour, firstly *Picture-Switching* ratio (indicating the proportion of dwells at each picture which were followed by an immediate fixation at another art object) was predicted by the following model:

```
Picture-Switching ~ VCA + Potential CoVis. +
  (1|ids) + (1|loc)
```

Model 15.5: Picture-Switching Ratio

Substituting the predictor of *VCA* by *Isovist Area* did not improve the overall performance of the model. Normality and homoscedasticity of residuals were not violated. *Kappa* score of the final model was 1.3, *Marginal* R^2 was 0.21 and *Conditional* R^2 was 0.5. Table 15.3 and Figure 15.5 summarise the model.

	Model 15.4
(Intercept)	$3.04\ (0.10)^{***}$
VCA	$0.11\ (0.03)^{***}$
Iso. Jaggedness	$0.04\ (0.03)^{\cdot}$
Total Time Inside	$0.73\ (0.07)^{***}$
cond2	$-0.10\ (0.14)$
AIC	653.87
BIC	681.92
Log Likelihood	-319.94
Num. obs.	406
Num. groups: ids	29
Variance: ids.(Intercept)	0.12
Variance: Residual	0.23

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.2.: Model 15.4.

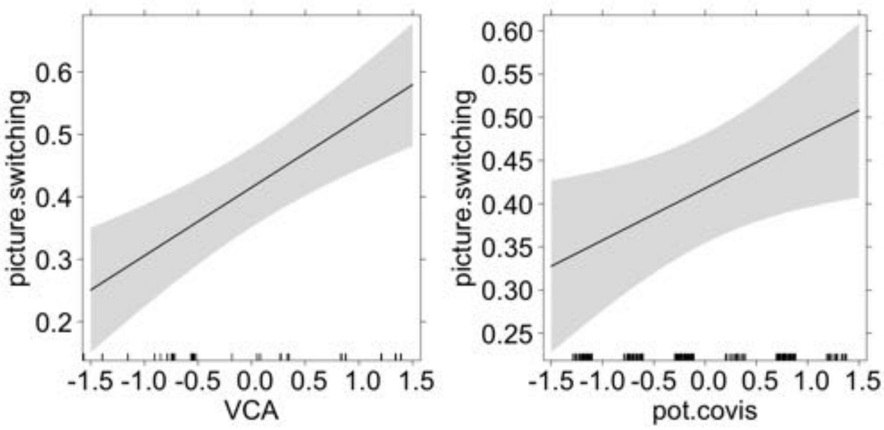


Figure 15.5.: Visualisation of the effects for Model 15.5.

	Model 15.5
(Intercept)	0.42 (0.03)***
VCA	0.11 (0.03)***
Potential CoVis.	0.06 (0.03)*
AIC	2.89
BIC	26.87
Log Likelihood	4.56
Num. obs.	402
Num. groups: ids	29
Num. groups: loc	28
Variance: ids.(Intercept)	0.01
Variance: loc.(Intercept)	0.01
Variance: Residual	0.04

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.3.: Model 15.5.

Long-to-Short Dwell Ratio measure had a bi-modal distribution caused by a noticeable proportion of interaction consisting of long dwells only. It could have not been transformed to approach normality through traditional data transformation techniques. Thus, the general fit of the model might be reduced, but the assumption of the model is not violated, as long as the distribution of residuals remains normal[9].

The following formula provided the best explanatory performance[10]:

```
Long-Short Dwell Ratio ~ cond + VCA + Iso. Jaggedness+
 Potential CoVis. + Salience Rating + gender +
 (1 | ids) + (1 | loc) + cond*Salience Rating
```

Model 15.6: Long-to-Short Dwell Ratio

Substituting the predictor of *VCA* by *Isovist Area* again did not improve the overall performance of the model. The impact of *ET Sequence* was significant but negligible, with each next rank order in the viewing sequence increasing *Long-to-Short Dwell ratio* by a value of 0.007[11]. Since the significance of such small effect might be achieved due to a large number of degrees of freedom (182 for this particular effect), it was not included in the final model. Residuals' distribution was nearly-normal and their homoscedasticity was not violated. *Kappa* score of the final model was 3.6, *Marginal* R^2 was 0.25 and *Con-*

9 Mixed effect models are generally considered robust under the lack of normality (Winter, 2013).

10 Main effect of *Condition*was not significant, but needs to be included for the significant interaction factor to be considered.

11 This means, that the difference in *Long-to-Short Dwell ratio* caused by the sequence of viewing was only about 0.01 between the first and the last picture in the sequence, and therefore negligible.

	Model 15.6
(Intercept)	$0.44\ (0.04)^{***}$
cond2	$0.03\ (0.04)$
VCA	$-0.08\ (0.02)^{***}$
Iso. Jaggedness	$-0.04\ (0.02)^{**}$
Potential CoVis.	$-0.05\ (0.02)^{**}$
Salience Rating	$0.04\ (0.01)^{**}$
gender-male	$-0.08\ (0.03)^{*}$
cond2:Salience Rating	$-0.05\ (0.02)^{**}$
AIC	-28.24
BIC	15.75
Log Likelihood	25.12
Num. obs.	403
Num. groups: ids	29
Num. groups: loc	28
Variance: ids.(Intercept)	0.00
Variance: loc.(Intercept)	0.00
Variance: Residual	0.04

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.4.: Model 15.6.

ditional R^2 was 0.39. Table 15.4 describes the model in detail and Figure 15.6 demonstrates the significant effects visually.

The four above models are key to understanding how space affects the quantity and dynamics of art viewers' oculomotor behaviour. Based on them, additional models were constructed to consider more specific elements of the phenomenon.

In order to account for Hypothesis (B7), the model predicting *Number of Dwells* has been re-run with one additional factor - *Long Dwells Level*. The factor was constructed based on the *Long-to-Short Dwell ratio* variable: all interactions between individual users and separate pictures which consisted of *long dwells* in more than 50% were assigned 'high' *Long Dwells Level*, and those which consisted primarily of *short dwells* (i.e. their *Long-to-Short Dwell ratio* <= 0.5) were assigned 'low' level. *Long Dwells Level* (LDL) has been included as a fixed-effect, as well as a factor interacting with all other fixed-effect variables, resulting in the following formula:

```
log(No. of Dwells) ~ LDL + VCA + Total Time Inside+
View. Sequence + (1|ids) + (1|loc) +
LDL*(VCA + Total Time Inside + View. Sequence)
```

Model 15.7: Logarithm of Number of Dwells analysed with respect to Long Dwells Level

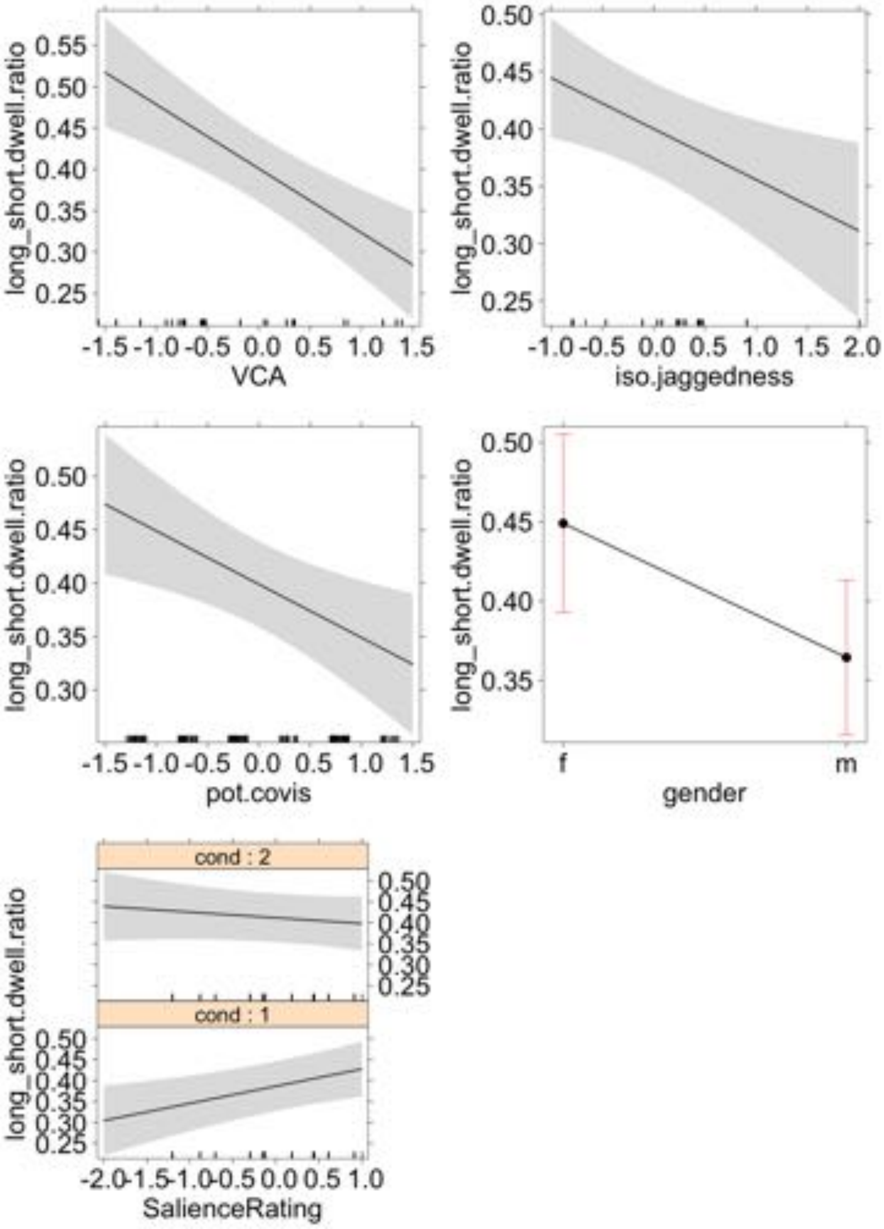


Figure 15.6.: Visualisation of the effects for Model 15.6.

Model's assumptions were not violated but none of the interaction effects were significant.

To account for Hypothesis (B8), Model 15.6 has been re-run with a newly constructed variable - *Number of Pictures on a Single Wall*. The minimum value of this variable was 1 (for locations which had no adjacent pictures on the same wall) and the maximum value was 4 for location x208-x211 in Condition 2). The following amended model[12]:

```
Long-Short Dwell Ratio ~ cond + VCA + Iso. Jaggedness+
Pictures per Wall + Salience Rating + gender +
(1 | ids) + (1 | loc) + cond*Salience Rating
```

Model 15.8: Long-to-Short Dwell Ratio with respect to Number of Pictures Per Wall

decreased the overall fit compared to 15.6, therefore the detailed statistics are omitted.

In context of suggestions made by Wineman et al. (2006) and Hypothesis (D3), space shall have a higher influence on those visitors, who remain inside the museum for longer. To account for this, participants were median-split based on the time they spent inside the gallery and this factor (named *inside factorial*) has been included into two models considering the dynamics of the visual attention - Models 15.5 and 15.6 as a fixed effect interacting with spatial predictors. Neither for *Picture-Switching ratio* nor for *Long-to-Short Dwell ratio* was this interaction significant, thus the models are not reported in detail. The procedure could not be conducted for models predicting quantitive aspects of the oculomotor behaviour (Models 15.3 and 15.4), as in them, *Time Spent Inside* has been a highly significant factor linearly predicting the outcome variable already.

To tackle this limitation, two time-independent measures: *Normalised Number of Dwells* and *Normalised Total Dwell Time* were also modelled as outcome variables, based on the procedure described at the beginning of the current section, but without the *Time Spent Inside* predictor.

Firstly, logarithmic (+1) transformation of *Normalised Number of Dwells* was predicted by the following model:

```
log(Norm. No. of Dwells) ~ cond + inside +
View. Sequence + VCA + Targeted CoVis. +
(1 | loc) + inside(factorial)*(View. Sequence+VCA)
```

Model 15.9: Logarithm of Normalised Number of Dwells

The model reached *Kappa* score of 11.3, *Marginal* R^2 of 0.47 and *Conditional* R^2 of 0.55. The interaction effect of *inside* factor (with values 'long' or 'short' depending on the median-split of the visitors' *Time Spent Inside*) with *VCA* and *ET Sequence* was significant. Figure 15.7

12 Due to high correlation of *pictures per wall* with *Potential Co-Visibility*, only one of the two measures can be used in a single model.

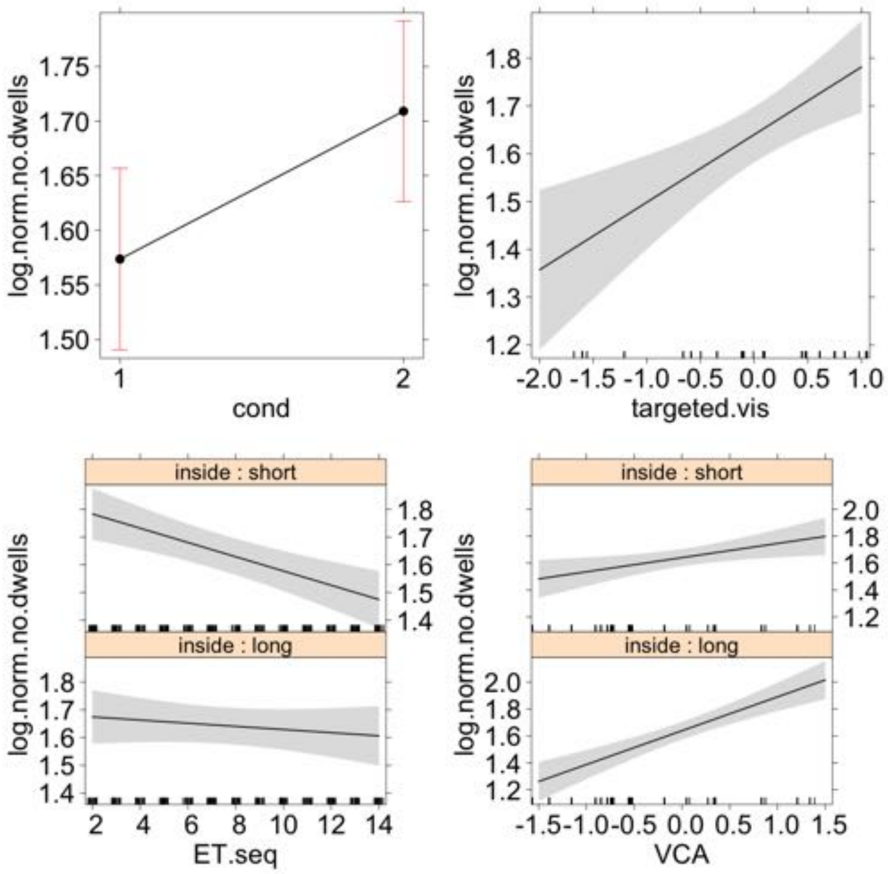


Figure 15.7.: Visualisation of the effects for Model 15.9.

visualises the model and Table 15.5 presents the descriptives. As can be noticed, the influence of *VCA* on the proportion of dwells a visitor would allocate to a given picture was more noticeable if a person stayed in the gallery for longer. Interestingly, pictures seen further along the individual viewing sequence were allocated smaller proportion of dwells by those participants who spent little time inside, but the factor had on impact on participants who were inside for longer.

Secondly, logarithmic (+1) transformation of *Normalised Total Dwell Time* was explained by the model:

```
log(Norm. Total Dwell Time) ~ inside + VCA +
Potential CoVis. + age + (1|loc) +
inside(factorial)*Potential CoVis.
```

Model 15.10: Logarithm of Normalised Total Dwell Time

The model's *Kappa* was 3.2, *Marginal* R^2 was 0.1 and *Conditional* R^2 was 0.15. Figure 15.8 visualises the relations between significant factors and Table X provides the statistical description. The interaction effect of factorial time spent *inside* and *Potential Co-Visibility* demonstrates that higher number of co-visible pictures did cause

	Model 15.9
(Intercept)	$1.61\ (0.07)^{***}$
cond2	$0.14\ (0.06)^{*}$
insideshort	$0.15\ (0.07)^{*}$
View. Sequence	$-0.01\ (0.01)$
VCA	$0.25\ (0.04)^{***}$
Targeted CoVis.	$0.14\ (0.04)^{***}$
insideshort:View. Sequence	$-0.02\ (0.01)^{*}$
insideshort:VCA	$-0.15\ (0.03)^{***}$
AIC	273.00
BIC	312.99
Log Likelihood	-126.50
Num. obs.	403
Num. groups: loc	28
Variance: loc.(Intercept)	0.02
Variance: Residual	0.09

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.5.: Model 15.9.

longer-exploring participants to spent proportionally more time fixating on the given artwork, but had an opposite impact on participants who were in the gallery shorter than the median time.

15.3 RECOGNITION MEMORY

The first part of the Recognition Memory analysis considered all responses to the Recognition Task (without the Eye-Tracking data and even for those participants, for whom ET data was unavailable). The following linear mixed-effect model was proposed to predict inverted, standardised and mean-centred Reaction Times[13]:

```
RT(inv) ~ cond + VCA + Iso. Jaggedness + Potential CoVis.+
Total Time Inside + Salience Rating + age +
gender + RTtrial + proceedingRT + proceedingRTacc+
(1|ids) + (1|pic) + (1|loc)
```

Model 15.11: Reaction Times (inverted)

Table 15.7 presents the relevant statistics. *Kappa* score of the model was 3.7, Marginal R^2 was 0.18 and Conditional R^2 was 0.47. As it can

13 Please note that since these models are constructed with a more careful selection of predictors guided by a deeper analysis in the previous section, no step-wise procedure is employed. As a result, not all predictors listed in the formula are significant and the reader should always refer to the relevant table for that information. Throughout the thesis the step-wise predictor selection is only used for exploring initial set of relevant factors. Where the `step` function is not mentioned, the formula of the model **is not** equivalent to statistical significance of the listed predictors.

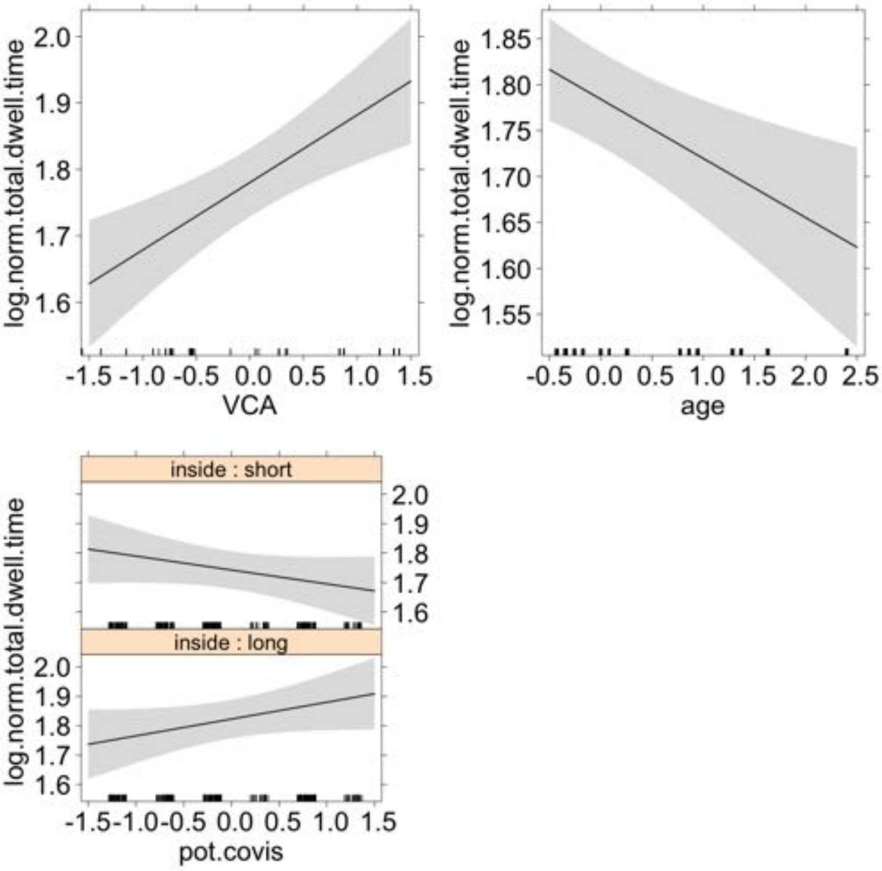


Figure 15.8.: Visualisation of the effects for Model 15.10.

	Model 15.10
(Intercept)	1.82 (0.03)***
insideshort	-0.08 (0.04)*
VCA	0.10 (0.03)***
Potential CoVis.	0.06 (0.03)$^{\cdot}$
age	-0.06 (0.02)**
insideshort:Potential CoVis.	-0.10 (0.04)**
AIC	433.64
BIC	465.69
Log Likelihood	-208.82
Num. obs.	406
Num. groups: loc	28
Variance: loc.(Intercept)	0.01
Variance: Residual	0.15

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.6.: Model 15.10.

be noticed in the detailed model description, fixed-effects other than those related to the nature of the Recognition Task (*RTtrial*, *proceedingRT*, and *proceedingRTacc*) were not significant. *Time spent inside* and *gender* effects explaining the by-participant variability were relatively high, but not significant. Spatio-visual factors, including the main effect of *Condition* failed to reach significance. The random by-location variance was also low compared to other random effects and to the residuals' variance. This means that there was little consistent variation in the output variable across separate locations. Consulting the random-effects structure further leads to a conclusion that the variance accounted to unexplained by-participant differences (marked in the model as (1 | ids)) is particularly high when to compare its Standard Deviation to Standard Errors of individual fixed-effects. This suggests that the by-participant variability is the main factor driving the overall performance of the model. Since linear mixed-effect models are constructed based on the model's *overall* performance, random effects might 'steal' unexplained variance from the fixed-effect factors if this leads to a higher Conditional R^2 value. Simultaneously, while calculating the by-participant variability, the model does not consider the variability within separate conditions, but the variability across all participants who have provided a response. For low sample sizes, like in this study, it is probable that this high variance in by-participant response reduces the fixed-effect of *Condition*. In other words, all people who took part in the experiment differ so much between each other, that this variability is larger than the variability which has been caused by *Condition*. This, however, is not equivalent to saying that the effect of *Condition* does not exist. It simply might not be strong enough to reach significance under the presence of such high by-participant variability in Reaction Times.

To test for the main effect of *Condition*, while accounting for the by-participant variance *within* each condition separately, another method was required. Since multiple responses were collected from each participant, using a standard t-test on this dataset would violate the assumption of independence. There are two solutions to this problem: a) aggregating the data into participant means (so that only one value is associated with each participant and therefore the independence assumption is met), or b) using a technique which accounts for clustered data, but does not generalise the by-participant variance across the conditions[14]. A statistical technique which meets this second requirement is a t-test for clustered data (C. Roberts & Sibbald, 1998), as it controls for intercluster correlation. For the purpose of the current analysis, inverted Reaction Times were used as dependent variable, *participant id* as clustering factor, and *Condi-*

14 The latter option should be considered superior due to higher statistical power.

	Model 15.11
(Intercept)	-0.25 (0.22)
cond2	-0.25 (0.21)
VCA	0.05 (0.05)
Iso. Jaggedness	-0.04 (0.05)
Potential CoVis.	-0.04 (0.05)
Total Time Inside	0.15 (0.11)
Salience Rating	0.07 (0.08)
age	-0.10 (0.10)
gender-male	0.34 (0.21)
RTtrial	-0.22 (0.04)***
proceedingRT	-0.14 (0.06)*
proceedingRTacc-TRUE	0.16 (0.10)$^{\cdot}$
AIC	813.36
BIC	873.10
Log Likelihood	-390.68
Num. obs.	309
Num. groups: ids	31
Num. groups: loc	28
Num. groups: pic	14
Variance: ids.(Intercept)	0.24
Variance: loc.(Intercept)	0.01
Variance: pic.(Intercept)	0.05
Variance: Residual	0.54

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.7.: Model 15.11.

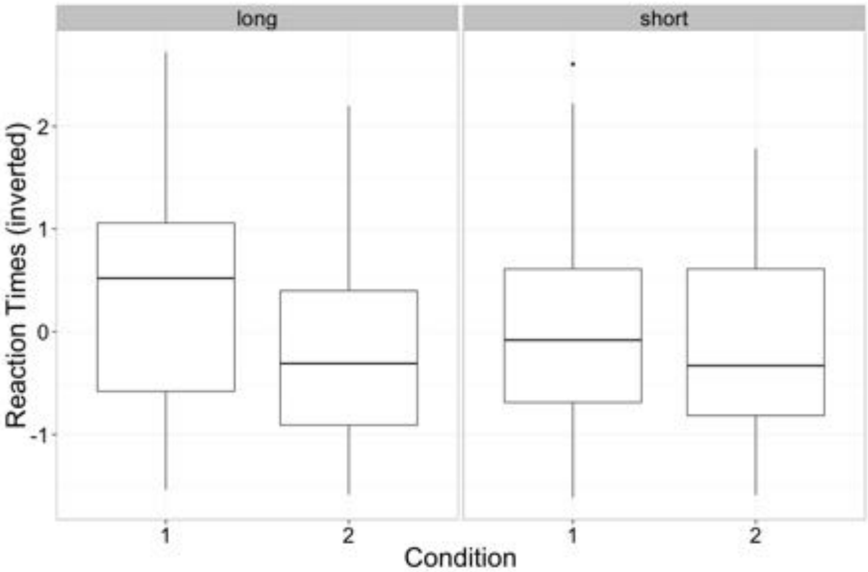


Figure 15.9.: Visualisation of the Recognition Memory test split by the time spent inside the gallery.

tion as the grouping factor[15]. `T.test.cluster` function from `Hmisc` R package was used (Harrell Jr & Dupont, 2006). *Inverted RTs* were mean-centred and standardised, so that the difference between means of two groups was equivalent to the effect size described in Standard Deviations (Cohen's D). The analysis showed a significant effect of *Condition* on *Reaction Times*, with participants in Condition 1 answering faster than participants in Condition 2 (p=.049, Cohen's D = -0.39, CI_d[-0.78, -0.001]).

To account for Hypothesis (D3), a similar comparison was made on a dataset split on the median of *Time Spent Inside*. This comparison showed, that the significance of the effect is primarily driven by the subset of visitors who stayed 'long' (i.e. longer than median) inside the gallery (p=.042). When the subset of participants exploring the gallery for a 'short' (shorter than median) period of time was analysed, the effect of *condition* fails to reach significance (p=0.598). Figure 15.9 demonstrates this relation graphically.

Such framing of the results, however, limits our ability to infer the exact reasons for those faster Reaction Times. Also, high by-participant variance shown in the linear mixed-effect model puts in doubt the applicability of the finding if the effect is weaker than the by-participant variations in the population. Although this might be due to the low sample size, it implies the question of the suitability of Reaction Time measures for real-life spatial cognition studies, where the observed effects might be relatively weak, and not often observed

15 One participant who had less than 2 responses valid was removed from this analysis due to the requirements of the test.

on an object-by-object basis. A further argument in this respect are low by-picture and by-location random-effect variances, showing that there was little repeatability in the response patterns to the same pictures or to the same locations.

There exists a possibility that this by-participant variation could be explained with the addition of eye-tracking data, which can precisely describe individual differences in the visual attention of separate participants. The disadvantage of this approach is that such a model can only be fit to the subset of the entire dataset, as there were some participants for whom the eye-tracking recordings were not available (two participants in the case of this study). For such a subset, the above model was modified to include the available eye-tracking variables, resulting in the following formula (for details consult Table 15.8):

```
RT(inv) ~ Quantity of ET + View. Sequence +
Time to First Fix. + Picture-Switching +
Long-Short Dwell Ratio + cond + VCA +
Iso. Jaggedness + Potential CoVis. +
Total Time Inside + Engagement Ratio +
Salience Rating + age + gender +
RTtrial + proceedingRT + proceedingRTacc +
(1|ids) + (1|pic) + (1|loc)
```

Model 15.12: Reaction Times (inverted) with respect to Eye Movement

Kappa score was 17.7, Marginal R^2 was 0.19 and Conditional R^2 was 0.45. The by-participants random-effect variance (1|ids) was only marginally decreased compared to the previous (15.11) model, and none of the eye-tracking effects even approached significance (with the limited exception of *Long-to-Short Dwell Ratio*). This can be interpreted in two ways: either there is no connection between the oculomotor behaviour and Recognition Memory, or Reaction Time procedure is not reliable for lower sample sizes (further limited by the lack of ET data) and high by-participant variability in the context of the current study. A lager sample size in further studies might help to tackle these issues.

To further investigate the dataset, additional analysis of *RT accuracy* was conducted. This output variable is binomial (the response can be either 'correct' or 'incorrect'), and thus a modification of linear mixed-effect models known as mixed logit models was used to predict it. Mixed logit models decrease Type I and Type II error risk compared to classical ANOVAs for categorical data and allow the researcher to describe the results in the standard language of linear modelling by the 'output predictors' syntax (Jaeger, 2008). All relations between random- and fixed-effect factors discussed earlier for linear mixed-effect modelling remain valid. The analysis was conducted with `glmer(family=bimodal)` function of `lme4` R package

	Model 15.12
(Intercept)	−0.86 (0.97)
Quantity of ET	0.05 (0.05)
View. Sequence	0.01 (0.02)
Time to First Fix.	−0.07 (0.07)
Picture-Switching	−0.00 (0.06)
Long-Short Dwell Ratio	0.07 (0.05)
cond2	−0.22 (0.24)
VCA	0.08 (0.06)
Iso. Jaggedness	−0.02 (0.05)
Potential CoVis.	−0.04 (0.06)
Total Time Inside	0.12 (0.12)
Engagement Ratio	0.68 (1.12)
Salience Rating	0.06 (0.08)
age	−0.03 (0.11)
gender-male	0.32 (0.21)
RTtrial	$-0.22\ (0.05)^{***}$
proceedingRT	$-0.14\ (0.06)^{*}$
proceedingRTacc-TRUE	0.14 (0.10)
AIC	804.35
BIC	885.32
Log Likelihood	-380.18
Num. obs.	293
Num. groups: ids	29
Num. groups: loc	28
Num. groups: pic	14
Variance: ids.(Intercept)	0.21
Variance: loc.(Intercept)	0.00
Variance: pic.(Intercept)	0.05
Variance: Residual	0.56

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.8.: Model 15.12.

	Model 15.13
(Intercept)	1.54 (0.32)***
cond2	0.11 (0.41)
VCA	−0.09 (0.14)
Iso. Jaggedness	0.05 (0.14)
Potential CoVis.	−0.05 (0.14)
Total Time Inside	0.71 (0.23)**
Salience Rating	0.12 (0.16)
AIC	432.39
BIC	473.12
Log Likelihood	-206.20
Num. obs.	434
Num. groups: ids	31
Num. groups: loc	28
Num. groups: pic	14
Variance: ids.(Intercept)	0.63
Variance: loc.(Intercept)	0.00
Variance: pic.(Intercept)	0.14
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.9.: Model 15.13.

(Bates et al., 2014). Effect sizes are estimated with the modification of *r.squaredGLMM* function from `MuMIn` package made by Lefcheck and Casallas (2014).

Running the following model for all participants (i.e. with no ET data included):

```
RT Accuracy ~ cond + VCA + Iso. Jaggedness +
Potential CoVis. + Total Time Inside + Salience Rating+
(1|ids) + (1|pic) + (1|loc)
```

Model 15.13: Recognition Test Accuracy

resulted in Marginal R^2 of 0.11, Conditional R^2 of 0.28 and detailed statistics displayed in Table 15.9. As can be noticed, *Time Spent Inside* was the only significant fixed-effect predictor of *RT accuracy*. Little by-location random-effect variance suggests negligible consistent by-location variability in *RT accuracy*, although the by-picture variance is relatively high. Similarly to the analysis of *RT(inverted)* measure, the by-participant variation was the most impactful.

To explain the influence of *Time Spent Inside*, another analysis of *RT accuracy* was conducted, this time considering the Eye-Tracking data available for the subset of all participants. Because *Time Spent Inside* is a measure highly correlated with *Number of Dwells* (r = 0.53) and *Total Dwell Time* (r = 0.70), three separate models were run (each containing

	Model 15.14
(Intercept)	1.83 (0.34)***
Total Dwell Time	1.77 (0.46)***
Time to First Fix.	0.12 (0.24)
Picture-Switching	−0.17 (0.16)
Long-Short Dwell Ratio	0.18 (0.16)
cond2	0.34 (0.39)
VCA	−0.04 (0.17)
Iso. Jaggedness	0.04 (0.15)
Potential CoVis.	0.03 (0.16)
AIC	379.65
BIC	427.61
Log Likelihood	-177.83
Num. obs.	402
Num. groups: ids	29
Num. groups: loc	28
Num. groups: pic	14
Variance: ids.(Intercept)	0.37
Variance: loc.(Intercept)	0.00
Variance: pic.(Intercept)	0.32
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.10.: Model 15.14.

only one of the above variables additionally to other spatio-visual, and eye-tracking predictors). The model performing the best was the one involving *Total Dwell Time*, expressed by the following formula:

```
RT Accuracy ∼ Total Dwell Time + Time to First Fix.+
Picture-Switching + Long-Short Dwell Ratio + cond+
VCA + Iso. Jaggedness + Potential CoVis. +
(1|ids) + (1|pic) + (1|loc)
```

Model 15.14: Recognition Test Accuracy with respect to Eye Movement

Marginal R^2 was 0.67, Conditional R^2 was 0.73 and other statistics are described in Table 15.10. The clearly visible difference in effect size compared to Model 15.13 (p. 191) comes from including the eye movement measures in the analysis, justifying their predictive power of Recognition Memory accuracy.

15.4 SPATIAL MEMORY

As previously mentioned in Section 10.1, the risk of using *Back-to-the-Wall* measure lies in the fact that it does not control for wall length.

For this reason, poorly performing participants might consistently attempt to choose longer walls for the miniature locations in order to increase one's chance of success. This would result in 'false positive' errors being overrepresented on longer walls compared to shorter walls.

To control for such situation, each wall length was linked to the correct number of pictures factually hung on it within each experimental condition[16]. So obtained dataset unsurprisingly did not show a linear relation between wall length and the number of pictures, since wall length was not used as a guide to choose original picture locations (Figure 15.10). Consequently, even if participants did extract wall length from the Miniature Task layout, using it as a cue would not increase their performance in the test.

For each participant, each picture miniature was assessed based on whether it was placed on the correct wall ('correct'), or on a wrong one ('false positive'). The resulting number of correct answers and errors were summarised for each wall (separately for both conditions) and divided by the number of participants in the given condition. This created a list of walls, separate for both conditions, with associated means of correct answers and 'false positive' errors. Mean number of correct answers did not show any consistent relation with wall length (Figure 15.11). The presence of such a relation would suggest that despite the lack of correlation between wall length and the actual number of pictures hanging on them (Figure 15.10), participants were more likely to provide correct responses to longer walls. The data collected in Experiment 1 did not show the presence of such a bias.

Nevertheless, the number of 'false positives' was positively correlated with wall length ($r(38) = .54$, $p<.001$; Figure 15.12). This suggests that when participants were not sure about the location of a picture, they were more likely to choose a longer wall for its placement. Although it is important to note that using wall length as a deliberate strategy for placement of poorly remembered objects is not the only possible explanation. It is equally likely that unsure participant would make a random choice, within the constrains offered by the layout. This would also result in 'false positive' error rate being significantly correlated with wall length, as such length represents the probability of making this choice as a random decision, relative to all other possibilities. The fact that the exact strategy used by poorly performing participants cannot be established without making addition enquiries, does not affect the reliability of the Miniature Task.

16 Wall lengths were obtained from the layout's CAD drawing. Walls containing exit and entrance to the gallery were not included. In a single case where an, otherwise continuous, wall is interrupted by a passage, the sum of two stretches was used to represent its total length. In all other cases a single stretch of a wall was consider between two points of its corners or endings.

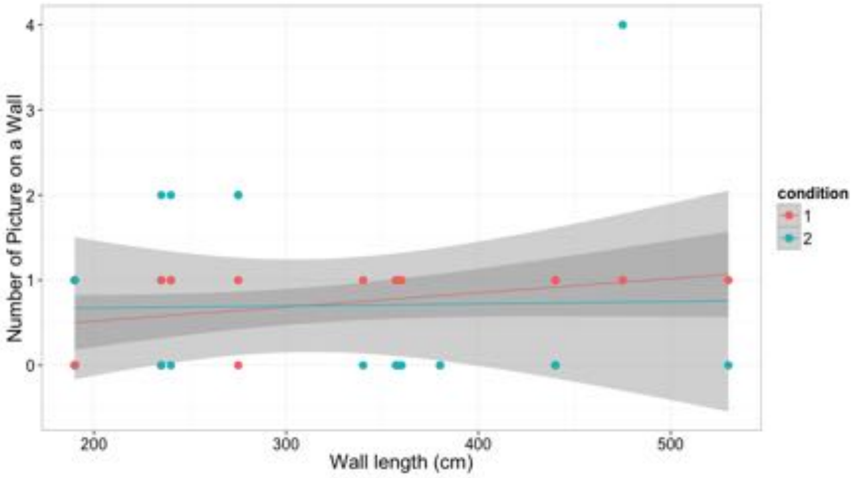


Figure 15.10.: Wall length plotted against the number of pictures contained on that wall.

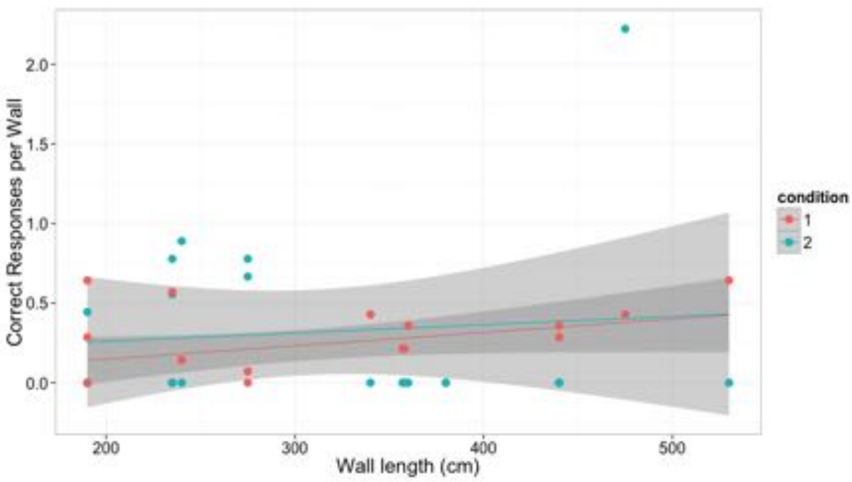


Figure 15.11.: Wall length plotted against the number of correct responses per wall.

	Model 15.15
(Intercept)	-1.42 $(0.61)^{*}$
cond2	1.23 $(0.60)^{*}$
VCA	0.19 (0.18)
Iso. Jaggedness	0.00 (0.17)
Potential CoVis.	-0.24 (0.18)
Total Time Inside	0.73 $(0.27)^{**}$
Salience Rating	0.02 (0.21)
age	-0.50 $(0.27)^{\cdot}$
gender-male	0.46 (0.55)
cond2:Salience Rating	0.34 (0.25)
AIC	505.08
BIC	557.60
Log Likelihood	-239.54
Num. obs.	420
Num. groups: ids	30
Num. groups: loc	28
Num. groups: pic	14
Variance: ids.(Intercept)	1.41
Variance: loc.(Intercept)	0.36
Variance: pic.(Intercept)	0.18
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.11.: Model 15.15.

This is because the correct answers—which are the subject of further analysis—seem unaffected by the wall length.

Similarly to Recognition Memory analysis, Spatial Memory can be explained based on the full dataset, or its subset limited to the participants for whom eye-tracking data is available. Firstly, the analysis without ET data was conducted for the entire dataset. The following model was formulated:

```
BttW ~ cond + VCA + Iso. Jaggedness + Potential CoVis.+
Total Time Inside + Salience Rating + age + gender+
(1|ids) + (1|pic) + (1|loc)
```

Model 15.15: Back-to-the-Wall Ratio

Marginal R^2 was 0.14 and Conditional R^2 was 0.46. Please refer to Table 15.11 for detailed statistics. Based on it, it is visible that participants' Spatial Memory was significantly better in Condition 2, increased with *Time Spent Inside* and decreased with *age*. The cross-conditional result is contrary to that observed for Recognition Memory, where it was Cond. 1 which resulted in better performance.

	Model 15.16
(Intercept)	$-0.98\ (0.50)^{*}$
Quantity of ET	0.06 (0.15)
Time to First Fix.	$-0.27\ (0.19)$
Picture-Switching	0.20 (0.18)
Long-Short Dwell Ratio	0.21 (0.17)
cond2	1.09 (0.68)
VCA	0.06 (0.22)
Iso. Jaggedness	$-0.06\ (0.19)$
Potential CoVis.	$-0.22\ (0.20)$
Total Time Inside	$0.66\ (0.31)^{*}$
Salience Rating	0.16 (0.17)
AIC	471.14
BIC	526.60
Log Likelihood	-221.57
Num. obs.	388
Num. groups: loc	28
Num. groups: ids	28
Num. groups: pic	14
Variance: loc.(Intercept)	0.45
Variance: ids.(Intercept)	1.99
Variance: pic.(Intercept)	0.15
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 15.12.: Model 15.16.

To further explore the causes of this misalignment, Model 15.16 was constructed, containing the eye-tracking data available for the subset of all participants:

```
BttW ~ Quantity of ET + Time to First Fix. +
Picture-Switching + Long-Short Dwell Ratio + cond + VCA+
Iso. Jaggedness + Potential CoVis. + Total Time Inside+
Salience Rating + (1|ids) + (1|pic) + (1|loc)
```

Model 15.16: Back-to-the-Wall Ratio with respect to Eye Movement

Achieving Marginal R^2 of 0.12 and Conditional R^2 of 0.51 it did not, however, improve our understanding of the dataset. Under the presence of larger number of factors, *Condition* also lost its statistical significance (see Table 15.12).

Looking back at the Model 15.15, we can notice that similarly to Recognition Memory analysis, the by-participant variance is very large. For this reason, aggregating the data by individual visitors' mean performance might provide additional insight into the above models. So transformed dataset allows us to investigate the amount of variance present between individuals in each condition separately. In other words, it can demonstrate how much each visitor differed from other visitors *within their condition*. Visual analysis of boxplots from Figure 15.13 alone leads to a conclusion that the differences in Spatial Memory performance between visitors from Condition 2 were much larger than those from Condition 1.

Primacy and recency effects in Spatial Memory were also investigated. As can be seen on Figure 15.14 visualising mean frequencies of correct *Back-to-the-Wall* responses against *ET sequence*, the picture seen as first clearly stands out, with a visible decline in Spatial Memory performance for the last pictures. A contrast chi-squared test (the first artwork in *ET Sequence* against all other artworks) showed that difference was statistically significant ($\chi^2(1) = 4.33$, $p = .037$).

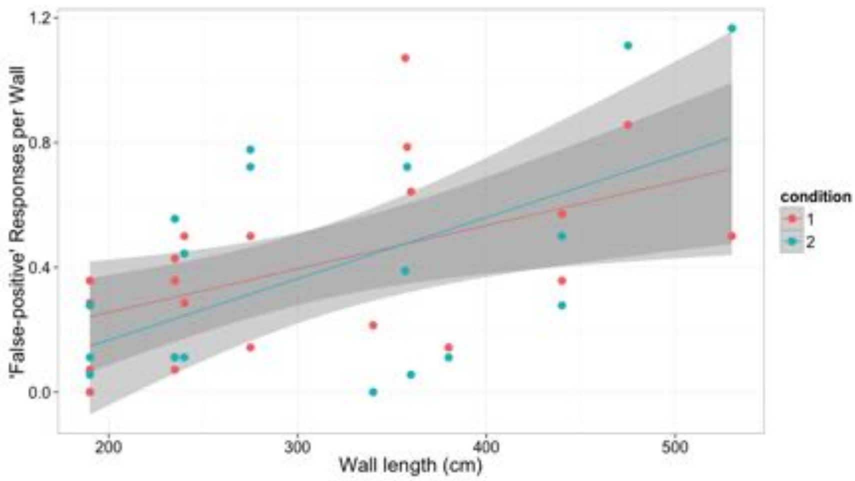


Figure 15.12.: Wall length plotted against the number of 'false positive' responses per wall.

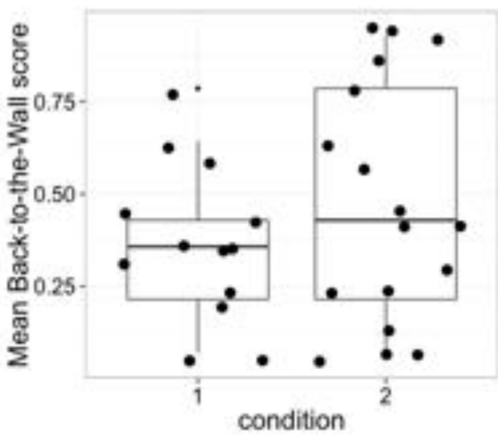


Figure 15.13.: Average Back-to-the-Wall score in two experimental conditions.

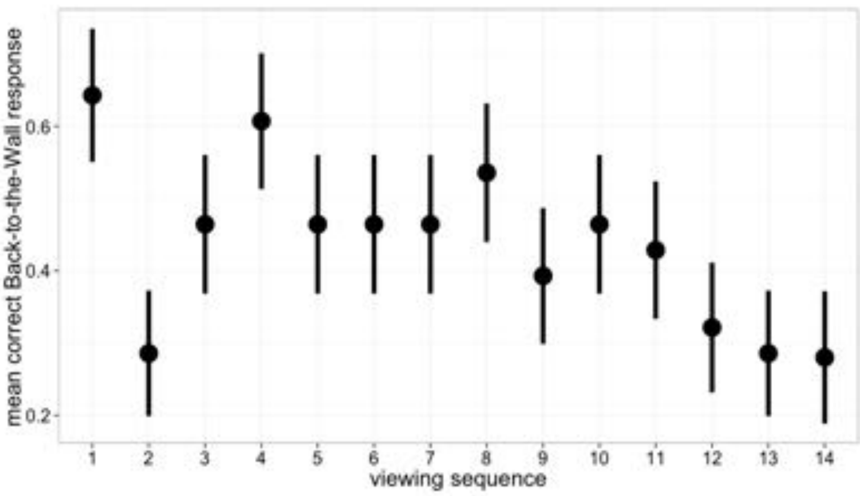


Figure 15.14.: Mean frequencies of correct Back-to-the-Wall responses plotted against the viewing sequence (whiskers indicate standard error).

16

DISCUSSION

This section will discuss the results of Experiment 1 in relation to the hypotheses stated previously in Section 12. The study here described had an exploratory character and its analysis will form a basis for further work in this thesis.

16.1 VISUAL ATTENTION

The distribution of dwell lengths was heavy-tailed, as it was predicted by Hypothesis (A1). A steep slope of the distribution's density (Figure 15.2, p. 170) can be observed at around 2-second length, reflecting the theory-based division made earlier in this thesis between *short* and *long* dwells potentially contributing to qualitatively different elements of the art perception process. *Long-to-Short Dwell ratio* reaching the value of 0.33 demonstrates that for each *long dwell* performed, visitors performed about two *short dwells*. High proportion of *short dwells* can indicate that—during the exploration of an unknown environment—those quick glimpses aid the decision making process about the further visual engagement and that this process remains under constant revision. Such interpretation would follow the suggestions made by J. K. Smith and Smith (2001), Bitgood (2010), Locher et al. (2007) in regard to the decision-making component of the visual interaction with art.

Considering the cumulative lengths of all engagements, it has been noted that *Total Dwell Time* per picture ranged from 24 to 40 seconds. With the mean of 31.71 sec. it therefore was not significantly different from the 30 seconds-mark previously found in other studies (J. K. Smith & Smith, 2001; Locher et al., 2007) and described by Hypothesis (A2).

The proportion of time visitors spent looking at pictures and the time they spent looking off-pictures was significantly different from the proportions which could be expected based on the artworks' share in total length and visibility of all walls. This, consistently with Hypothesis (B1), demonstrates that visual attention of participants from both experimental conditions was modified by the presence of visual targets in that space. It thus can be assumed to be task-specific in an environment where looking at pictures constituted a natural, 'default' cognitive goal of the entire visit.

Quantitive aspects of the oculomotor behaviour, i.e. *Number of Dwells* and *Total Dwell Times* were predicted by two models:

```
log(No. of Dwells) ~ VCA + View. Sequence + Total Time Inside +
(1|ids) + (1|loc)
```

and

```
log(Total Dwell Time) ~ VCA + Iso. Jaggedness + Total Time Inside +
 (1|ids)
```

All effects except *ET Sequence* were positive, meaning that the increase in *VCA*, *Isovist Jaggedness*, and *Time Spent Inside* increased *Number of Dwells* and/or *Total Dwell Times*.

Hypothesis (B2) stated that the size of *VCA* will have a positive impact on the *Number of Dwells* and *Total Dwell Times*. In both cases, this statement has been confirmed. At least two alternative explanations can be offered: (a) that this influence is due to 'top-down' factors such as perceived prominence of those locations with larger *VCA* attracting more attention from the viewer; and (b) that it is due to 'bottom-up' factors, such as increased probability of fixating on more visible objects during a random visual exploration of space[1]. However, under the lack of predictive advantage of the *Isovist Area* factor, the former explanation (a) becomes more justified. Since visitors explored the gallery through the variety of individual trajectories, assuming fixations on the basis of visibility alone would yield a higher suitability of the entire *Isovist Area* measure over its restricted 60 degree visibility cone[2]. This variable incorporates a much wider angular spectrum, thus better reflecting the visibility potential of individual locations. Knowing that the visitors' eye movement was guided by *Visibility Catchment Areas* instead, suggests that viewers either associated such locations with higher importance, or actively preferred straight-on interactions while minimising the oblique ones. In either case, this fact demonstrates that the oculomotor behaviour of the gallery visitors has been guided by the 'top-down' cognitive goal of a comfortable visual engagement with an artwork.

1 As we have already shown, participants' visual attention was not fully random and was attracted to pictures much more than it could be expected purely from their width and visibility compared to the empty stretches of wall surfaces. Here, we consider a case in which we assume that visual attention is random *given* that artworks already constitute a legitimate 'visual target'. The randomness here assumed would lie in the choice of one visual target over another, not in the choice of a wall segment with and without a visual target.

2 This is due to the fact that random fixations are similarly likely to fall on objects at very oblique angles, as they are to happen straight-on since *visibility* of a stimulus would be the only criterion considered.

On the contrary, increased *Isovist Jaggedness* (also associated with prolonged cumulative viewing times) represents a less-ideal viewing conditions by offering a potentially non-stable viewing experience for each viewer who walks through space. An explanation of why this effect is significant might be the fact that longer cumulative viewing times are required for a picture to be processed to a satisfactory level (e.g. until an aesthetic judgement can be formed Locher et al., 2007) under such non-stable viewing conditions. This suggestion, however, bears a dissimilar theoretical assumptions underlying the oculomotor behaviour happening inside art galleries compared to the ones presented in the previous paragraph in regard to the influence of *VCAs*. One argument links longer and more often viewing with a consciously made pre-selection (e.g. based on the perception of presumed importance of an artwork), while the other associates it with increased cognitive challenge. The data here analysed does not offer a satisfactory answer to the question of which one is true. From the theoretical viewpoint it seems entirely possible, however, that two—somewhat conflicting factors—have presence in an art gallery context. This raises a question in relation to memory measures further investigated in this thesis: if longer *Total Dwell Times* are the result of higher interest, they will most likely result in deeper cognitive processing and contribute to better *Recognition Memory* for the particular artworks. If longer *Total Dwell Times* are the result of difficulty in cognitive processing when an image is being analysed to a satisfactory level, no change in *Recognition Memory* would be predicted based on this variable. We will refer to this problem as Hypothesis (C8) later in the Discussion.

The negative influence of viewing sequence demonstrates that artworks which were seen later attracted less dwells. And yet, no such effect on cumulative viewing times was observed, suggesting that despite the fact participants might have devoted less glimpses to artworks located further away on their path, these less numerous interactions became longer, resulting in similar cumulative fixation times. This finding might add an interesting perspective to the issue of 'museum fatigue effect' (Bitgood, 2009), although remains beyond the scope of the current work, as it is unclear how exhibition layout might affect viewing sequences. The only relevant data here available is a generic measure of *viewing sequence similarity*. This analysis has shown significantly lower *MNLD* values in Cond. 2 what indicates that viewing sequences were more similar between individual participants in Condition 2 than they were between participants in Condition 1. Such result is in line with Hypothesis (A4), since the difference in spatial arrangement of artworks[3] unified the sequences through which visitors explored the gallery in Condition 2. One reason for such an effect might be the fact that artworks under this arrange-

3 The content of artworks and the wall layout being identical across the conditions.

ment were grouped in a clearer manner. While a person exploring any room in Cond. 1 had the possibility to turn the head into any desired direction in order to follow one engagement with another, visitors from Cond. 2 were much more likely to engage with those artworks, which were directly adjacent to the one they were already engaged with. Thus, a 'local' emergence of 'exhibition visit scripts' was much more probable. From the data available, little can be understood, however, about how global trajectories differ in respect to the likelihood of entering one room prior to another.

Both models are under a heavy influence of *Time Spent Inside* variable. In both cases, this has been the most impactful factor, and due to the nature of this statistical technique, it would be misleading to analyse the influence of space without considering so highly predictive measure in the model. The dominance of this predictor is especially visible in the model describing *Total Dwell Time*, which is unsurprising, given that being inside the gallery visitors rarely engage in activities alternative to viewing artworks.

Initially however, there was no evidence in favour of Hypothesis (D3), which stated that the oculomotor behaviour of those participants, who stayed inside the gallery for longer will be affected by space to a larger extent. Such an effect is difficult to quantify, as *Time Spent Inside* already is the strongest predictor of the quantity of eye gaze deployed to separate artworks. However, considering the measures of *Normalised Number of Dwells* and *Normalised Total Dwell Time* made it possible to control for the impact of time. The above two variables describe relative proportions of eye movement falling onto each individual picture, compared to all other pictures during a single visit. Predicting relative proportion of dwells seems to have little applicability in practice, as the final subjective experience of every visitor is heavily dependent on the time spent inside an exhibition. Therefore, the reader should bear in mind this statistical transformation has been conducted only to test Hypothesis (D3). It remains questionable, how measuring the influence of space on the proportion of attention allocated to separate objects can reliably guide design practices if other, more important factors remain ignored. Low R^2 values of the Model 15.10 further expose this issue. However, the models—to a limited extent—were able to confirm the assumption that space has a more dominant influence on the oculomotor behaviour of those visitors, who remained in the gallery for longer. Both, *Normalised Number of Dwells* and *Normalised Total Dwell Time* showed a significant interaction effect with factorial *Time Spent Inside* and one spatial measure (*VCA* for *Normalised Number of Dwells* and *Potential Co-Visibility* for *Normalised Total Dwell Time*). The influence of the above spatial factors will be more dominant on those participants, who stay inside for longer. As Wineman et al. (2006) suggest, this is due to the fact that visitors exploring the gallery for longer have more opportunities to become aware of spatio-visual relations between individual artworks (such as a more prominent visibility of some of them) and in result engage with them more. This is in an interesting point, as it suggests that more interested visitors, willing to spend more time on interacting with the art, remain under a heavier influence of the curatorial intervention, instead of building a more independent, individually unique experience. Prevailing popularity of physical art gallery visits in the age of the internet shows, that being *influenced* by these curatorial spatial interventions is a desired part of the visitor experience. As spatial layout is a legitimate curatorial mean for creating such experiences, its increased impact on more engaged visitors is of little surprise. On the contrary, *Number of Dwells* allocated to pictures seen later in the sequence decreased for those participants who explored it faster, but did not decrease for participants who allocated more time for the entire visit. Longer visits facilitated less dependence on the order of viewing, but more on visual characteristics of artwork locations. Again, no similar effect was found for *Total Dwell Times*, meaning that the decrease in the *Number of Dwells* for the works viewed later has no straightforward impact on cumulative viewing times.

It is a separate issue to what extent *Time Spent Inside* an art gallery can be modified by a curatorial intervention (such as altering spatial layout), and how much of its variation emerges from differences in individual exploration styles instead. Spatial interventions intending to affect average visit durations could for instance include changes in total floor area (however, there is some counter-evidence to that: Serrell, 1997; Yoshimura et al., 2012) or the configuration of artworks imposing more complex walking trajectories (Wineman & Peponis, 2010; Stavroulaki & Peponis, 2003). Individual exploration styles can, on the contrary, remain beyond the external control in a typical art visit situation. Informal observations of art enthusiasts suggest that differences in the agendas and preferences of their visitors might be a strong moderating factor. After all, an art gallery visit is a voluntary experience, being primarily the result of one's own wish to satisfy a personal cognitive need. The physical context of the exercise adds further components affecting the time spent inside it, such as physical tiredness, time constrains in one own's schedule, or the attitude of other members of the group. Perhaps, the limited impact a curator might have on this set of factors could be improved by making an effort towards diversifying the physical possibilities of exploration, e.g. by including various types of benches.

In accordance with the previously stated Hypothesis (A3), separate pictures and their *Salience Rating* did not introduce a significant variability into the quantity of visitors' oculomotor behaviour, meaning that the impact of individual picture's content was negligible. This is also visible on the boxplot Figure 14.3, p. 162. None of the by-picture random effect was significant, indicating that the variability in cumulative viewing times across pictures could not be explained by any other difference in their content. However, the influence of random by-subject variability was significant. These combined facts suggest that even if personal preferences did play a role in how long participants looked at individual images, they were so dissimilar, that cannot be generalised on a by-picture basis. It should here be noted that those findings can be the result of the used selection consisting of relatively similar stimuli and of the specific style of hanging employed in both conditions. By-picture variability could, for instance, happen to be significant under more radical hanging arrangements such as displaying multiple pictures hanging within a closer proximity to each other.

The fact that *experimental condition* was not a significant factor predicting *Number of Dwells* or *Total Dwell Time* of art gallery visitors suggests that the spatial fixed-effect predictors used in the models accounted for the sufficient amount of spatial variation introduced across conditions. This demonstrates, that there was no significant influence of spatial factors other than the one described by *VCA*, *Isovist Jaggedness*, and *Potential Co-Visibility* variables. Alternative variables which could be considered, such as number of pictures per wall or metric distance between pictures, either had no significant effect under the employed experimental conditions, or have already been sufficiently accounted for by the used measures. Concurrently, one of the two above models includes a significant random by-location effect, meaning that a considerable part of the variance across separate locations remains unexplained even after considering the fixed-effect spatial predictors. Even then this by-location variability does not seem, however, to be systematically divided across the experimental conditions. The impact of the imposed experimental manipulation on visitors' oculomotor behaviour perhaps did not reach statistical significance because is was not radical enough. Experimental manipulation increased the spread of spatial variables describing the positioning of artworks, but did not have much other influence. The fact that 3 spatial variables are able to sufficiently describe the entire significant spatial variation between 2 distinct exhibition arrangements poses two comments: (1) that the selected spatial measures are very well fit for purpose, at least when visual attention inside art galleries is considered; and (2) that the impact of space on oculomotor behaviour is restricted when the configuration of physical walls remains unchanged. Floor area, entrance and exit direction, as well as the set of 'all possible' trajectories are identical no matter what hanging arrangement is employed in such a study. The latter point will be one of the main reasons for modifying physical wall configuration in Experiment 2, which we will review further in the thesis. The former suggests, that those spatial effects which remain meaningful even in the presence of large individual differences can be controlled by a curator with the use of a limited number of easily quantifiable variables. At least in an art gallery as simple as the one used in the Experiment 1 and for non radically unusual hanging arrangements.

High *Conditional* R^2 values suggest that about 75% of all variance in the visitors' *Number of Dwells* and *Total Dwell Time* per individual object can be predicted based on the individual differences in viewing style, and spatial arrangement of the paintings alone. *Marginal* R^2 values demonstrate that about 52% of all variance in *Number of Dwells* and about 62% in *Total Dwell Time* has been explained by the fixed-effect factors which bear the potential to inform future design of the yet-unbuilt exhibitions.

Two additional models explored the possibility of predicting the 'dynamics' of visitors' oculomotor behaviour. The following formulae described the best-fit predictors:

```
Picture-Switching ~ VCA + Potential CoVis. + (1|ids) + (1|loc)
```

and

```
Long-Short Dwell Ratio ~ cond(n.s) + VCA +
Iso. Jaggedness + Potential CoVis. + Salience Rating +
gender + (1 | ids) + (1 | loc) + cond*Salience Rating
```

With *Marginal* R^2 values at 0.21 and 0.25 respectively, the dynamics of the visitors' eye movement was much less predictable based on the factors of interest. It did not depend on *Time Spent Inside*, as both outcome variables here considered were ratio measures calculated relative to all interactions conducted by individual participants.

Interestingly, despite the fact that *Potential Co-Visibility* had no significant effect on either of the measures describing the quantitive aspect of the oculomotor behaviour (Models 15.3 and 15.4), it did affect the dynamics of viewing. It is not *where* people allocate their attention, but *how* they utilise it, which remains under the influence of the number of potentially distracting objects in the visitor's surrounding. It is an important aspect to consider for future experimental aesthetic and museum visitor studies. Primarily, because it exposes how major a limitation are all observational techniques insensitive to the more subtle patterns of visual attention. This relates to all studies where visual engagement with art is approximated based on third-person observation (typically from behind the visitor's back), or on the spatial trajectory of the individual alone. It remains to be seen in this thesis, whether the change in the dynamics of the oculomotor behaviour also affects what visitors 'get out' (or remember) of an art exhibition.

Hypothesis (B4) which stated that only a 'complete isolation' of an artwork will result in lengthening its viewing times could not be tested precisely, as all artworks from Condition 1 and only two paintings (closest to the entrance and exit) in Condition 2 were isolated 'completely', possibly confounding such detailed interpretation. However, there was no linear influence of *Potential Co-Visibility* on *Number of Dwells* or *Total Dwell Time*. *Potential Co-Visibility*, did however, had a negative influence on the dynamics of visual attention, decreasing the proportion of long dwells and increasing the *Picture-Switching ratio*. Visitors performed lower proportion of *long dwells*, and were more likely to quickly 'switch' between alternative artworks the more of them were *co-visible* with the given painting. According to the principles stated in Hypotheses (B5) and (B6), *Co-Visibility* seems to have caused 'spotty-focused' attention, as the stimuli drew the engagement from each other. The limitation of the employed statistical approach and the limited variance of *Potential Co-Visibility* measure does not allow us, however, to investigate if this influence is *truly* linear. Larger-scale studies should consider the possibility that increasing *co-visibility* ceases to cause any more *picture-switches* and *short dwells* at some level, as it would become uncomfortable and impractical to react yet more dynamically to an infinitely increasing number of stimuli in the nearby environment.

According to Model 15.5, *Picture-Switching Ratio* indicating how often a dwell was followed by a quick, immediate fixation on another object, was higher for locations with larger *VCA* and higher *Potential Co-Visibility*. Larger *VCA* can be linked to a higher opportunity for 'catching' that object while switching between multiple pictures (i.e. the target of a potential 'switch' is automatically more likely to be the image from which one is 'switching on' to something else). The impact of higher *Potential Co-Visibility* is unsurprising, as this variable was designed as a measure of potential distraction one might encounter while engaging with separate artworks. Having a larger number of objects in one's periphery vision thus increases the chance of changing the object of attention more directly. When an artwork has no other artworks 'co-visible', such a direct 'switch' is simply impossible - the viewer needs to look at other fragments of space while traveling with gaze to the next stimulus. What is interesting is the linear characteristic of this relationship. It appears that the presence of additional objects does not simply provide the opportunity to quickly switch between them, but most likely stimulates this type of behaviour. An extreme example of such a fact is to imagine an art gallery populated very densely with pictures, almost adjacent to one another. It can be predicted that despite large floor, ceiling, and wall surface areas available above and below the line of displayed artworks, potential visitors would spend close-to-none fixations on anything else than pictures and their *Picture-Switching ratio* would be very high. Whether the recalled experience of such a visit would be different from walking between individual rooms with one painting in each space only, proceeded and followed by completely empty wall surfaces is still an open question.

Both *Picture-Switching ratio* and *Long-to-Short Dwell ratio* models include a significant random effect of picture location. This suggests that not all variability in the dynamics of visitors' eye movement can be explained using the specified spatial predictors (namely *VCA*, *Potential Co-Visibility*, and *Isovist Jaggedness*) and that there are other significant factors playing an important role, although they don't seem to differentiate the two experimental conditions in a systematic manner. This might include factors related to the metric distance between individual artworks and their orientation angles relative to each other. More theoretical work is required in order to build measures which would sufficiently describe the aspects of the environment relevant for human visual attention, while avoiding collinearity of such measures. For example, *Picture-Switching* was higher in Cond. 2, where pictures were located closer to each other, and the possibility of viewing many of them without moving one's head was much higher. A measure quantifying this feature would however depend on the current location of the viewer and be path-dependent. As stated previously, the precision of path-independent measures in predicting human viewing fields must be limited and it is important to highlight that measures which are easily generalisable for multiple exploration styles have a lower predictive power. *Potential Co-Visibility* used instead to encompass similar aspect is an example of such limitation.

Interestingly, despite being more likely to change between pictures quickly, visitors from Condition 2 executed more similar viewing sequences than those in Condition 1. This supports the earlier stated explanation that higher similarity of viewing sequences is most likely the result of viewing pictures in local groupings. Where multiple artworks are present in front of the viewer, the likelihood of comparing them through 'picture-switching' is higher, but the resulting dwells form clusters which distinguish it from the viewing mode of one-picture-per-wall situation arranged in Cond. 1. Further work focused on discovering the exact factors which facilitate 'picture-switching' could contribute to the designer's ability to plan and predict the 'local' occurrence of 'exhibition visit scripts'.

Long-to-Short Dwell ratio was affected by the highest number of significant predictors from all models considered above. Interestingly, it was the only measure where gender had a significant impact on the visual attention of art gallery visitors, with the behaviour of males generally consisting of lower proportion of dwells longer than 2 seconds[4]. Locations with larger *VCAs* generated lower proportion of *long dwells*, which—together with a significant negative influence

4 Gender differences in visual attention are not the focus of this thesis. They also tend to be limited to a certain range of age spectrum (Miyahira, Morita, Yamaguchi, Nonaka & Maeda, 2000).

of *Potential Co-Visibility*—can be explained by the higher chance of being the 'distracting' object while participants remained engaged with other artworks (and consequently one visitors get distracted 'from'). Also, the more 'jagged' the isovist was, the lower the proportion of longer dwells, which is in line with the theoretical underpinning of this concept. A more jagged visual field results in a less stable, intermittent visual experience consisting of shorter interactions[5]. Lastly, pictures ranked higher on *Salience Rating* generated higher proportion of long dwells, but the effect was only present in Cond. 1. This can suggest, that the impact of individual picture content only affected the dynamics of the oculomotor behaviour when viewing conditions across the entire gallery were arranged in a manner emphasising the distinct character of the stimuli by isolating them in a one-per-wall fashion. This suggests that the power of 'complete' isolation Robinson (1928) wrote about has a much more subtle contribution to the final gallery experience and affects the dynamics, but not the quantity, of the attention deployed.

In none of the four models described above, including *Isovist Area* instead of *Visibility Catchment Area* as a predictor improved the overall performance of the model. This fact can suggest that despite covering a smaller proportion of space, *VCA* bears a predictive power equal to, or better from, considering the entire *Isovist Area* generated from a particular artwork's location. Since considering the fraction of isovists lying outside *VCAs* did not result in an increase in fit of the considered models, it can be concluded that *VCA* is a measure as suitable, or more suitable from the *Isovist Area*, while remaining more theoretically plausible, and correlated less with other spatial predictors considered in this thesis. While the exact *VCA*, or *true* 'area of comfortable viewing' can vary depending on the context and size of the stimuli (Xie et al., 2007), the main theoretical principle stating that the area located directly in front of a picture bears higher significance for human cognition during the unrestricted exploration of space, seems to find confirmation in the empirical data.

As stated in Hypothesis (B7) eye-tracking data did not confirm the ideas of Bitgood (2010), who suggested that the influence of spatial factors might be less important for longer, more diligent interaction with artworks. A model predicting *Number of Dwells* altered to contain the interacting factor of *Long Dwells Level* has not shown a significant interaction effect of *VCA* and *Long Dwells Level*. A dataset aggregating all visual interactions between individual participant and each separate artwork, however, is not well-suited for such an

5 Note this does not only apply to extreme examples of isovists 'jagged' due to the presence of columns or other free-standing obstacles. An isovist 'jagged' because of encountering multiple open room corners is also more likely to be viewed intermittently than the one which does not. Vistas expanding across multiple rooms are visual tools very commonly used in many world-renown museum institutions (Tzortzi, 2007).

analysis. Inquiring whether a single diligent interaction has or has not been affected by the environmental factors could potentially be proposed within a dataset containing exact walking trajectories and viewing directions. More robust confirmation of Hypothesis (B7) lies therefore beyond the technical possibilities of this thesis.

For similar reasons, Hypothesis (B8) stating that multiple pictures adjacent to each other might be perceived as a holistic creative entity and result in lower proportion of *long dwells* per picture could not be verified with the data available. The result showing that *Number of Pictures on a Single Wall* decreased the overall fit of the model compared to the standard *Potential Co-Visibility* measure speaks against more specific Hypothesis (B8). However, a more detailed dataset involving quantified viewing directions joint with exact walking trajectories could potentially feed a more sophisticated statistical model[6].

The spatial variation introduced across the experimental conditions had a limited impact on the visual attention, most probably because of the same wall configuration. These was one of the main reasons inspiring the design of Experiment 2. However, before reviewing the next study, let us first discuss the memory results of Experiment 1.

16.2 recognition memory

In general, Recognition Memory performance of individual participants for separate pictures could not be well explained with the gathered data. In relation to the research question of this thesis and Hypotheses (C4), (C5), and (C6), no direct linkage between Recognition Memory and spatio-visual factors describing the location of pictures inside the gallery were found. Little by-location random-effect variance in the described models demonstrates that there was no noticeable consistent variability across those locations in the Recognition Memory performance. Some participants performed better then others and in some cases separate pictures triggered varying memory performance, but the location of the individual artworks seems to play a limited role in this aspect of human cognition. However, a significant cross-conditional effect on *Reaction Times* demonstrates that the artwork locations can *cumulatively* modify the memory performance of gallery visitors. Nevertheless, since more than one spatio-visual factor differed across the conditions (Section 14.1), the interpretability of this result is limited. It can potentially be argued that the cause of this difference is a somewhat less ‘optimal’ setting of Condition 2 consisting of smaller *VCAs* and larger *Co-Visibility* in relation to Cond. 1. As it is known from the previous section, Condition 2 caused more ‘spotty-focused’ attention by decreasing the proportion

6 A theoretical concept of ‘sight vectors’ proposed by Müller-Feldmeth et al. (2014) might offer such a possibility in the future and it will be reviewed in more detail further in the thesis.

of long dwells and increasing the *Picture-Switching ratio*. According to Hypothesis (C1), the resulting decrease in Recognition Memory performance is in line with laboratory-based memory studies of divided attention. It must be noted however, that real-life setting is very different from a computer screen and any effect of *Potential Co-Visibility* on human behaviour and cognition in 3-dimensional space might come from a much broader set of factors.

Hypotheses (C4—6) considered the relation between isovist characteristics and Recognition Memory. Even though—as it has been noted above—no direct linkage has been shown statistically significant in the data, the effect can be analysed *cumulatively*. As we known from Section 14.1, the distribution of *isovist area* and *isovist jaggedness* variables across the conditions was similar. Therefore, only Hypothesis (C6) can be verified, to a limited extent suggesting that larger *VCAs* can contribute to reducing the cognitive load associated with cognitive processing of individual artworks.

Total Dwell Time was a significant predictor of Recognition Memory operationalised by *RT accuracy* variable. This confirms the general assumption of the linkage between eye movement and Recognition Memory expressed in Hypothesis (D4). As the majority of participants' responses were 'correct' (78%), the influence of oculomotor behaviour should be rather interpreted as *inhibiting* in a limited number of circumstances, than linearly predictive of better Recognition Memory. In other words, very low *Total Dwell Time* decreased the chance for successful memorisation while the memorisation was nevertheless successful in the majority of cases. It thus can be concluded that the influence of space on Recognition Memory is *indirect* and manifests itself through the space's impact on the oculomotor behaviour proceeding memorisation. In order to interpret this significant effect of *Condition* on *RT* measure, we can refer back to the spatio-visual factors which have previously been shown to modify *Total Dwell Times* on a location-by-location basis. These are *VCA* and *Isovist Jaggedness* (see Model 15.4 in Section 15.2). Since the effect of *Isovist Jaggedness* was smaller, and artwork locations from Conditions 1 and 2 did not differ substantially in this measure, but did in their *VCAs*, we can conclude that Recognition Memory has been likely affected by cumulatively smaller *VCAs* in Condition 2, as a consequence of the fact that they decreased *Total Dwell Times* for individual artworks.

According to Hypothesis (C8) additionally stated in Section 16.1, *Total Dwell Time* was linked with better Recognition Memory, at least on the aggregate level. As we have previously discussed, this fact might clarify the ambiguous relation between the quantity of visual attention deployed to individual artworks and the underlying reasons for such behaviour. The fact that longer *Total Dwell Times* were correlated with better *RTs* suggests that longer viewing deployed to an artwork might be connected with deeper cognitive processing of these

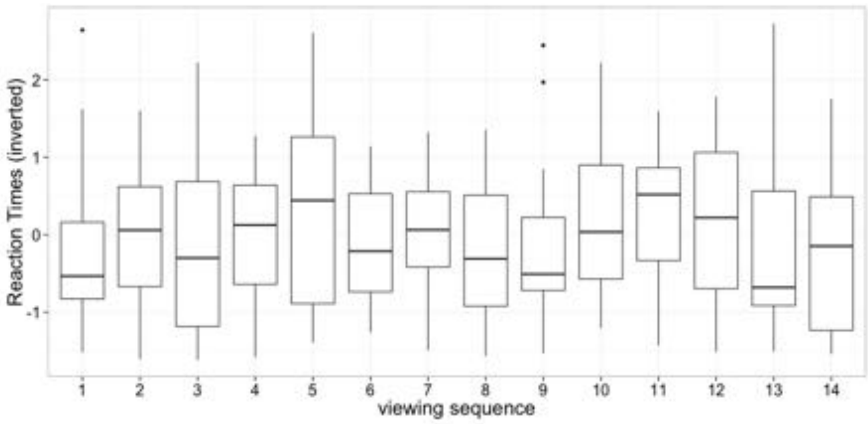


Figure 16.1.: Reaction Times (inverted - the higher the faster) and viewing sequence plot demonstrating the lack of primacy effect on Recognition Memory.

works, and not—what was the alternative hypothesis—with higher difficulty in processing the picture to a satisfactory extent. Such interpretation is in line with Hollingworth (2012) who argues that different tasks trigger different patterns of oculomotor behaviour, but once attention is deployed to a particular object, memory benefits from this allocation with little consideration to the initial task. In line with this argument, we could consider spatio-visual characteristics as external factors guiding visual attention similarly as task specificity is described by Hollingworth (2012). Those spatio-visual characteristics, in a similar manner to the task, bear the potential to influence eye movement patterns. It is thus not surprising that their impact on Recognition Memory happens in a similar, *indirect* manner - through the allocation of eye gaze.

To test for Hypothesis (C7) predicting better memory performance for first (primacy effect) and last (recency effect) object seen by each participant, *RT(inv)* was plotted against *ET Sequence*, indicating at which order the pictures were seen by each participant. Figure 16.1 shows these boxplots.

Boxplot figure revealed that there is no primacy or recency effect in the dataset and thus further statistical tests were omitted. There was no evidence to support Hypothesis (C7).

Significant effect of *Time Spent Inside* on *RT accuracy* and a stronger effect of *Condition* on *Reaction Times* for participants who stayed inside the gallery for 'longer' can be interpreted in line with Hypothesis (D3). The longer visitors spent exploring, the higher was the cumulative impact of space on their cognition. Participants who spent relatively little time inside can be seen as those performing 'the minimal required' engagement to fulfil their goal (which was to see all pictures inside the gallery). During this phase, space has a very limited influence, and it is only in the later phase of the exploration when the impact of small differences in spatio-visual characteristics of individual objects start to accumulate.

We can summarise this section by stating that the effect of space on human recognition memory is *indirect* (happens through the allocation of eye gaze), *cumulative* (is better observable in aggregate across two distinctly designed environments than on a location-by-location basis), and most likely *inhibiting* (rather than enhancing) memory in comparison to a theoretical 'neutral' situation.

16.3 SPATIAL MEMORY

Comparing Models 15.15 and 15.16 suggests an interesting conclusion that eye movement *is not* an important predictor of Spatial Memory on an object-by-object basis. *Time Spent Inside* and experimental *condition* had the strongest influence. The lack of influence of eye movement on Spatial Memory in the context here described can be further supported by referring to Section 16.1 which contains a detailed analysis of visual attention results. Under the presence of precisely defined spatial characteristics of individual locations, *Condition* was not found to be a significant predictor. The cross-conditional difference in Spatial Memory observed above lies therefore *not* in accumulated differences in visual attention patterns but in other factors, not accounted for in the current analysis. There are multiple reasons for which the spatial configuration of Condition 2 was easier to recall, one of which being that multiple co-visible artworks could serve as reference points in one's mental map of the environment. However, since the conditions contained different number of pictures per wall, this interpretation is limited by the potential bias in the measure. For instance, participants from Condition 2 answering randomly had a higher chance of putting a picture onto a correct wall, as long as they remembered which wall contained any pictures at all. On the contrary, a person not being able to recall which walls were occupied with pictures would face the risk of scoring multiple 'false hit' errors. This issue will be tackled in a further study, where number of pictures per wall will be kept constant across the conditions.

Insofar, the lack of relation between eye movement and Spatial Memory is contrary to Hypothesis (A5). It was based on the findings made by B. W. Tatler, Gilchrist and Land (2005), who found that memory for object's *position* accumulated with number of fixations in a study where participants were asked to view a real-world scene for 5 seconds. Tatler's explanation is that information about other object properties measured in the study (e.g. its *colour* and *presence*) can be gathered outside of fovea, while the information about its position is not. In light of this explanation, the effect might be much larger in an experiment allowing only very short viewing times, when the information gathered in general is scarce. In a free-walking scenario, where participants are almost unlimited in the viewing times they wish to devote to exploring a gallery other factors seem to play a role in facilitating their Spatial Memory and no linear effect of eye movement can be observed any more. Contrary to the previous expectations 'spotty-focused' attention happened to have no negative effect on the spatial aspect of this understanding.

The finding that Spatial Memory performance increased in Condition 2, where—as we saw in Section 16.1—visual attention was more 'spotty-focused' adds an interesting depth to the discussion of Hypothesis (B6). More interrupted viewing is generally believed to decrease the 'understanding of the exhibition' (Bitgood, 2010). On the cumulative level of the entire gallery visit, indeed such a viewing mode caused by the setting of Condition 2 decreased Recognition Memory (describing how well participants remember individual objects), but also increased Spatial Memory (potentially describing how well visitors remember spatial relations between objects). The 'understanding of the exhibition' has therefore at least two facets which do not necessarily increase or decrease simultaneously and linearly throughout the visit.

This confirms earlier studies of landmarks where these two memory systems were found to depend on different set of factors, as it was predicted in Hypothesis (C2), even though no direct linkage between any of these factors was clearly apparent. Contrary to the mentioned landmark studies, however, the experimental design of the study here described was not planned specifically for investigating this single effect.

In opposite to Hypothesis (C3) these factors were not observable on a location-by-location basis and thus it cannot be verified whether Spatial Memory for particular objects was affected by its perceived higher spatial importance. Low by-location variability suggests however, that this was not the case, and that the differences are only visible in an aggregate cross-conditional comparison.

Large variance in the results from Condition 2 (compared to Cond. 1) indicates, that significantly better performance observed in Cond. 2 is due to a group of visitors for whom this particular exhibition setup was easy to comprehend. Cond. 1 lacked such a subgroup, and the performance of all visitors was relatively similar. We can conclude by saying that the layout presented in Condition 2 can work very well for some, and very badly for others, when it comes to understanding spatial relations between separate pictures. Individual differences however seem to play a much smaller role in Condition 1, where 'an optimised' spatial design in fact resulted in relatively poor performance of all visitors.

In line with Hypothesis (C7) the position of the first picture seen inside the gallery was remembered best, although this effect might be due to the particular layout of the studied galleries, where no distractors were present at all in the corridor the participants explored as first. Further confirmation of this result in a more complex environment is required before its implication for Spatial Memory studies can be discussed.

Another result worth noting is that, as predicted by Hypothesis (D1), Spatial Memory performance decreased with age. Although this not surprising considering a large body of research showing that all spatial skills decrease as people get older (Kirasic, 2000).

Part V

EXPERIMENT 2

Experiment 1 had shown that spatial factors responsible for the 'quality' of the visual experience alone were able to create a difference in the subsequent memory performance of the gallery visitors. The factors we focused on were *Potential Co-Visibility* and *Visibility Catchment Areas*. Despite the fact that the total coverage of isovist areas was the same in both conditions of Experiment 1, and that the average size of the pictures' isovist was similar, this 'quantity' of visual stimulation did not provide equal conditions for engaging with and processing of artworks.

Building on the findings from exploratory Experiment 1, a more restricted study was designed to investigate the specific relationships between *Visibility Catchment Areas* and *Potential Co-Visibility* of picture locations. By arranging two experimental conditions where *VCAs* were similar, and *Potential Co-Visibility* highly differed, it was possible to disentangle their individual impact on the gallery experience.

17

METHOD

In Experiment 2, participants were invited to explore a non-public art gallery arranged in a psychological laboratory at the Open University in Milton Keynes, UK. The space is located in one of the newly built university buildings and it is purpose-built for experimental psychology studies. Despite being located in a public building the access to the laboratory area is restricted and a separate reception room for briefing and de-briefing participants is available. The study was conducted between the 22nd of January and 5th of February 2014.

17.1 PROCEDURE AND PARTICIPANTS

The procedure was identical to the one employed in Experiment 1 (Section 13.1) with a single exception. Informal conversations with the participants of Experiment 1 suggested that some of them assumed the longer time they spend inside the gallery the 'better' it is for the study. Thus, the instruction presented in Experiment 2 (Appendix D) was amended to put a stronger emphasis on the individual preference regarding the time spent inside. Again, due to the eye-tracker's battery capacity, 30 minute time limit was imposed. One participant did not comply to the time limit and was interrupted by the researcher and asked to leave after having spent 35 min inside the gallery. This person's data was not removed from the final dataset as eye-tracking recordings would still allow to assess the sequence and quality of his or her visual engagement with the artworks.

Forty-five participants were recruited through university email systems, regional internet discussion forums, and a local newspaper article for an advertised fee of £6. Four participants were interrupted by unusual events (lighting failure, or a picture falling off the wall) and were not considered in the data analysis. The remaining 41 participants (27 female, 14 male) were aged between 18 and 69 (M = 40.64, SD = 15.92).

Participants were randomly assigned to one of two experimental groups depending on the scheduled day of the meeting. Seventeen participants explored the gallery in experimental Condition 1, and 24 in Condition 2. Following the calibration of the eye-tracker, each

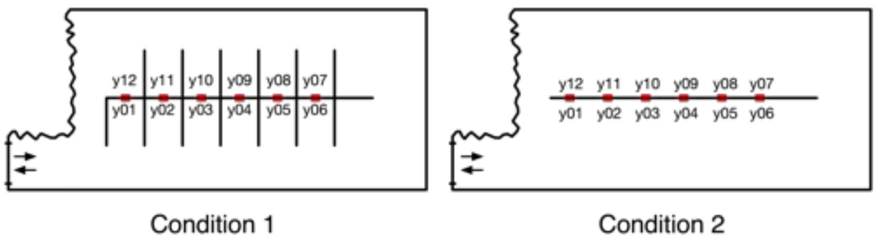


Figure 17.1.: Two layouts arranged in Experiment 2. Ragged line in top left corner represents a black curtain stretching from floor to ceiling, which covered laboratory's technical equipment behind it. At its longest and widest point the room is 17 x 7.6 m, and the wall partitions added in Condition 1 are 2m (vertical) and 1.6m (horizontal ones) each.

participant was brought in front of the gallery entrance and asked to use the same door as an entrance and an exit.

17.2 SPACE AND MATERIALS

Contrary to Experiment 1, spatial conditions of Experiment 2 differed by the shape of the gallery layout, but did not by the coordinates of the pictures' locations (Figure 17.1, 17.3, 17.2, and 17.4).

An empty wall segment was added at the end of the artwork sequence in both Conditions (following locations y06/y07) to prevent participants from turning very rapidly around the wall and spotting the first image on the other side from a very close distance. Such a situation would potentially distinguish these two edge locations from all others, as visual encounter was more likely to start from a larger distance in all other cases. This addition of an extra partition was not possible on the opposite end of the wall and will be considered during the analysis of the results.

Each experimental condition contained 12 artworks which were previously used in Experiment 1 (Bellamy, 2012). Compared to the previous study however, two artworks were not used: pictures 'G' and 'H'. The size of the image set had to be reduced in order to ensure equal spatial distribution of individual objects within the available length of wall segments. The selection of those two images was based on the previously established *Salience Rating* (Section 13.6): 'G' and 'H' were two objects which obtained *Salience Rating* closest to the mean of the entire dataset. Removing them created higher variation in the perceived salience of the used image set. This was desired

Figure 17.2.: Experimental set-up in Condition 1.

Figure 17.3.: Experimental set-up in Condition 2.

Figure 17.4.: Experimental set-up in Condition 1, a view on a single picture.

due to the placement of a large number of pictures on only two continuous wall surfaces. Under these spatial conditions there was a risk of cancelling out any potential influence of perceived salience if multiple highly salient pictures are placed next to each other. For this reason, random hanging sequences unique for each participant (similarly to Experiment 1) were created with the additional rule of interchangeability between highly- and lowly- salient pictures. *Salience Rating* median value was used to differentiate between highly- and lowly- salient pictures and each random hanging sequence contained no two artworks from the same group on the adjacent locations. For half of participants sequences started with a highly salient picture at location y01, and for another half with a lowly salient one.

Similarly to Experiment 1, no labels or textual information about the artworks were provided. Additionally, the experimental space contained a one-way mirror built into the wall in front of locations y01 and y02 which is typically used for conducting unobtrusive observations of the experimental procedure from the adjacent room. This room was not used during the Experiment 2.

Due to a non-uniform nature of the lighting conditions, the number of LUX units were measured for each location, separately in both experimental conditions. The measurements were performed with an analogue lighting sensor, from the centre of each picture, twice for each case (and the mean of those two measurements was taken).

17.3 eye-tracking recordings

'Tobii Glasses 1' eye-tracker was used in the experiment. The calibration was successful (and recorded above 60% of valid fixations) for 29 participants, whose recordings were extracted and coded according to the procedure described in Appendix A.1. The coding procedure considered instances of each picture dwell, as well as joint dwells on walls, floors and the ceiling. Additionally to Experiment 1, the dwells falling on the one-way mirror were also coded in order to control for the level of potential distraction introduced by this object. Any fixations on the artwork's reflection in the mirror were however coded equally to a dwell occurring on the artwork.

Those participants whose eye-tracking data could not be used were unaware of the equipment failure and therefore still performed all tasks in the same context as all other participants. For this reason, the data obtained from their memory tests was used in all analyses where it did not require matching with eye-tracking data.

17.4 recognition memory test (reaction times)

The Recognition Memory Test was presented via OpenSesame software (Mathôt et al., 2012) on a 13" laptop with two keys ('A' and 'L') labeled 'YES' and 'NO' by black-and-white printed stickers. The entire task was identical to the one used in Experiment 1, with the exception that pictures 'G' and 'H' were not included, and therefore the task consisted of 3 training stimuli, followed by a randomly ordered projection of 12 correct artworks and 12 new objects.

17.5 spatial memory test (miniature task)

The Miniature Task was similar to the one used in Experiment 1, but in order to record not only the final result, but also the sequencing of the answers, it was prepared on a computer and presented on the same laptop used for the Recognition Memory Test. An empty Apple Keynote slide was displayed in the fullscreen mode, without any toolbars visible. The layout of the gallery was displayed as a non-clickable, non-movable slide background. The layout contained the locations of all walls and of the entrance (in black), as well as the locations of all pictures (in red). Picture Miniatures were randomly scattered outside the layout, in equal numbers above and below it. Figure 17.5 presents a sample starting set up of the task, and Figure 17.6 its sample solution.

The Miniatures were manipulated by 'drag and drop' actions performed with a mouse. Participants were instructed to take as much time as they need to complete the task, but it was suggested that they place all pictures on their presumed location (Appendix D.4). The en-

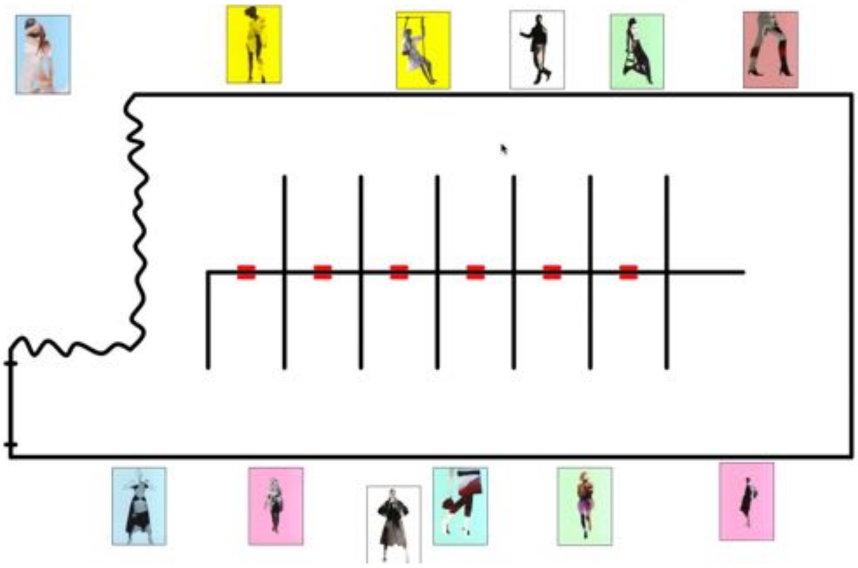


Figure 17.5.: Sample set-up at the beginning of the computerised Miniature Task.

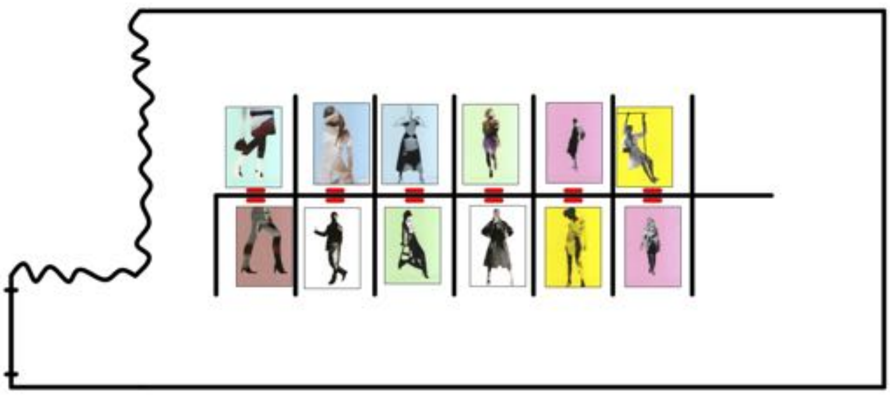


Figure 17.6.: Sample solution of the computerised Miniature Task.

tire solution was recorded with Mac OS X native screencast function and a screen shot of the final solution was taken for the purpose of a back-up.

All participants produced valid solution. If ambiguities in miniature placement were present, the researcher asked to specify which exact wall location the participant considers correct and to adjust the position of the miniature.

17.6 REVISED HYPOTHESES (E)

Experiment 2 differentiated the *Co-Visibility* of pictures across two separate experimental conditions, while keeping other influential spatial factors identical (either within- or across conditions). It therefore constituted a research setting well suited to test for hypotheses referring to the influence of *Co-Visibility*. Asterisks indicate key contributions proposed by this work.

- *(E1*)*: Considering the findings of exploratory Experiment 1, the quantity of visual attention deployed to individual pictures was expected to remain similar across the conditions, as *Potential Co-Visibility* had no significant effect on *Number of Dwells* or *Total Dwell Times* in Exp. 1. However, the dynamics of visitor oculomotor behaviour is expected to differ, resulting in 'spotty-focused' attention under the high co-visibility of Condition 2.

- *(E2)*: Simultaneously, there were factors which were expected not to vary significantly across the conditions. *Time Spent Inside* was strongly affecting the quantity of oculomotor behaviour in Experiment 1, and so it was expected to equally strongly affect the results in both conditions of Experiment 2.

- *(E3)*: It remains to be seen whether the impact of individual pictures and their *Salience Rating*, which has not been observed in Experiment 1, rises under more radical differences in hanging arrangement. The current study excluded two stimuli which had their *Salience Rating* closest to the mean value of the entire set. As mentioned previously, hanging sequences were also modified to interchangeably display highly and lowly salient paintings. It is expected, that under lower spatial variability within conditions, these circumstances will facilitate higher impact of visual salience, and result in significant random effects of individual pictures and a significant fixed effect of *Salience Rating*. The effect should be particularly visible in Condition 1, which facilitates isolated viewing of individual artworks. In Experiment 1, 'completely' isolated and highly salient artworks attracted larger proportion of *long dwells*. It might be expected that the dynamics of the oculomotor behaviour in Condition 1 of the current study will therefore also be linked to the *Salience Rating* of individual artworks.

- *(E4*)*: The importance of *VCA* in guiding visitors' eye movement as well as a significant cross-conditional difference in Recognition Memory performance in Experiment 1 (where *VCAs* were the major factor differing the conditions) suggests that *VCA* might indeed formally describe a more comfortable conditions for processing artworks. In the current study, where *VCAs* were kept nearly identical, it is thus expected that Recognition Memory performance will not differ between the conditions.

It is important to note that Condition 2, despite keeping *VCAs* similar to Cond. 1, also provided more diverse opportunities for viewing pictures from multiple angles and distances. This difference can be described by a larger *isovist area*, which did not significantly improve fits of the models explaining the results of Experiment 1, but remains a factor not to be ignored in Experiment 2. The interpretation suggested to explain the influence of *VCA* in Experiment 1 was that it guided participant's visual choice, possibly indicating importance of the given spatial location. This assumption cannot be tested in the current study, which contains little variation across separate locations within the conditions and thus is most likely to result in two different strategies of exploration very similar within each condition.

Higher *Potential Co-Visibility* of some locations in Experiment 1 was linked with 'spotty-focused' dynamics of visual attention, decreasing the proportion of long dwells and increasing *Picture-Switching ratio*. This effect, already considered by Hypotheses (B5) and (B6), is expected to prevail across both conditions of Experiment 2.

18

DATA ANALYSIS

18.1 CROSS-CONDITION DIFFERENCES

Following the findings from Experiment 1, the spatial conditions were prepared in order to provide a differentiation in *Potential Co-Visibility* while keeping the *Visibility Catchment Areas* as similar as possible across the conditions. Due to the length of the available wall segments, preparing Condition 1 with unrestricted 60°*VCAs* was not possible. The segments created between *VCAs* restrict them to 46°. Thus, the areas of comfortable viewing were identical across conditions up to 46°only. Figure 18.1 presents these relations visually. As the distance between each picture's surface and the external wall of the gallery is identical in all cases, the size of so defined *VCAs* is also equal (both within and across conditions). Despite a narrower angle range compared to the original suggestion Stavroulaki and Peponis (2003), the uniformity of 46°*VCAs* was expected to create similar conditions for 'comfortable' processing of pictures. Therefore no significant difference between Recognition Memory performance was predicted between the two experimental conditions.

The presence of the additional wall partitions in Condition 1 limited *co-visibility* so that there was no location in the entire space from which more than 2 pictures can be viewed at the same time. In fact,

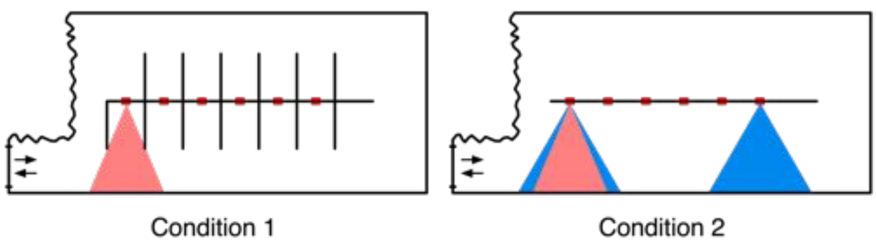


Figure 18.1.: Comparison of VCAs across the conditions (blue - VCA as defined in Exp. 1; red - VCA restricted in Cond. 1 of Exp. 2).

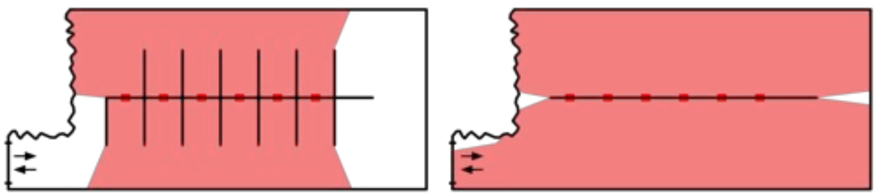


Figure 18.2.: Total area making it possible to engage with at least one object.

the majority of the entire gallery area allows its visitor to view only a single picture at a time, with the only exception being the 'transition zone' between two separate segments. The *potential co-visibility* parameter in this layout is equal to 1 for each picture. In can be therefore suggested that Condition 1 was 'optimised' for viewing pictures individually. On the contrary, a visitor exploring the gallery in Condition 2 is permanently exposed to multiple visual stimuli - the majority of the gallery area provides visual access to 6 pictures at a time. In this layout, the *potential co-visibility* measure is equal to 6 for each picture. If the impact of *potential co-visibility* comes indeed from facilitating less-focused, more distracted viewing, than the difference between those two experimental conditions exemplify this spatial aspect.

Additionally to *potential co-visibility* and *VCAs*, the variability of possible viewing angles and distances for each visitor-artwork interaction was either highly restricted (Condition 1), or broad and flexible (Condition 2). This is a factor which can be formally described by the average size of pictures' isovist areas. These differed substantially (Cond. 1: M = 350, SD = 57; Cond. 2: M = 2636, SD = 153[1]). Thus, while in Condition 1 almost all visual engagements with any artwork had to occur within its *VCAs* (i.e. in a narrowly defined range of angles and distances), the variability of any potential visual engagement in Condition 2 was much broader.

Moreover, the isovist coverage area, or the proportion of the total gallery space making it possible to engage with at least a single artwork was higher in Condition 2 (98%) than in Condition 1 (67%; see Figure 18.2).

The last two spatial factors (*mean isovist size* and *isovist coverage area*) contribute to the 'quantitative' aspect of the visual experience. Through them, two situation were created where participants were able to potentially engage with artworks for different proportions of

1 Arbitrary Depthmap units.

time. In Condition 1, the gallery's floor area is used 'less efficiently', leaving out 'blank spots' of unused space. In consequence, Condition 2 provided higher 'quantity' of the afforded visual engagement, while Condition 1 provided its lower 'quantity' but presumably a higher 'quality'. The aim of Experiment 2 is to investigate whether the 'quality' outweighs the 'quantity'.

18.2 WITHIN-CONDITION DIFFERENCES: SPACE SYNTAX AND ISOVIST ANALYSIS

Within-condition variability of spatio-visual measures was minimised as all locations within each condition had similar spatio-visual properties. For this reason, a detailed isovist analysis was not required. The only inconsistency existed in the *isovist size* of the edge location 'y12' in Condition 1. This however was unlikely to create a significant distinction, as observations of the eye-tracking recordings revealed that the room corner to which this isovist area stretched out typically remained unexplored by the visitors. Small differences in isovist area sizes of the bottom wall surface in Condition 2 must also me noted, as the subsequent locations were located at decreasing angular positioning to the gallery entrance area. The presence of eye-tracking recordings made it possible however to control for the sequence of visual engagement with individual pictures, which was more critical in a situation when the entrance area remained unexplored during periods other than the entrance and exit phases.

18.3 EYE-TRACKING MEASURES

Eye-tracking measures described in Section 8.2 were obtained from the coding logs, using a custom-built R script. Similarly to the results of both memory tests, oculomotor variables were associated with individual pictures and with individual locations.

The recordings were also used to calculate total *Time Spent Inside* the gallery by each participant (from the moment of the first dwell occurring inside the gallery to the moment of opening the exit door). *Time spent inside* the gallery was also manually tracked by the researcher in order to estimate it for those participants, whose eye-tracking recordings could not be analysed due to technical reasons.

18.4 RECOGNITION MEMORY TEST (REACTION TIMES)

For Reaction Time data analysis, only the correct 'yes' answers were taken into consideration (86% of the whole dataset). However, the accuracy was also recorded and will be referred to as *RT accuracy*. Observations lying below 250 ms or further than two standard deviations from the mean of the whole dataset were removed (Ratcliff, 1993; Whelan, 2008), leaving 80% of all responses valid (i.e. correct and within two Standard Deviations).

18.5 SPATIAL MEMORY TEST (MINIATURE TASK): STRING MATCHING ANALYSIS

When participants' Miniature Task performance is considered, *Bidimensional Regression* would not be a suitable tool for analysing this experiment's results. This is due to multiple pictures being located exactly on the opposite side of the same wall which is a problematic situation for this method (see Section 10.1). Also *Back-to-the-Wall* score is unreliable in the current situation - it would not provide the sufficient differentiation between the scores, as all 12 pictures were allocated on one of two wall surfaces and the chance of correctly placing any of them by luck would equal to 50%.

For this reason, String Matching algorithm previously described in Section 8.3 was used to score Miniature Task results. Custom-designed 'Adrien-Kuba Distance' makes it possible to assign weights for each action type. For the purpose of this analysis, the cost of a single *swap* was set at 0.2, and the cost of a *substitution* at 1.0. All participants made use of the red line cues as place-holders for miniatures, therefore all solutions had string lengths equal to the correct reference strings and no *insertion* or *deletion* actions were required. In result, each Miniature Task solution can be coded as a set of two independent strings, each corresponding to one of the two wall surfaces (upper or lower side on the layout), and averaged to obtain the final score named *Mean Adrien-Kuba Distance*. Under this operationalisation, a potential mistake of substituting two pictures lying on two furthest ends of the same wall surface (i.e. mistaking picture located at y01 for that from y06 and vice versa) requires 9 *swaps* at the cost of 0.2 to rectify. The resulting score of 1.8 (assuming all other picture locations being solved correctly) after averaging between 2 wall surfaces equals to 0.9. This result is still lower than the score which would be obtained from the simplest error made *across* wall surfaces (as this would require 2 separate *substitutions* for each wall, divided by 2 walls, giving the score of 1.0). Therefore, under these weights, the biggest error possible *within* a single wall surface is punished lower than making any error *across* the wall surfaces.

This operationalisation reflects hierarchical organisation of spatial memory with a vista space serving as a basic unit of distinction (see Section 3.2.2). However, while this might seem plausible in Condition 2, it is not certain that participants in Condition 1 used two sides of the wall to differentiate between distinct units of space. As there was no single point in space of Condition 1, from which the visitor would be able to comprehend the entire wall surface (but only 1 or 2 *VCAs* at a time), artworks might have been remembered as a list of objects, based on the sequence of encounters with their discrete subspaces. In such a case, there would be no reason to suggest any difference between a mistake made on a pair of pictures located at one side of the wall (e.g. y05-y06), and one made across two wall surfaces (e.g. y06-y07). To control for such a situation, a variable reflecting this fact was also calculated. Adrien-Kuba Distance algorithm was modified to allow single *swaps* between first and last character of the string, and *Wrapped Adrien-Kuba Distance* was created with the input provided as a continuous string of characters representing pictures from location y01 to y12. In *Wrapped Adrien-Kuba Distance* the use of *substitutions* is not theoretically justified, therefore only *swaps* were used. The biggest possible error in this variation would be represented by the misplacement of two pictures originally located on the furthest positions *within the string* (and not within a single wall), e.g. on locations y01 and y07. Correcting this error, assuming all other placements being accurate, would require 10 *swaps*. To make such an error the reference point for further interpretation, the weight of 0.1 was assigned to a single *swap* action, which would result in the value of 1.0 associated with the biggest single error possible. Other weights were set at 1.0 in the algorithm, so that it would always be more justifiable to perform *swaps* than *substitutions*.

However, it had to be decided which measure—*Mean Adrien-Kuba Distance*, or its *Wrapped* version—is more appropriate for the purpose of this analysis. Therefore it had to be established whether participants in both conditions encoded the content of the gallery in a hierarchical memory structure. This would result in each edge of the wall surface constituting a 'barrier' to their memory more significant than any single VCA partition in Condition 1. To control for such possibility, total number of errors irrespective of their string distance to the correct location was extracted from the Adrien-Kuba Distance script. The errors were classified based on whether they were performed *within* the correct wall surface, or *across* it. The number of errors for each wall was summarised per person. As a result, two new variables were created describing *sum of errors across the wall* and *sum of errors within the wall* made by each participant. The variables were then divided by each other to create *across-to-within error ratio*. If a given participant conducted equal number of errors *across* the wall and *within* the wall, the ratio would be equal to 1. This would reflect a situation where all incorrect answers were placed randomly on the layout (there was equal number of possible locations on upper and lower side of the wall, therefore the random chance of making an *across-wall* error was equal to that of making a *within-wall* error). Ratio lower than 1 would indicate that a participant conducted more *within-wall* errors, i.e. possibly remembered the correct side of the wall even if could not identify the exact location of the picture. Ratio higher than 1 would indicate that the participant has committed more *across-wall* errors - an unlikely situation, unless he or she wrongly mapped the spatial memories from the gallery to its representation on the Miniature Task.

If the two wall surfaces were indeed used as distinct units for storing spatial memories of its content, it can be expected that the proportion of errors made across the wall surfaces to those made within wall surfaces will be significantly lower than 1.0. This was indeed true both in Condition 1 (M = .46, SD = .34; One Sample t-test agains the value of 1.0: $t(16) = -6.7$, $p < .001$) and in Condition 2 (M = .5, SD = .52; Wilcoxon signed test agains 1.0: $V = 26$, $p < .001$) and the difference between the conditions was non-significant (W = 206.5, p = .957). This proves that Miniature Task participants were influenced by the presence of two wall surfaces in a similar way in both conditions and that *Mean Adrien-Kuba Distance* is a more appropriate measurement of their performance, equally adequate in both experimental conditions.

18.6 MINIATURE TASK SEQUENCE ANALYSIS

One common criticism of sketch map-type exercises is that their correct solution does not necessarily require the participants to construct an allocentric reference frame of the explored environment. Simply remembering the sequence of one's own movement and the objects encountered on the way (known as the 'route map') could serve to produce a sketch map. If this was true, Miniature Task would not be a reliable measure of Spatial Memory. This can be particularly the case in both experimental conditions arranged in Experiment 2, where the shape of the galleries encouraged their circular exploration. For this reason, 'route maps' potentially constructed by the visitors would be based on almost identical sequences of viewed artwork locations.

To account for such a possibility, screen recordings of participants' 'drag and drop' actions (performed while they were solving the Miniature Task) were coded into a spreadsheet. Each time a miniature was moved to one of the picture locations (indicated by thick red lines on the miniature layout), the order of this movement was recorded. All movement to unmarked positions was ignored. Each time a miniature was moved from an already taken position to a new one, its order in the sequence was changed to the last. For example, if a sample participant moved 4 pictures in the order [C,F,M,D] and then changed the decision and moved picture 'F' to a new location, the modified sequence was coded as [C,M,D,F]. This procedure of sequence coding also assumes the level of certainty with which each decision is being made - a modification of the already made choice is treated as a sign of lower level of certainty. The same coding scheme was used for all participants, even when inferring the actual level of certainty was more problematic (e.g. when participants moved a group of pictures into an empty space to reorganise them; in which case the coding procedure would ignore all movement until pictures were placed to one of the marked locations).

This data was then correlated with *ET sequence* variable to test whether there was a consistent relation between the order of *looking at* an artwork and the order of *moving* the artwork in the Miniature Task. The presence of such a relation would indicate that participants provided answers to the Miniature Task according to the 'route map' built during the gallery exploration. In such a case, the sequence of those answers should not be used as a reliable measure of certainty of the Miniature Task answers.

As both variables of interest were on a rank-order scale, Spearman's rho statistic was used. The analysis showed that there was no monotonic relation between *ET sequence* and *Miniature Task sequence* in either of the experimental conditions (Cond. 1: $\rho = 0.09$, p=.350; Cond. 2: $\rho = -0.02$, p=.782). However, a closer examination of separate frequencies revealed that there was some consistency in the order of filling the Miniature Task which followed the organisation of the space. Across the conditions, location y01 was solved as *first* in the Miniature Task in 24 cases (out of 41 participants for whom Miniature Task video recordings were available). Location y02 was filled as *second* in 10 cases, and location y03 as *third* in 11 cases. For over half participants, location lying the closest to the entrance was where they started to solve the Miniature Task. For about a quarter of visitors, the same was true for second and third location along the gallery wall. Figure 18.3 presents this relation visually. It indicates, that Miniature Task order extracted from screencast recordings can be potentially used as a measure of certainty.

Moreover, as pictures were scattered above and below the layout, there existed a risk that the default locations of the miniatures will be associated with the side of the wall they should belong to. The screencast recordings were analysed for the occurrence of an event in which participant 'drag and dropped' at least a single picture across the wall, relative to its default location. All participants analysed in the study have performed at least one such action, confirming that they did not associate the default miniature location with the correct side of the wall.

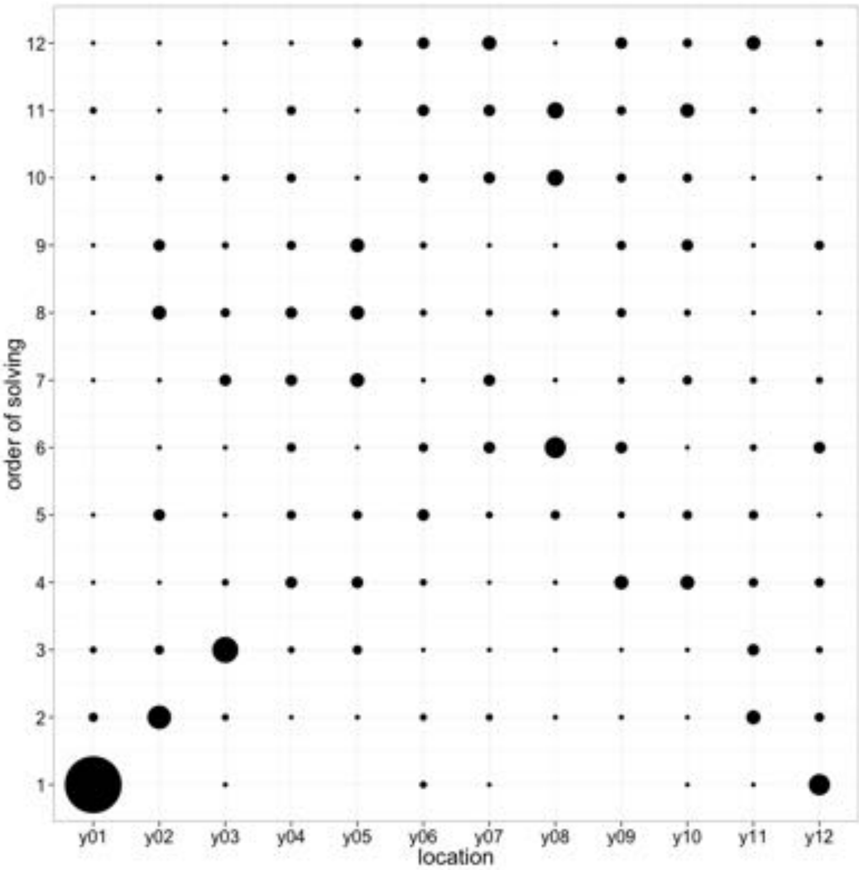


Figure 18.3.: Order of selection in the Miniature Task depending on the location. The size of each dot indicates how often a given location was filled in the Miniature Task in the particular order. It is visible, that locations y01, y02, and y03 were often solved in the same order (as first, second, and third respectively). The pattern was less consistent for other combinations. If all locations were solved in a linear order following their spatial order, larger dots would form a distinctive diagonal line across the graph.

19

RESULTS

19.1 TIME SPENT INSIDE

Participants spent M = 513 seconds (or about 8.5 minutes) inside the gallery (Cond. 1 M = 537.02 sec., SD = 216.65; Cond. 2: M = 495.98, SD = 453.35; W = 270, p=.083). Similarly to Experiment 1, the distribution of visit lengths shows a steep decrease at a specific point (around 10 minutes), which is below the imposed time limit. This indicates, that to the majority of participants, this period was sufficient to explore the gallery (Figure 19.1).

19.2 VISUAL ATTENTION

All participants fixated at least once on every picture. They performed 80.17 dwells on average during a single visit and the proportion of the total time inside spent looking at pictures (or average *Engagement Ratio*) was 0.84. It was significantly lower in Condition 1 (M = .78, SD = .1) than in Condition 2 (M = .87, SD = .06; $t(27)$ = 2.97, p = .006) indicating that participants in Cond. 1 spent, on average, 9% less of their time inside the gallery fixating on pictures. This is partially due to the entrance area of the gallery not allowing visitors to engage with any picture instantly, which resulted in higher average *Time to First Fixation* in Condition 1 (M = 6.88 sec., SD = 1.45) than in Condition 2 (M = 2.31, SD = 1.63; W = 182, p<.001). Additionally, as mentioned previously, the percentage of all floor surface area allowing the visitors to engage with at least one picture was higher in Cond. 1 (98%), than in Cond. 2 (67%; Figure 18.2, p. 230).

However, *Total Dwell Times* summarised in seconds were very similar across the conditions (Cond. 1: M = 407 sec., SD = 229.91; Cond. 2: M = 409.16., SD = 309.4; W = 102, p = .77). Consistently with the result presented in Section 19.1, this shows that visitors in Condition 1 had to stay inside for longer, but still looked at pictures for the same amount of time. In line with Hypothesis (A2), average *Total Dwell Time* per picture was close to half a minute and did not differ across conditions (Cond. 1: M = 33.92., SD = 23.03; Cond. 2: M = 34.1., SD = 27.4; W = 14434.5, p = .398) nor was it different from the mean

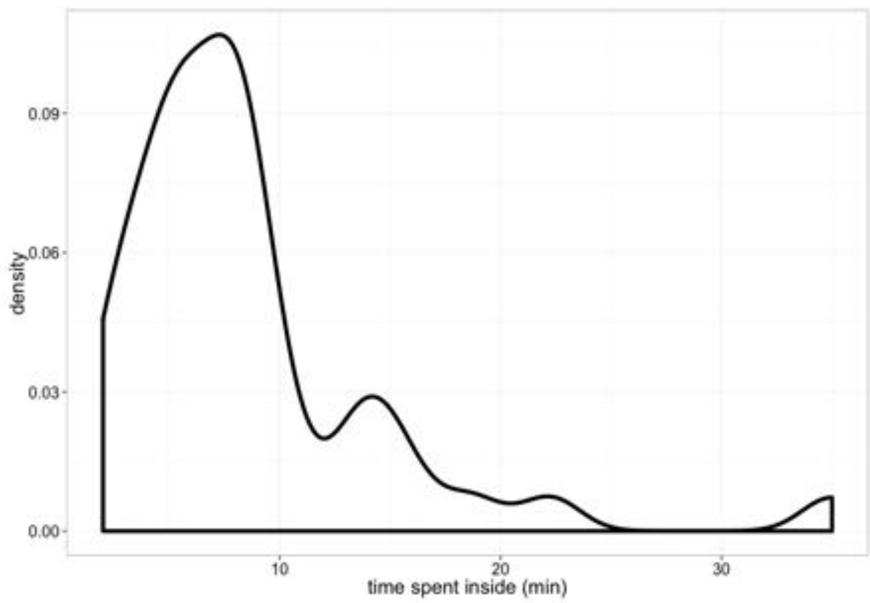


Figure 19.1.: Distribution of time spent inside the gallery by the experiment participants.

value obtained in Experiment 1 with a similar set of stimuli (Wilcoxon signed rank test against 31.71 V = 27328, p = .106).

Each picture was a subject of M = 5.92 dwells during a single gallery visit. Mean *Long-to-Short Dwell Ratio* was 0.46, meaning that an average participant performed about as many *short dwells* (i.e. dwells lasting less than 2 sec.) as *long dwells* (>2 sec.). As indicated by *Picture-Switching* metric equal to 0.35, about one third of all dwells were followed by an immediate fixation on another picture (although the means largely differed for Cond. 1 M = .04, and for Cond. 2 M = .51).

Mean *dwell length* was 5.09 sec. and the cumulative distribution of dwell lengths was heavy-tailed (in accordance to Hypothesis A1), although it differed across the conditions, indicating the presence of much larger proportion of longer dwells in Condition 1 (Figure 19.2; this effect will be investigated further below). The longest single dwell on a separate picture lasted for 131.35 seconds.

Engagement Ratio was 0.84, meaning that 84% of all dwell lengths inside the galleries (78% in Cond. 1 and 87% in Cond. 2) were deployed onto pictures which in total occupied only 2.9% of the entire wall length in Condition 1 and 4.9% in Condition 2[1]. *Isovist area size* value was used similarly to Experiment 1 to obtain expected *Engagement Ratio* proportions based on visibility and length of each object alone. In Cond. 1, *Isovist area* of picture location grid cells constituted 1.38% of the summed *Isovist area* of all grid cells adjacent to any vis-

1 Unlike in Experiment 1, total wall length differed across the conditions.

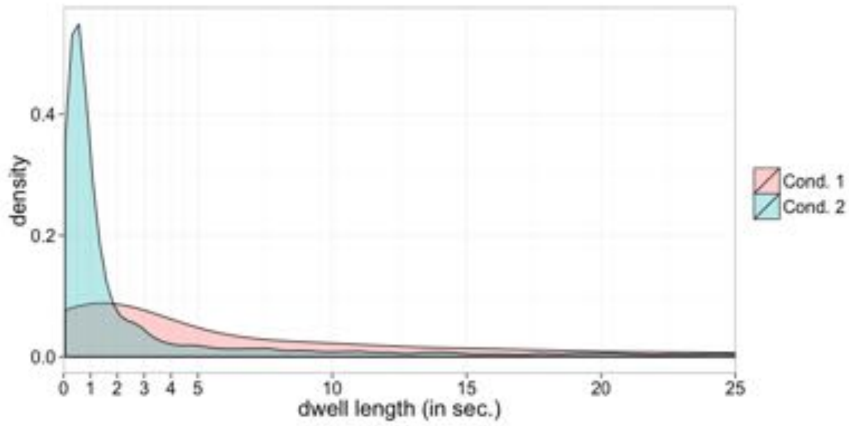


Figure 19.2.: Distribution of dwell lengths in Experiment 2 (cropped to 25 seconds).

ible wall surfaces[2], and in Cond. 2 it constituted 6.27%. Chi-Squared Goodness-of-Fit test was again used to compare the expected proportions to those which were actually observed. The observed proportions were significantly different from the expected ones, both in Condition 1 ($\chi^2(1)$ = 42.51, p<.001) and in Condition 2 ($\chi^2(1)$=11.01, p<.001).

Hypothesis (A4) was tested based on the analysis of *viewing sequence similarity* as described previously in Section 8.3. Personal *MNLD* values were compared across the conditions. The difference was significant ($W = 34$, $p = 0.005$), indicating that *viewing sequences* in Condition 1 were more similar to each other than those in Cond. 2.

In order to assess how much variability in visitors' oculomotor behaviour was due to spatial aspects, and how much due to by-picture, or by-participant random effect, linear mixed-effect models were once again used for the statistical analysis. Compared to Experiment 1, individual spatial aspects such as *VCA*, or *Co-Visibility* were not included in the model, as these variables varied very little by-location within conditions. That is to say, spatial properties of individual locations belonging to the same condition were very similar. Their *Co-Visibility* and *Isovist Areas* were identical within- but very different across-conditions. The presence of a significant random by-location effect (after accounting for the fixed effect of experimental condition) would suggest that other spatial factors (e.g. being located at one of the wall edges) should be considered to explain variability in participants' eye movement. Compared to Experiment 1, physical coordinates of picture locations remained unchanged (it were the walls surrounding them that varied across experimental conditions). There-

2 Note that these values take into account the width of each picture, as the calculations were based on the values of 3 grid cells for each artwork, corresponding to its real width in the used scale.

fore, locations from Condition 1 and from Condition 2 were both described in the data with the same identifiers (y01-y12). This made the models 'aware' that there is some degree of commonality expected for the same locations across the conditions[3]. However, because the influence of experimental condition was expected to be predominant even in the effect individual locations might have on visual attention, the random effect of location was entered to the models as random slope and random intercept effect (similarly to the expected but not confirmed effect of separate pictures in Experiment 1; see Section 15.2). That is to say, the model 'expected' that if any location is somewhat unique in the way it affects the visitors, this 'uniqueness' might have a stronger impact—and a different direction—in one condition than another[4].

Similarly to Experiment 1, *Number of Dwells* and *Total Dwell Time* were modelled as two key variables describing the quantitative aspect of the visitors' eye movement. To predict *Number of Dwells* falling on each picture, the following model was constructed:

```
log(No. of Dwells) ~ cond*Time Spent Inside +
cond*Salience Rating +
View. Sequence + age + gender +
(1|ids) + (1+cond|loc) + (1|pic)
```

Model 19.1: Logarithm of Number of Dwells (preliminary formula)

Experimental *condition*, as well as its interaction effect were predicted to bear the strongest influence on the *Number of Dwells*. Considering the findings from Experiment 1, *Time Spent Inside* was expected to be the second most important predictor.

The formula above was subject to the *step* function from `lmerTest` R package (Kuznetsova et al., 2014)[5]. Logarithmic (+1) transformation of *Number of Dwells* for visitor-artwork interactions was best predicted by the model described by the following formula:

```
log(No. of Dwells) ~ cond + Time Spent Inside +
View. Sequence + age + (1 | ids) + (1 | loc) +
cond*Time Spent Inside
```

Model 19.2: Logarithm of Number of Dwells

Normality and homoscedasticity of residuals of the final model were not violated. *Kappa* score was 5.2, *Marginal* R^2 was 0.55 and

3 No matter which condition participant explored the gallery in, location y01 was always on the left edge of the wall, location y03 was always in the centre of the sequence, and so forth.

4 For instance, it could be imagined that location y01 will differ from the rest, as it is likely to be explored as first, but in Condition 1 this effect might be smaller - due to the additional wall partitions the artwork cannot be seen immediately from the entrance area.

5 Please consider notes in Section 15.2 regarding the detailed procedure used throughout the modelling process and the limited deterministic role of stepwise model selection procedure.

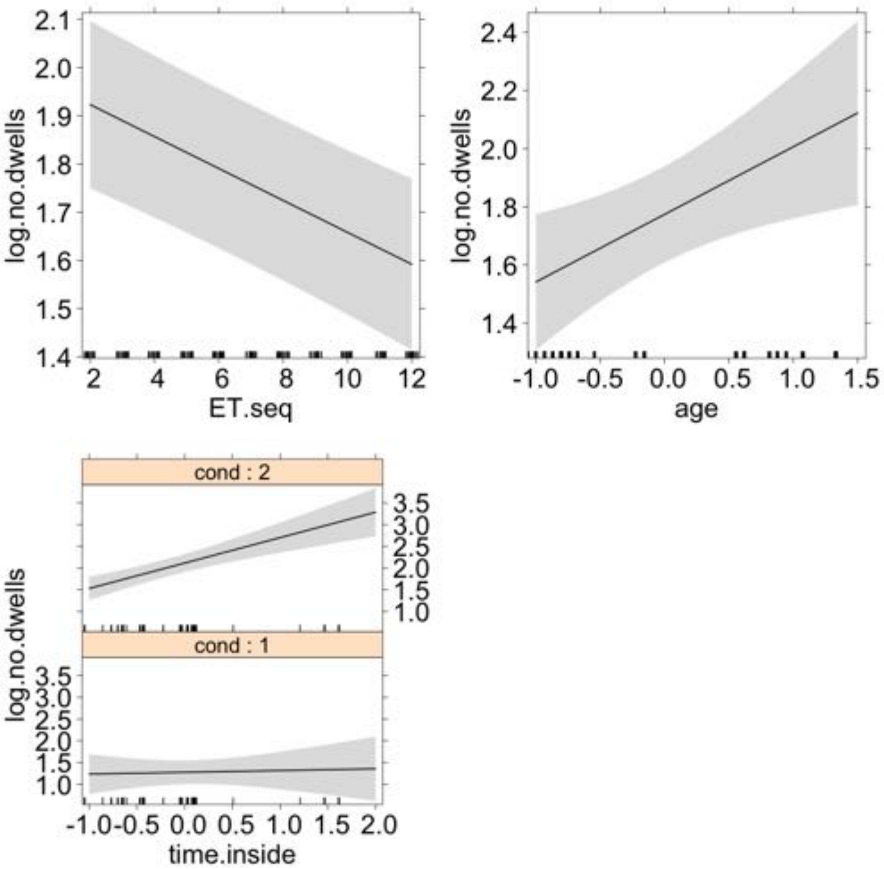


Figure 19.3.: Visualisation of the effects for Model 19.2.

Conditional R^2 was 0.86. Table 19.1 describes the model in detail and Figure 19.3 visualises the effects.

Total Number of Dwells was best predicted by the model:

```
log(Total Dwell Time) ~ Time Spent Inside +
(1|ids) + (1+cond|loc) + (1|pic)
```

Model 19.3: Logarithm of Total Dwell Time

meaning that the presence of wall partitions did not affect *Total Dwell Time*, except the fact that the size of the effect (slope) was different for individual locations across conditions. *Kappa* score was 1, Marginal R^2 was 0.70 and Conditional R^2 was 0.82. Normality and homoscedasticity assumptions were not violated.

Just like in Experiment 2, *Picture-Switching* and *Long-to-Short Dwell Ratio* were selected as key variables describing the dynamics of eye movement.

Picture-Switching ratio could not be reliably explained with a linear mixed-effect model, as a large proportion of data points within Condition 1 (84%) was equal to 0. In this case, unsurprisingly, mean *Picture-Switching ratio* significantly differed between the conditions ($W = 2185$, $p<.001$).

	Model 19.2
(Intercept)	$1.49\ (0.14)^{***}$
cond2	$0.84\ (0.17)^{***}$
Time Spent Inside	$0.04\ (0.18)$
View. Sequence	$-0.03\ (0.01)^{***}$
age	$0.23\ (0.09)^{**}$
cond2:Time Spent Inside	$0.54\ (0.21)^{*}$
AIC	235.83
BIC	270.18
Log Likelihood	-108.92
Num. obs.	336
Num. groups: ids	28
Num. groups: loc	12
Variance: ids.(Intercept)	0.17
Variance: loc.(Intercept)	0.01
Variance: Residual	0.08

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.1.: Model 19.2.

The following Model 19.4 best predicted *Long-to-Short Dwell ratio*:

```
Long-Short Dwell Ratio ~ cond + View. Sequence +
(1|ids) + (1|loc)
```

Model 19.4: Long-to-Short Dwell Ratio

with *Kappa* score reaching 4.6, Marginal R^2 equal to 0.26 and Conditional R^2 reaching 0.55. Participants in Condition 2 had lower *Long-to-Short Dwell Ratio*, meaning the percentage of *long dwells* in their average interaction with an artwork was smaller. Table 19.2 presents the statistical summary.

Informal observations of the eye-tracking videos suggested that while the one-way mirror (typically used for experiment observation) located near the gallery entrance brought little attention from most of the visitors, observing one's own reflection with the eye-tracking device seemed engaging to some. In order to control whether this fact had an influence on any other aspects of participants' oculomotor behaviour, visitors from both conditions were split into separate subgroups based on the median total time of dwells spent on the mirror. All aforementioned analyses were repeated to control whether participants who looked at the mirror for longer differed from those who looked at the mirror for less time. Adding median-split factor of *time spent glimpsing at the mirror* did not significantly improve the fit of any of the above models, suggesting that the potentially distractive presence of the mirror had no meaningful effect on participants' oculomotor behaviour.

	Model 19.4
(Intercept)	$0.63\ (0.06)^{***}$
cond2	$-0.30\ (0.06)^{***}$
View. Sequence	$0.01\ (0.00)^{**}$
AIC	-51.25
BIC	-28.13
Log Likelihood	31.62
Num. obs.	348
Num. groups: ids	29
Num. groups: loc	12
Variance: ids.(Intercept)	0.02
Variance: loc.(Intercept)	0.00
Variance: Residual	0.04

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.2.: Model 19.4.

Adding lighting conditions as a predictor (described by the number of LUX registered at each location) did not significantly improve the fit of any of the models, even though locations in Condition 2 were slightly better lit on average, due to the lack of additional walls casting shadows (Cond. 1: M = 1171 lx, SD = 644; Cond. 2: M = 1408 lx, SD = 571).

The explanation of significant random by-location effect was attempted by classifying locations as 'edge' (y01, y06, y07, y12) and 'non-edge' (all others) locations. The assumption being that such artworks were likely to become first or last points of engagement in—otherwise linear and symmetric—viewing sequences. Including this factor did not significantly improve the fit of any of the models, demonstrating that some other spatial factors played a role in significant by-location variability of eye movement. However, the analysis of the by-location boxplot visualisations of all dependent variables for which random by-location effect significantly improved the fit of the model, could not bring any clear conclusions. The only exception is a noticeable distinction of the locations y01-y03 in Condition 1 and location y12 in Cond. 2.

Hypothesis (B7) assumed that longer, more diligent visual interactions with artworks will be equally dependent on space as the shorter ones. Similarly to Experiment 1, the possibility of testing this statement reliably in the set-up of Experiment 2 was limited. Additionally to technological limitations mentioned previously, the spatial by-location variability within-condition of Exp. 2 was minimised and the across-condition variability had a strong impact on *all* aspects of oculomotor behaviour. However, the average number of *long* *dwells* per picture was not significantly different across the conditions (W = 12698.5, p=.248).

19.3 RECOGNITION MEMORY

Similarly to Experiment 1, the first part of the Recognition Memory analysis considered all responses to the Recognition Task (without the Eye-Tracking data and even for those participants, for whom ET data was not available). The following linear mixed-effect model was proposed to predict inverted, standardised and mean-centred Reaction Times:

```
RT(inv) ∼ cond + Time Spent Inside +
Salience Rating + LUX + age + gender +
RTtrial + proceedingRT + proceedingRTacc+
(1|ids) + (1|pic) + (1+cond|loc)
```

Model 19.5: Recognition Times (inverted)

This model, with *Kappa* score reaching 4, has achieved *Conditional* R^2 of 0.45. However, *Marginal* R^2 was only 0.06 and it was driven primarily by the statistically significant effects of *RT trial* and *proceeding RT*, as it is visible in Table 19.3.

Similarly to the equivalent analysis in Experiment 1, the by-participant random effect was the strongest factor affecting high *Conditional* R^2. This, once again, presented a risk that the differences between individual participants in their Recognition Memory performance are so large that they 'steal' explained variance from relatively strong but still insignificant fixed effects (compare t values in the table for *age*, *time spent inside* and *condition*). To test this assumption, a clustered t-test was conducted (similarly to Experiment 1) but its result was far from reaching significance (p=.583) [6]. Visual analysis of boxplots confirmed, that this was true both for the group which explored the gallery quickly and the one which spent longer (i.e. longer then median) time inside (Figure 19.4).

An analogous logit mixed-effect model was constructed for *RT accuracy* measure in the following form:

```
RT Accuracy ∼ cond + Time Spent Inside +
Salience Rating + LUX + age + gender +
(1|ids) + (1|pic)
```

Model 19.6: Recognition Test Accuracy

Its *Marginal* R^2 was 0.05 and *Conditional* R^2 0.17, meaning that *RT accuracy* could not be well predicted using the spatial factors alone. As visible in Table 19.4, again it was the by-participant variance which

6 Assuming that the true population effect size of VCA on Reaction Times is 'large' and equals to d = 0.8, the statistical power of this test was β = 0.69. See Section 26.6 for the discussion of why the true population effect might be even larger.

	Model 19.5
(Intercept)	0.05 (0.23)
condwithout	−0.20 (0.23)
Time Spent Inside	0.10 (0.11)
Salience Rating	0.05 (0.08)
LUX	0.06 (0.05)
age	−0.10 (0.11)
gender-male	0.06 (0.23)
RTtrial	−0.08 (0.04)*
proceedingRT	−0.24 (0.07)***
proceedingRTacc-TRUE	−0.09 (0.11)
AIC	1000.00
BIC	1062.74
Log Likelihood	-484.00
Num. obs.	373
Num. groups: ids	39
Num. groups: pic	12
Num. groups: loc	12
Variance: ids.(Intercept)	0.36
Variance: pic.(Intercept)	0.05
Variance: loc.(Intercept)	0.01
Variance: loc.condwithout	0.00
Variance: Residual	0.58

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.3.: Model 19.5.

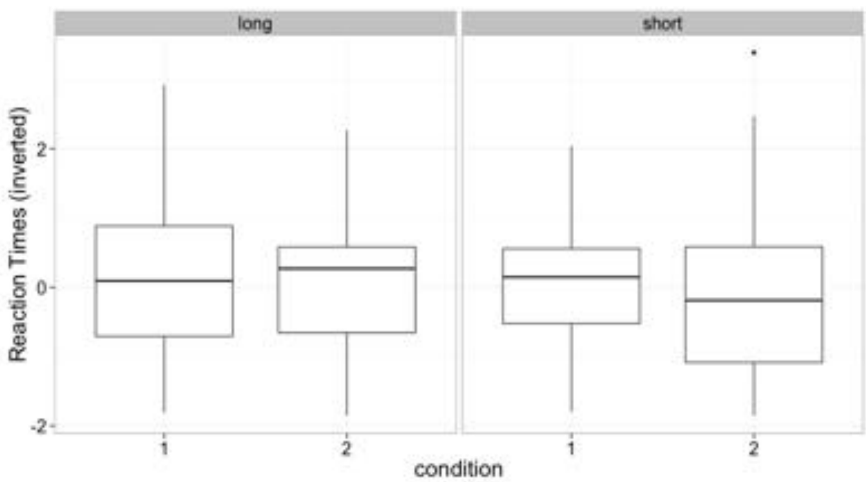


Figure 19.4.: Recognition Memory performance across the conditions, with distinction between participants staying inside for longer and shorter.

	Model 19.6
(Intercept)	$2.03\ (0.35)^{***}$
cond.2	$-0.02\ (0.37)$
Time Spent Inside	$0.41\ (0.25)$
Salience Rating	$0.03\ (0.17)$
LUX	$-0.20\ (0.14)$
age	$-0.10\ (0.17)$
gender-male	$0.02\ (0.36)$
AIC	391.10
BIC	428.43
Log Likelihood	-186.55
Num. obs.	468
Num. groups: ids	39
Num. groups: pic	12
Variance: ids.(Intercept)	0.33
Variance: pic.(Intercept)	0.13
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.4.: Model 19.6.

resulted in the largest differences in the dataset, although even this did not contribute to a higher than 'medium' effect size of *Conditional* R^2. This can suggest that there was generally little consistency in the *accuracy* of participants' Recognition Memory within the controlled factors. It is important to note, however, that in this simplified spatial setting—and in the presence of slightly smaller number of pictures (12 instead of 14)—the ratio of correct *RT* responses was generally higher compared to Experiment 1 (86% in Exp. 2 vs. 78% in Exp. 1). In this case, *RT accuracy* is not a reliable measure for investigating effects others than the very strong inhibiting ones.

To account for the available Eye Movement data, the above analyses were repeated for the subset of participants for whom eye-tracking recordings were available. Note that in this study, *Total Dwell Time* and *Number of Dwells* correlated only to a moderate extent (r = 0.34), and therefore there was no reason to de-correlate the variables by transforming them into the *Quantity of ET* measure. The reason for the lack of strong correlation might be the very distinctive spatial setting employed across the experiments. As it has been observed in the previous section, especially in Condition 1, the viewing patterns consisted typically of long (but very few) dwells per picture. Compared to Experiment 1, where correlation between the two measures was strong, it seems that the lack of it is the result of this unnatural spatial situation enforced on the viewers. Moreover, *Time to First Fixation* variable was highly correlated with *ET sequence* and thus was removed

	Model 19.7.
(Intercept)	0.25 (0.30)
cond.2	−0.08 (0.32)
Total Dwell Time	0.10 (0.13)
No. of Dwells	0.03 (0.11)
Picture-Switching	−0.04 (0.08)
Long-Short Dwell Ratio	0.11 (0.08)
View. Sequence	−0.03 (0.02)
Time Spent Inside	0.03 (0.25)
Salience Rating	0.06 (0.08)
LUX	0.12 (0.06)$^{\cdot}$
age	−0.21 (0.14)
gender-male	−0.15 (0.29)
RTtrial	−0.12 (0.05)*
proceedingRT	−0.21 (0.07)**
proceedingRTacc-TRUE	−0.12 (0.13)
AIC	718.52
BIC	793.29
Log Likelihood	-338.26
Num. obs.	260
Num. groups: ids	28
Num. groups: pic	12
Num. groups: loc	12
Variance: ids.(Intercept)	0.40
Variance: pic.(Intercept)	0.06
Variance: loc.(Intercept)	0.00
Variance: loc.cond.2	0.01
Variance: Residual	0.54

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.5.: Model 19.7.

from the model. Also *Engagement Ratio*—as a measure aggregated by participants—strongly correlated with *experimental condition* and was therefore omitted.

```
RT(inv) ~ cond + Total Dwell Time + No. of Dwells+
Picture-Switching + Long-Short Dwell Ratio +
View. Sequence + Time Spent Inside + Salience Rating+
LUX + age + gender + RTtrial + proceedingRT +
proceedingRTacc + (1|ids) + (1|pic) + (1+cond|loc)
```

Model 19.7: Reaction Times (inverted) with respect to Eye Movement

The model had *Kappa* score of 5.6, *Marginal* R^2 of 0.11 and *Conditional* R^2 was 0.52. Table 19.5 describes it in detail.

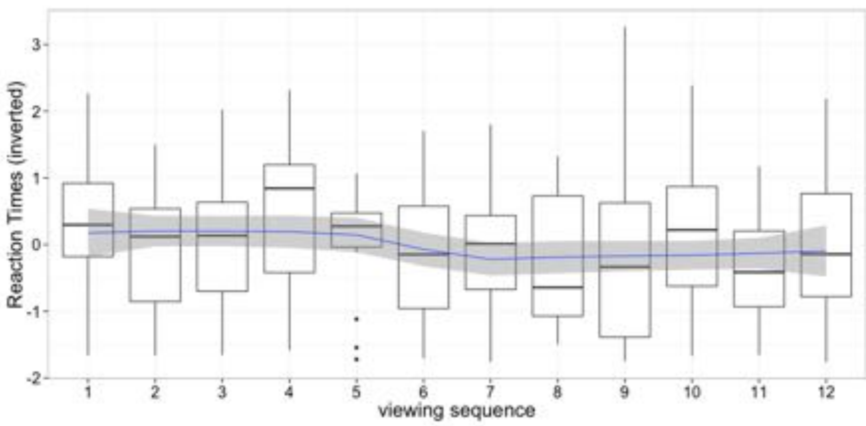


Figure 19.5.: Variability in Recognition Memory performance depending on the sequence of seeing each artwork.

As the factor *ET sequence* had a relatively low t-value, it could be expected that its potential influence on Recognition Memory is non-linear but still relevant for further interpretation (as it was for Spatial Memory in Experiment 1). Figure 19.5 presents how quickly participants reacted to each picture depending on the place in the sequence they saw it at. The visible difference is too small to reach significance, especially under such large variation across participants.

Accounting for eye-tracking data for *RT accuracy* led to the construction of the following model:

```
RT Accuracy ~ cond + Total Dwell Time + No. of Dwells+
Picture-Switching + Long-Short Dwell Ratio +
View. Sequence + Time Spent Inside +
Salience Rating + LUX + age + gender +
(1|ids) + (1|pic) + (1+cond|loc)
```

Model 19.8: Recognition Test Accuracy with respect to Eye Movement

Its *Marginal* R^2 was 0.16, and *Conditional* R^2 0.31. Table 19.6 presents the detailed summary.

19.4 SPATIAL MEMORY

Spatial Memory performance in this Experiment could only be analysed on the basis of an aggregated result for each participant, based on a String Matching calculation technique (Section 18.5). To compare the participants' Miniature Task performance across conditions, a t-test was used. It showed no significant difference ($t(39)$ = 0.4, p=.689). A visual analysis of the dataset (Figure 19.6) revealed that, similarly to Experiment 1, the variance of Spatial Memory results in Condition 2 was noticeably larger compared to Condition 1.

	Model 19.8
(Intercept)	2.65 (0.65)***
cond.2	−0.53 (0.66)
Total Dwell Time	1.39 (0.55)*
No. of Dwells	0.38 (0.42)
Picture-Switching	0.31 (0.27)
Long-Short Dwell Ratio	0.42 (0.26)
View. Sequence	−0.03 (0.06)
Time Spent Inside	−1.28 (0.65)·
Salience Rating	0.01 (0.20)
LUX	−0.25 (0.19)
age	−0.22 (0.25)
gender-male	−0.23 (0.48)
AIC	299.97
BIC	364.86
Log Likelihood	-132.99
Num. obs.	336
Num. groups: ids	28
Num. groups: pic	12
Num. groups: loc	12
Variance: ids.(Intercept)	0.54
Variance: pic.(Intercept)	0.15
Variance: loc.(Intercept)	0.00
Variance: loc.cond.2	0.00
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.6.: Model 19.8.

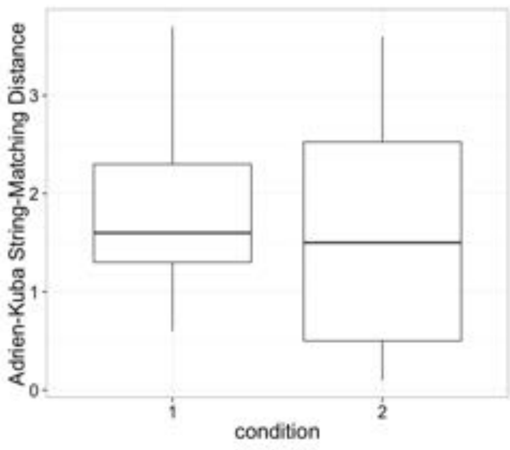


Figure 19.6.: Spatial Memory performance (the higher, the worse) across the conditions.

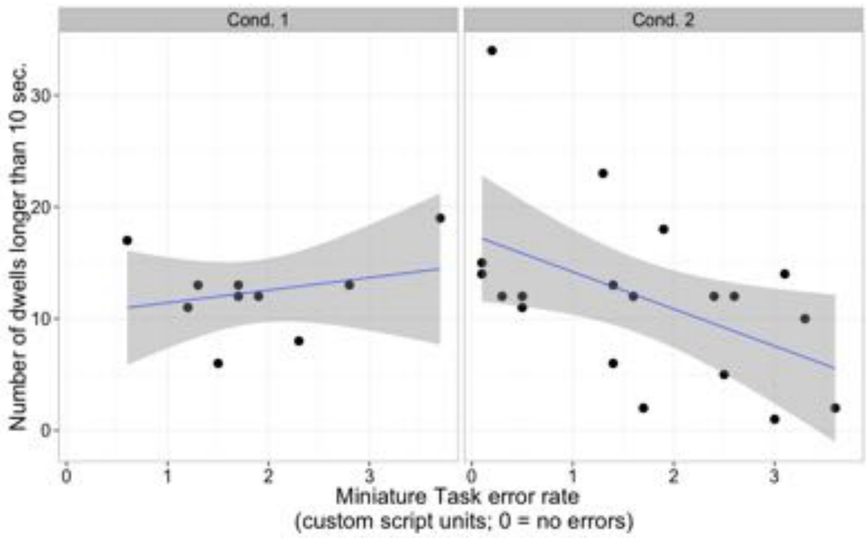


Figure 19.7.: Spatial Memory performance (the higher, the worse) across the conditions.

In order to further explore this relation, detailed correlation patterns were investigated. They showed that Spatial Memory performance was correlated with the *number of long dwells* (i.e. longer than 10 seconds), but only in Condition 2 ($\rho = -0.53$; p=.019). The relation, however, was not significant in Condition 1 ($\rho = 0.14$; p=.702). Figure 19.7 presents this visually.

To perform the Miniature Task Sequence Analysis, once again linear mixed-effect model was used. Table 19.7 presents the results.

Only two predictors were significant: *Total Dwell Time* and *Time Spent Inside*. The latter variable was included to control for the error coming from different times participants spent inside the gallery. After correcting for this factor, *Total Dwell Time* was also significant, with an estimator meaning that an increase in 1 SD on this measure for a particular artwork resulted in this picture being located about 2.09 places earlier in the sequence of the Miniature Task solution[7]. The longer participants looked at an individual picture, the more likely they were to solve it earlier in the computerised Miniature Task. The effect explained 11% of the variance. Moreover, the effect of *RT inverted* was approaching significance. When tested on a larger part of the dataset (excluding the eye-movement factors, and therefore including even participants for whom ET data was unavailable), the effect reached significance (p = .013). This indicates that there was a direct linkage between how fast participants answered to the question in Recognition Memory Test and how early they solved the location of the particular artwork in the Miniature Task.

7 Note that the estimator is negative since a higher order number was associated with a 'worse' performance (i.e. a more distant placement in the sequence)

	Model 1
(Intercept)	6.01 (0.77)***
cond.2	0.47 (0.73)
Total Dwell Time	−2.09 (0.52)***
No. of Dwells	−0.21 (0.26)
Picture-Switching	0.30 (0.30)
Long-Short Dwell Ratio	0.31 (0.27)
View. Sequence	0.09 (0.09)
Time Spent Inside	2.77 (0.66)***
Salience Rating	0.17 (0.20)
LUX	−0.48 (0.42)
age	0.07 (0.24)
gender-male	−0.09 (0.44)
RT(inv)	−0.33 (0.20)
AIC	1387.11
BIC	1454.77
Log Likelihood	-674.56
Num. obs.	260
Num. groups: ids	28
Num. groups: pic	12
Num. groups: loc	12
Variance: ids.(Intercept)	0.00
Variance: pic.(Intercept)	0.00
Variance: loc.(Intercept)	0.95
Variance: loc.cond.2	0.76
Variance: Residual	9.85

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 19.7.: Model of the Miniature Task Sequence Analysis.

There was no significant impact of *age* on mean Miniature Task score.

20

DISCUSSION

This section will review the results of Experiment 2 in relation to the hypotheses stated previously in Sections 12 and 17.6, as well as to the results of Experiment 1, discussed in Section 16. The discussion is based on the structure introduced in the previous study, and where possible, generic interpretations will be made in consideration to both Experiments 1 and 2. As backward comparisons across multiple studies are made throughout the thesis, these sections form the main part of the result's discussion.

20.1 VISUAL ATTENTION

The similarity of visit lengths to those observed in Exp. 1 indicates that even under radically different spatial arrangement, a similar set of artworks is likely to result in a similar 'visit-threshold' length. Moreover, the length of average cumulative viewing time per picture, and the distribution of dwell lengths were largely similar to those observed in Experiment 1 on a similar set of stimuli, thus confirming Hypotheses (A1) and (A2). Condition 2 promoted longer dwells, as it was expected in such a restricted environment where little alternative stimuli is available to the viewer at any given moment.

Once again, the proportion of dwells falling on- and off-pictures was compared with the proportions expected from the artworks' width and visibility alone. As predicted by Hypothesis (B1), and similarly to Experiment 1, the proportion of dwells falling on pictures suggests that the visitors' oculomotor behaviour was attracted to objects hanging on the walls and therefore was task-specific. Looking at pictures seemed to constitute the 'default' cognitive aim of the art gallery visits. This is an important observation in the experimental conditions hosted at a research building of a large university. Movable wall partitions, a one-way mirror as well as other details left no doubt to the lack of authenticity of this arranged study situation. And yet, it appears that the visitors had no difficulties with following the instruction of 'exploring the gallery as you would explore any other art gallery'.

Similarly to Experiment 1, Hypothesis (A4; p. 141) was confirmed through the analysis of *viewing sequence similarity*. *Viewing sequences* in Condition 1 were more similar to each other than were those performed by the visitors of Cond. 2. This fact indicates that additional wall partitions decreased the variability in the choice of visual targets. When exploring a simple gallery with decreased *Co-Visibility*, participants were more likely to use homogenous trajectories. As previously observed in Experiment 1, similarity of trajectories in an open-plan art gallery was there most likely caused by the local grouping of artworks, encouraging the visitor to engage with those groupings. This explanation is less convincing in Experiment 2. Here, the similarity of *viewing sequences* seems to be possible more on the global level - the shape of the gallery encouraged circular movement, and the limited choice of visual targets under Condition 1 contributed to rather homogenous global sequences.

Model 19.2, described by the following formula

```
    log(No. of Dwells) ~ cond + Time Spent Inside + age +
(1 | ids) + (1 | loc) + cond*Time Spent Inside
```

explained about 85% of all variance in the *Number of Dwells* variable. Similarly to Experiment 1 conducted on the same set of stimuli, random effect of pictures and fixed effect of *Salience Rating* did not improve the fit of the model, meaning that the variation between individual artworks did not contribute to explaining the variance in the number of eye glimpses. According to the expectations, experimental *condition* (representing spatial differences) had the largest impact. Participants exploring the gallery without the wall partitions (Cond. 2) performed more dwells per picture. Basing the interpretation on Experiment 1, it can be noted that multiple spatial factors differentiating the conditions must have played a role in this result, as *Co-Visibility* had no impact on the *Number of Dwells* in the previous study. Since *VCAs* were kept similar across the conditions, most likely larger *isovist areas* (affording more diverse viewing angles and distances) were a considerable factor. Such an interpretation might seem contrary to the one provided after the Experiment 1, where the quantity of eye movement is believed to have been consciously guided towards larger *VCAs*, as they provided more comfortable viewing conditions. It must be noted however, that the spatial differences arranged in the currently analysed study have been much more extreme. Throughout this Discussion the reader should bear in mind that in Experiment 2 spatial differences were present not only in what the gallery space 'afforded', but rather what it 'made possible'. As it might be the case with the result describing *Number of Dwells*, perhaps the significant effect should not be read as 'Condition 2 generated higher *Number of Dwells*' but rather 'Condition 1 limited *all possible Number of Dwells* to minimum'. This effect, enforced by drastically modified spatial arrangement was planned in order to test its influence on visitors' memory, and further part of this Discussion should be read along this criterion.

Interestingly, *Time Spent Inside* only had a significant, positive influence on *Number of Dwells* in Condition 2, and no significant effect in Cond. 1. This suggests that in the presence of additional wall partitions in Condition 1, participants staying inside for longer did not use the additional time to perform larger *Number of Dwells*. So restricted space, 'normalised' number of dwells performed per picture, no matter how long the visitor was inside. Since *Time Spent Inside* did significantly increase *Number of Dwells* performed in Condition 2, participants staying for longer rose the mean *Number of Dwells* in this condition and created a statistically significant effect on the output variable. It thus can be stated, that the cross-conditional difference in *Number of Dwells* is primarily due to the restrictive character of Cond. 1 which minimised number of glimpses for all participants, no matter how much time they were prepared to devote to the visit. By consulting the variance of random effects we can notice that the majority of the gain in *Conditional* R^2 compared to *Marginal* R^2 comes from unexplained individual differences across participants, and not differences across locations. This emphasises the importance of individual exploration strategies employed by the visitors. The random effect of location was, however, significant, and thus will be investigated further on. Also, age became a significant, positive predictor demonstrating that older participants performed more dwells, in both conditions.

In Model 19.3:

```
log(Total Dwell Time) ~ Time Spent Inside +
(1|ids) + (1+cond|loc) + (1|pic)
```

up to 70% of participants' variance in *Total Dwell Time* was predicted by the single fixed effect of *Time Spent Inside* (*Marginal* R^2). Considering the lack of this factor's influence on *Number of Dwells* in Cond. 1, we can conclude that participants staying inside the gallery for longer spent cumulatively more time fixating on pictures, even though the average number of these engagements was constant (and therefore their average length increased) in Condition 1. This finding is in line with Hypothesis (E2) predicting strong influence of this factor. However, the lack of significant cross-conditional difference in *Total Dwell Times* is contrary to Hypothesis (B4), which stated that cumulative viewing time will be longer when 'complete isolation' is ensured.

This finding can be interpreted more clearly, when we turn again to the very first assumption underlying the design of experimental spaces arranged in this study: the fact that all *VCAs* are similar across conditions. It thus seems, that the aforementioned 'optimal viewing conditions' are provided by sufficient areas of comfortable viewing in front of the artworks, and that the influence of *Co-Visibility* is less prevailing, at least on *Total Dwell Times*. Due to the large proportion of variance being explained by *Time Spent Inside* alone, the unexplained effect of space (present in the significant random-effect factor) remains relatively small.

Picture-Switching ratio differed noticeably across the conditions - 84% of all interactions in Condition 1 were not followed by an immediate engagement with another object. Previously, Experiment 1 had shown that the only predictors of *Picture-Switching ratio* considered in this thesis which remained significant were fixed spatial and random individual factors. Thus, the significant difference in *Picture-Switching ratio* between the two conditions of Experiment 2 is unsurprising given the spatial amendments purposefully designed to limit this type of behaviour. The result shows, that the manipulation was successful.

Another Model, 19.4 attempted to further explain the dynamics of visitors' oculomotor behaviour:

```
Long-Short Dwell Ratio ~ cond + (1|ids) + (1|loc)
```

Condition 1 caused participants to perform higher proportion of long dwells, thus decreasing the 'spotty-focused' (using Bitgood's term) attention and facilitating more diligent interaction with individual artworks. The presence of random by-participant and by-location predictors improved the fit of the Model 19.4 from Marginal R^2 0.26 to Conditional R^2 of 0.55. This suggests that a lot of unexplained variance lies in interpersonal differences (yet other than age or gender), as well as in the by-location variability (although different than the presence of wall partitions). The by-participant variability can be linked to different exploration styles, easily noticeable to any third-person observer decided to informally investigate social differences inside a public art gallery. However, similarly to Experiment 1, this outcome variable was the most difficult to predict (even after considering random effects) and it must be interpreted with caution.

The above analyses, as well as informal observations of the eye-tracking recordings confirm that the cross-conditional manipulation successfully affected the quantity and dynamics of the visitors' visual attention. Participants in Cond. 1 engaged with the artworks through a higher proportion of longer viewing periods (i.e. less 'spotty-focused' attention) in the presence of wall partitions. *Picture-Switching ratio* was there minimised, demonstrating that only in a very limited number of cases the viewers were able (and willing) to immediately switch their eye gaze from one artwork to another. These situations were only physically possible while traversing past the wall partition.

Participants staying inside the gallery for longer cumulatively looked at pictures for longer, but in Condition 1 did it through almost equal average number of dwells. This suggests that participants preferring to explore the gallery in a slower manner did so most likely in a linear way, not increasing the number of circulations but prolonging the undisturbed engagement periods instead. The decision to stay inside for longer was therefore not spontaneous, but links directly to individual viewing styles and preferences.

Condition 1—as it was predicted by Hypotheses (B5) and (B6)—also allowed visitors to focus on individual pictures with less visual interruptions. To better understand this effect we can consider that the difference between spatial arrangements of Conditions 1 and 2 could be potentially described by an operationalisation different from *VCAs*, *Co-Visibility* and other isovist-derived measures mentioned previously. Compared to Condition 2, Condition 1 'linked' walking trajectories with the availability of visual targets. In both conditions participants were likely to walk a similar path but the fragmented character of the visual experience in Cond. 1 decreased the diversity of viewing possibilities, resulting in a more focused, but also less 'comparative' visual engagements. This alternative operationalisation however, once again would be path-dependent, and therefore would provide a limited practical value for the potential exhibition designer.

Importantly, experimental *condition* had no significant effect on *Time Spent Inside* by the visitors. Thus, the aforementioned differences in eye movement patterns are due to how space modified the oculomotor behaviour, and not how it modified the length of the visits. Even though larger proportion of that time in Condition 1 was spent looking at empty wall segments (quantified by significantly lower *Engagement Ratio*), *Total Dwell Times* per picture were better predicted by *Time Spent Inside* alone, without the interaction effect of experimental condition. Thus, it was more the by-visitor variability in time one wished to devote to exploring the gallery than the influence of generalisable spatial factors, that facilitated cumulative viewing times. In relation to the discussion we have already undertaken in Section 16.1, it is questionable to what extent visit durations might be affected by simple spatial modifications.

In line with the findings of Experiment 1, *ET Sequence* had a significant impact on decreasing *Number of Dwells* but no such influence on cumulative viewing times. Instead, *Long-to-Short Dwell ratio* increased for artworks viewed later, meaning that as participants explored the gallery, they tended to do it less dynamically, with less numerous, but more diligent dwells. Similarly to what has been observed in Experiment 2, this change in viewing behaviour took place without a sacrifice to *Total Dwell Time* devoted to pictures.

Regarding the role of the artworks' content, Hypothesis (A3) was again confirmed, as *Salience Rating* and the random by-picture effect had little effect on the oculomotor behaviour, slightly improving the fit of only one of the aforementioned models [1]. Compared to Experiment 1, more radical hanging arrangement introduced in this study as well as removing two stimuli closest to the mean *Salient Rating* did not—contrary to Hypothesis (E3)—increase the by-picture variance

1 Namely, random by-picture effect in Model 19.3.

irrespective of their location. Visual content of the artworks used in both experiments had negligible impact on the visual attention of the gallery visitors.

In line with Hypothesis (B7), Experiment 1 showed that when interactions consisting primarily of *long dwells* are considered, the influence of space on visitors' eye movement does not seem to be minimised. The results of Experiment 2 to a possible extent demonstrated that the number of *long dwells* did not differ significantly across conditions. Thus, the above described differences in the dynamics of visitors' eye movement across two experimental settings derive rather from the decreased number of *short dwells*, than an increased number of *long dwells* in Condition 1. It can be stated, that the gallery with wall partitions was 'less distracting' but not clearly 'more focusing'. As a result, participants performed about 2 *long dwells* per picture on average, no matter what type of setting they were in, thus providing strong evidence against Hypothesis (B7). In the presence of conflicting interpretation from the two experiments, the current one is preferred, as it is based on a spatial setting better suited for testing this hypothesis.

This finding remains also in contrary to Hypothesis (B8), which suggested that multiple pictures hanging on the same wall are viewed as holistic creative entities, thus resulting in smaller number of *long dwells* dedicated individually to each of those artworks. It must be noted, however, that the arrangement of Condition 2 only consisted of two long walls occupied by artworks, and in the lack of contrasting (more isolated) picture locations, such situation might not have encouraged the perception of six adjacent paintings as a single creative work.

It remains to be seen how these modified patterns of eye movement have affected Spatial Memory and Recognition Memory performance. It cannot be ignored, that despite the lack of significant effect of experimental condition on *Total Dwell Times*, visual interactions occurring in Cond. 2 were (a) more numerous, due to a 'spotty-focused' attention; and (b) conducted from much more diverse set of angles and distances than in Condition 1 (which only allowed participants to view artworks from within their *VCAs*). Intuitive assumption would suggest, that if cumulative viewing times are similar, but the conditions of viewing are much worse in Cond. 2, memory performance following such 'badly optimised' artwork viewing shall be worse. On the contrary, Hypotheses (C6) and (E4) predict that cognitively deeper processing of artworks during an unrestricted gallery visit occurs primarily within the painting's *VCA*. In such a case, the presence of dwells from less optimal distances and angles (primarily *short dwells* occurring at oblique angles, as informal observations of ET recordings would suggest), will have little effect on visitors' Recognition Memory, and thus no significant difference across the conditions is expected. This is due to the fact that *VCAs* allowing the visitors to comfortably engage with an artwork during their free exploration of space are very similar.

20.2 RECOGNITION MEMORY

The lack of significant effect of *Condition* on Recognition Memory lies in line with the previously stated hypothesis. Despite a very strong impact of space on participants' oculomotor behaviour (as described in Section 20.1), this difference in viewing patterns did not result in a consistent change in Recognition Memory performance. Accounting for Eye Movement data did not substantially improve the overall performance of the models. However, the Model 19.7 was able to account for over a half of the variance in *Reaction Time* data, even though the majority of this variation was due to the by-participant random effect.

In contrast to Hypothesis C5, larger *Isovist Areas* did not increase participants' Recognition Memory performance. Compared to the corresponding results from Experiment 1, this fact supports the statement that *Visibility Catchment Areas* are a description more relevant for capturing the influence of artwork's location on visitors' memory (Hypotheses C6 and E4). Even as *Co-Visibility* is affected by the presence (or lack) of the wall partitions, the size of the floor area from which each painting can be viewed comfortably (straight-on) is similar. The two layouts arranged in this experiment facilitated the type of movement which is most typically observed in traditional art galleries: a slow, linear progression across space, in a straight line from one picture to another. This space supported such viewing pattern to the maximum extent possible. Space having many pictures located near room corners or in smaller subspaces (with more varied *VCAs*) would 'break' this movement pattern more often and potentially contribute to the significant impact of layout on Recognition Memory observed in Experiment 1.

Total Dwell Time significantly predicted *RT accuracy* similarly to Experiment 1, showing a direct relation between viewing times and Recognition Memory. The result is in line with the previous findings, as well as Hypothesis D4. Once again, it cannot be interpreted, however, without considering the fact that the majority of all answers in the Recognition Memory test were correct (86% in this experiment). As it has been mentioned previously, this limits the applicability of this measure for more detailed statistical analyses and makes it only reliable for detecting very strong effects.

In all models describing the influence of space and oculomotor behaviour on Recognition Memory, *Time Spent Inside* consistently appeared as the factor of the highest importance. This is in line with Hypothesis E2. It is important to note however that—compared to Experiment 1—time had a much more linear influence. That is to say that the more participants stayed inside the gallery, the better their memory for the artworks was. However, the radical cross-conditional difference in spatial arrangement had no role to play in this effect, as it did in Experiment 1. Since both spatial layouts facilitated Recognition Memory in a similar manner, participants staying inside any of the two arrangements for longer, were equally likely to leave with stronger memories of the viewed artworks.

In all models described in Section 19.2, the consistent variation found within individual locations was negligible. This means, that the experimental set-up successfully managed to eliminate by-location variability other than this deriving from the probability of viewing some locations as the first/last ones in the sequence. Even this aspect, however, had no clear impact on participants' Recognition Memory. The aim of this particular set-up was therefore successfully achieved - each artwork location provided equal opportunities for processing its content. Compared to Experiment 1, it is important to note how the random by-location effect increases in space where artwork locations are more diversified. A potential curatorial strategy deriving from this finding would be to explicitly plan the exhibition with the aim of either providing *equal*, or *varying* 'processing opportunities' for each artwork location. On the cumulative level of multiple diverse gallery visits the main differentiation between *equal* or *varying* spatial settings might be one of the most impactful decisions a curator faces during the design process.

Contrary to Hypothesis (C1), higher *Co-Visibility*—despite modifying the dynamics of the oculomotor behaviour—did not reduce the Recognition Memory performance by the 'divided attention' memory effect. Perhaps, the inhibiting influence of space on memory is not as impactful if the 'minimal sufficient' conditions for comfortable viewing of artworks are provided. That is to say, the 'divided attention' effect often observed in laboratory-based studies, looses on importance when participants have a chance to explore the environment in a less restricted manner and this environment provides equally suitable opportunities for engaging with each object. In fact, a situation different from this would mean that our visual attention is not well-accustomed to the external surrounding densely occupied by potentially attractive visual objects. The ability to scan through them in order to orient in one's own environment and plan further actions is a necessity (B. W. Tatler & Land, 2011). The inability to maintain cognitive processing efficiency in a situation when unrestricted, prolonged viewing opportunities are available would be a major impediment to our everyday functioning in the visually rich environment. The 'minimal sufficient' comfortable conditions seem to overcome the distraction-related spatial factors, such as *co-visibility*. Once the former is provided, the latter looses on importance. An indirect linkage between the 'divided attention' effect and Recognition Memory established in Experiment 1 can therefore be the result of a correlation between two spatial factors (*VCAs* and *Co-Visibility*) there modified, and the experimental set-up of Experiment 2 shed more light on this relation.

Salience Rating was not a significant predictor in any of the Recognition Memory measures (contrary to Hypothesis C2). The influence of this particular factor was not shown clearly in Experiment 1 and therefore it can be concluded that no clear relation between Salience Rating (as it was defined in this thesis) and Recognition Memory is observable. It must be noted however, that the set of artworks used in these two experiments is relatively unvaried compared to many other art exhibitions. In the art gallery context, the effect therefore might be highly dependent on the particular group of exhibited artworks.

20.3 spatial memory

There was no significant difference in participants' Spatial Memory performance across the conditions. However, once again (following the Experiment 1) it can be noticed that condition which facilitated more 'spotty-focused' visual attention (Cond. 2), resulted in a larger spread of individual performances. This suggests that a space which is not 'optimised' for linear viewing, might present a major challenge to some visitors when retrieving its spatial organisation is concerned. Not only did Condition 2 generate much larger spread in Spatial Memory performance, but also (unlike Cond. 1) it resulted in a significant correlation of this performance measure with *number of long dwells*. The potential reason for such a difference might be that Condition 1 was better optimised for preserving spatial memories of the environment, no matter what the abilities or strategies of the individuals were. In a less-optimal spatial setting, these spatial strategies (such as the tendency for longer dwells, as well as other individual differences) played a much larger role in differentiating the outcome result. The exact nature of this relation, however, lies beyond the scope of this thesis.

Much higher *Co-Visibility* in Condition 2 did not result in better Spatial Memory performance as it was predicted in Hypothesis C3. This particular hypothesis, however, assumed that such an influence would arise from a higher meaning associated with the artworks hanging at highly *co-visible* locations with large *Visibility Catchment Areas*. In this experiment, both of these measures were highly uniform within each condition. In result, a person experiencing only a single gallery set-up (in only one of the two conditions) had no possibility to relate the perceived spatial importance of any of the viewed objects agains other, more or less spatially important ones. Confirming or negating Hypothesis C3 would therefore require a study designed around this particular research question.

A very specific connection between Spatial Memory and eye-movement can be inferred based on the *Miniature Task Sequence Analysis*. The longer participants spent looking at a particular picture, the more likely they were to solve it earlier in the Miniature Task. This linkage follows the prediction made previously in Hypotheses A5 and D4 and the conclusion drawn from the results of Experiment 1: that space affects memory in an *indirect* manner, by influencing the eye-movement patterns first. Thus, the influence of space on visual attention (described in Section 20.1) remains the central part of the analysis revealing in detail how spatial layout modifies our cognitive processes in space. The resulting impact on memory performance seems intuitive: the more visitors look at a given object, the better they remember it.

Based on *Miniature Task Sequence Analysis*, also the linkage between Spatial Memory and Recognition Memory was visible - the quicker a person was in responding to a painting in the Recognition Test, the more likely he or she was to move it earlier in the Miniature Task. This result however might be associated with the most probable strategy of solving the Miniature Task - participants are more likely to firstly focus on solving the location of those pictures, of which presence in the gallery they were the most sure of. Only later they would solve the location of objects, of which presence in the gallery they were not certain.

Salience Rating had no significant impact on any of the Spatial Memory, Recognition Memory, or Eye-Tracking measures, even despite the fact that pictures were hung alternately (highly-lowly-highly salient...). Detailed analyses also revealed no interaction effect with experimental condition. Based on the results of Experiment 1, Hypothesis E3 assumed that the more radical hanging conditions, especially those promoting complete isolation of individual artworks in Condition 1, will result in a significant effect of *Salience Rating*. However, it was not observed in the results. The reason for this might be similar to the one described in one of the above paragraphs - each participant experienced only one spatial set-up, and all artworks within this set-up had very similar spatial characteristics. In Experiment 1, it were highly salient *and* isolated pictures which attracted more longer dwells. Perhaps, when all pictures are isolated, their individual salience looses significance, as a specific viewing mode is undertaken by the visitors from the start of the visit and subsequently applied to all artworks. Moreover, even if the effect found in Experiment 1 was strong enough to prevail in the new experimental conditions, it would nevertheless have negligible impact on memory, as *long dwells* only affected memory performance in Condition 2 - where no artwork was isolated. Experiment 2 therefore demonstrates that if object's salience (defined by *Salience Rating*) is influencing the visitor experience, this is done in a very specific manner (e.g. as demonstrated in Experiment 1) and not invariably on all occasions. It can be assumed that quasi-objective saliency measures, such as *Salience Rating*, cannot therefore reliably describe how a particular object 'stands out' from the others, similar ones. Especially in the art context, individual preference might play a far more significant role in this respect.

20.4 summary

The average *isovist size* was larger in Condition 2 showing that the participants had the possibility of engaging with each picture from more diverse set of angles and distances. Also the *isovist coverage area* was larger, making it possible for the visitors to maintain significantly higher *engagement rates* throughout their visits. This resulted in a modified dynamics of the viewing behaviour across the conditions.

However, these behavioural differences were not reflected in the cognitive outcome of the visit such as *Mean Reaction Times*, showing that the depth of processing occurring in both spatial conditions was similar. This results seems to be linked to equal *Visibility Catchment Areas* - fragments of the gallery floor plan located directly in front of a picture.

The question of the relation between *VCA* and isovist area is the one of the nature of human visual attention and memory. Are our cognitive processes able to make use of additional spatio-visual opportunities for processing? It appears they do not. At least not in a situation where one has the possibility and ease of guiding his or her own eye gaze into the place providing the most comfortable viewing conditions. When such conditions exist, deeper processing from less comfortable locations would perhaps generate unnecessary cognitive effort. This is not equivalent to saying that humans are incapable of deeper art processing from diverse set of angles. Yet, it is an ability that seems not to occur spontaneously and perhaps makes itself visible only under conditions leaving the viewer no other (more cognitively comfortable) alternative. In other words, higher 'quantity' visibility potential does not seem to give an advantage. There is not enough evidence to say it makes the processing worse, but under spontaneous exploration it is not a source of significant improvement. As long as the 'basic spatial conditions' are met, the abundance of available visual information has a strong behavioural, but a limited cognitive effect.

Part VI

BALTIC CASE STUDY

Both of the already described experiments took place in an artificially arranged art gallery. This is potentially problematic as the importance of measuring the impact of art in its original setting has been emphasised in the literature on multiple occasions (Locher, 2011; Tschacher et al., 2012; Brieber et al., 2014). The main reason is that laboratory-based studies cannot account for the artist's intentions regarding the presentation of the work, and have no curatorial background to facilitate the set-up of a 'mock-up' art gallery. In result, an experimental gallery is unlikely to be very similar to a 'real' art gallery. Even if they look alike, the ideas underlying spatial relationships in an experimental gallery are different (often random), and on the level of the cognitive outcome they might not simulate the 'real' art gallery experience well (Locher, 2011). Moreover, it has been long acknowledged that a museum visit should not be analysed in separation to the visitors' prior expectations, as they play a major role in shaping the final experience (Falk & Dierking, 2004). A visit to a psychological laboratory (or even an art gallery located within a university campus - similarly to Experiment 1 described in this thesis) is likely to bear very different expectations from a visit to an actual art gallery. The latter bears strong linkages with its marketing image, previously publicised agenda, and often a prestigious location. The resulting motivation to attend the exhibits (Bitgood, 2003) can also differ. Hence, the importance of testing the experimentally-generated findings in the real world setting is undeniable.

21

METHOD

In this Case Study, participants were invited to explore a working art gallery open to the public in the city of Gateshead, UK. BALTIC Centre for Contemporary Art opened in 2002 in a refurbished flour mill. It hosts numerous temporary exhibitions of contemporary art on its 6 levels, as well as a gift shop, library, a cafe, restaurant, a conference floor and a riverside terrace for special events. Spatial layout of exhibition spaces can be modified according to the requirements of individual exhibition with the use of movable walls and ceilings. The Case Study conducted within this thesis focused on the exhibition of a contemporary German artist Thomas Scheibitz titled 'ONE-Time Pad', which was hosted on two floors of the BALTIC between 26.07.2013 and 3.11.2013. The study took place on weekdays between the 7th and 25th of October and considered only one out of two exhibition levels.

21.1 PROCEDURE AND PARTICIPANTS

Individual meetings with participants were scheduled in a gallery location assigned for each day by the gallery's Duty Manager. This locations were either a dedicated empty meeting room, the gallery's library, an empty riverside terrace, or a corner table in the cafeteria during off-peak hours[1].

Within the procedure employed in the two previously described studies, participants were asked to wear Tobii Glasses eye-tracker and explore Thomas Scheibitz's 'ONE-Time Pad' exhibition. The exhibition was hosted on 2 separate levels, but participants were specifically asked to visit BALTIC's Level 3 only, as it has consisted majorly of paintings. Following the calibration of the eye-tracker, participants were explained the route to the exhibition from the briefing point and navigated alone to the exhibition and back. When the meeting point was located on a level different from Level 3, they were free to choose between using the staircase and the lift. The researcher did not follow the visitor to the exhibition area and waited in the briefing room for the entire period of the visit. It must be therefore noted that

1 Despite the variability in lighting, sitting conditions, and ambient noise during the briefing and de-briefing phases, subsequent data analysis showed no significant influence of the room type on participants' results.

participants spent different, uncontrolled time between seeing the exhibition and participating in the memory tests, as they took different paths to come back to the starting position. One participant accidentally visited both gallery levels hosting the exhibition. Time limit of 40 minutes was imposed - given the vast size of the exhibition as long a period as possible was desired, although (similarly to Experiments 1 and 2) participants were free to leave as quickly as they preferred.

The museum staff was aware of the study taking place but there were no specific instructions given to facilitate their reaction to the participants. Thus, some participants engaged in a conversation with the staff members, while others joined a guided gallery tour group. Informal observation of the eye-tracking recordings did not reveal any unusual reactions from other gallery visitors towards the presence of visitors wearing an eye-tracking device. These unrestricted study conditions were purposefully arranged to simulate a 'real-life' art gallery visit as closely as possible, with all its potential variability of distractors.

Sixteen participants (9 female; all over 18 years old, non-artists, non-architects, with normal or corrected-to-normal vision, not colour blind) were recruited through university email system for a fee of 6 pounds. None of the participants has previously visited the studied exhibition. As spatial layout of the gallery was modified for its purpose, previous familiarity with the building played no role, even though 9 of them have been to the BALTIC building before. Participants were aged between 20 and 46 years (M = 28.5, SD = 8.24). All participants explored the exhibition under the same spatial condition, unmodified for the entire duration of the exhibition.

21.2 SPACE AND MATERIALS

The layout of the exhibition has been designed by the artist in collaboration with the gallery using a miniature maquette (BALTIC Centre for Contemporary Art, 2013). This fact differs the Case Study from both of the previously described Experiments, where the placement of artworks was established with no curatorial goals other than the research aims specified by the researcher. Figure 21.1 presents the layout of the Case Study exhibition.

Fully 3-dimensional sculptures are not analysed in this Case Study, since—compared to paintings—they display unique visual information from different angles and therefore are likely to encourage distinct viewing patterns. However, they are considered in Space Syntax analyses, as their presence modified visual fields of the individuals. The room located the furthest from the entrance of the gallery was a rectangular space containing multiple glass cabinets with small miscellaneous objects and a number of densely hung sketches on the surrounding walls. Due to the available precision of the eye-

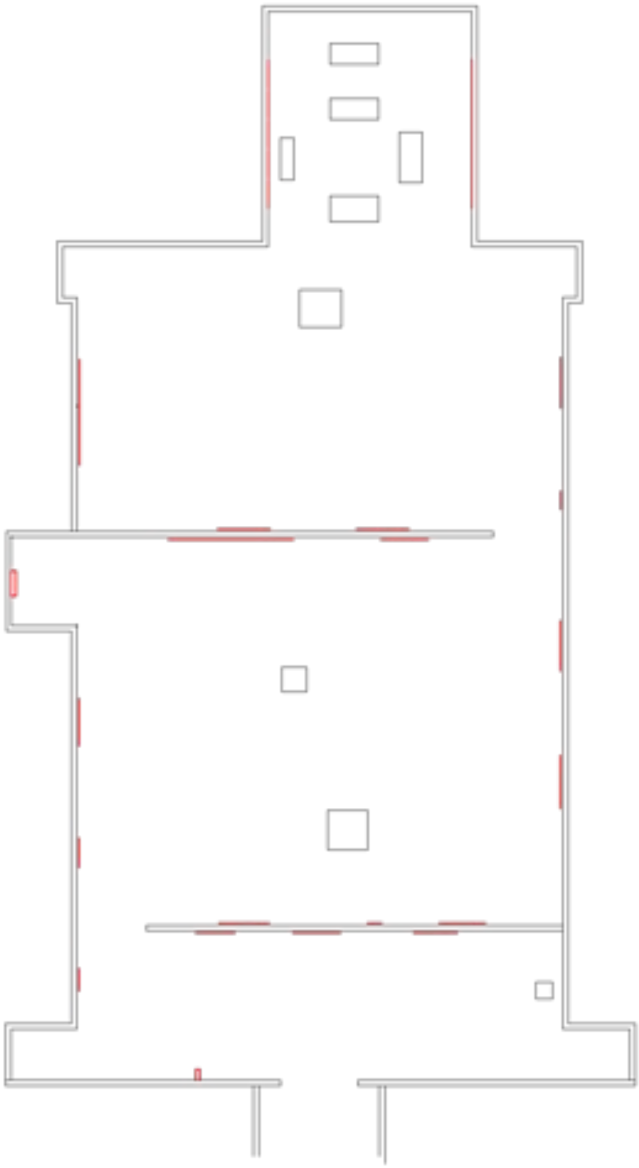


Figure 21.1.: Layout of the exhibition studied in the BALTIC. Except for the upper subspace, the exhibition floor contained 24 distinct artworks, 19 of which were paintings (or installations framed as paintings and hung on or supported by the wall); 5 were other sculptures. In its widest and longest space, the area was 19 x 38.5 m respectively.

Figure 21.2.: An overview of Thomas Scheibitz's 'One-Time Pad'. Image courtesy of BALTIC Centre for Contemporary Arts.

tracking device, all objects located in this room were also ignored in subsequent analyses, despite the fact participants spent noticeable amount of time in it. These objects were uniformly accounted for as 'other artworks' in the eye-tracking coding procedure and were not presented to the participants in the subsequent memory tests.

In total, 19 pictures, or flat sculptures were the subjects of eye-tracking coding and all subsequent analyses. Five 3-dimensional sculptures and all objects from the room containing glass cabinets were ignored. Figures 21.2, 21.3, 21.4, 21.5 and 21.6 present overview of the gallery and the contained artworks.

As in any other contemporary art gallery, labels and textual information about the author and the artworks were available. Participants received no instruction regarding these information and were free to explore them at their own discretion.

Other visitors were present in the gallery, in varying numbers depending on the time of the day. Conducting the study only during weekdays ensured off-peak activity in the gallery. Participants received no instructions regarding the engagement with other visitors and gallery staff.

For this Case Study, no independent Salience Study was conducted. As the paintings hanging in the gallery varied in size (from 3327 cm2 to 126000 cm2), assessing their salience in an isolated computer-based experiment would not provide a valid estimation of their factual 'salience'. Even if the objects were presented on a screen downsized to their relative scale, it is unlikely that the magnitude of the perceived difference in salience would accurately simulate the difference

Figure 21.3.: An overview of Thomas Scheibitz's 'One-Time Pad'. Image courtesy of BALTIC Centre for Contemporary Arts.

Figure 21.4.: An overview of Thomas Scheibitz's 'One-Time Pad'. Image courtesy of BALTIC Centre for Contemporary Arts.

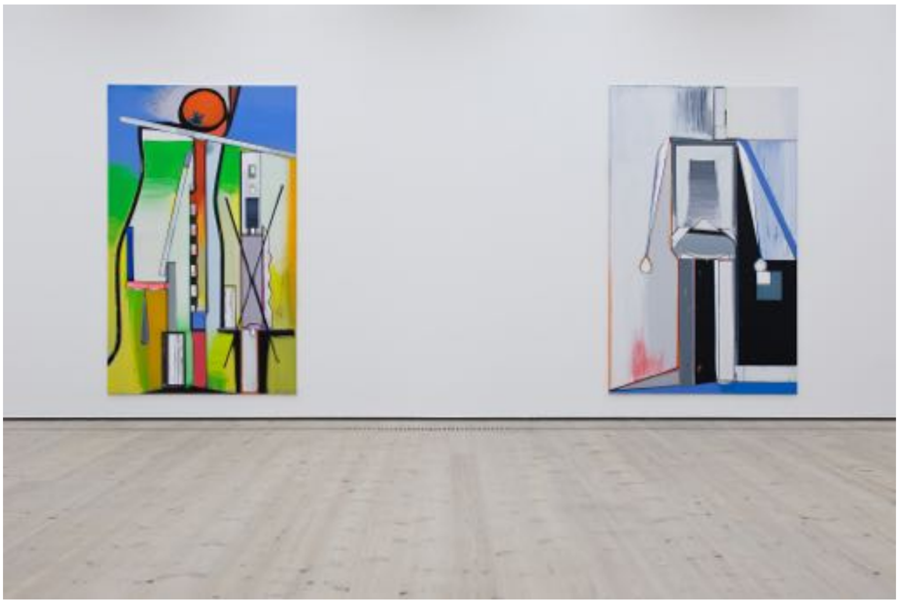


Figure 21.5.: An overview of Thomas Scheibitz's 'One-Time Pad'. Image courtesy of BALTIC Centre for Contemporary Arts.

Figure 21.6.: The back-room glass display cabinet at Thomas Scheibitz's 'One-Time Pad'. Image courtesy of BALTIC Centre for Contemporary Arts.

encountered in the real art gallery. For this reason, *size of the object* was used in further analyses instead (defined in centimetres-squared surface area).

Similarly to Experiment 2, lighting of each artwork was measured. Digital lighting sensor was used to obtain LUX measurements twice for each artwork (from about 10 cm from the its centre), which were then averaged. The lighting appeared to be uniformly distributed (M = 187 lx; SD = 13.5) and therefore was not included in subsequent analyses.

21.3 EYE-TRACKING RECORDINGS

'Tobii Glasses 1' eye-tracker was used in the experiment. The calibration was successful (and recorded above 60% of valid fixations) for 9 participants, whose recordings were extracted and coded according to the procedure described in Appendix A.1. The coding procedure considered instances of each picture dwell, as well as (jointly) dwells on walls, floors and the ceiling. Additionally to Experiment 1 and 2 dwells falling on other visitors were also coded in order to control for the level of potential distraction. Fixations on artworks not being the subject of further analyses (such as sculptures and glass display cabinets) were coded jointly under a separate code.

Due to the presence of a large number of stimuli remaining outside the scope of the main analysis, it is important to clarify how eye-tracking measures were calculated in the BALTIC study compared to the two previously described Experiments. Unless stated otherwise, all eye movement metrics further described relate to calculations which *excluded* fixations on non-analysed objects. For instance, where mean *Dwell Lengths*, *Number of Dwells*, or *Long-to-Short Dwell ratio* are reported, they all exclude the fixations at non-analysed objects. This is justified by the fact that viewing behaviour occurring within the back end space of the gallery (which contained large display cabinets of multiple small objects) most often attracted relatively long undisturbed viewing periods, rarely observable in other parts of the exhibition. Due to spatial isolation of these works, such interactions were relatively detached from the general 'viewing flow' of the main paintings. Including these fixations in the calculation would have a high impact on the distribution of the eye movement metrics, potentially decreasing their value as quantifications of the visual attention for works being the subject of the study. The only measure which considered fixations on non-analysed objects—as well as on other visitors—was *Picture-Switching ratio*. This metric would classify an immediate fixation after an interaction with a picture, even if the object being the subject of the subsequent fixation was a sculpture or a display cabinet in the back end room. Most importantly, the reader should note that *Engagement Ratio* is hereby equivalent to the percentage of total *Time Spent Inside* by the visitor, which was spent looking at any of the artworks being the subject of the analysis. It thus excludes the time spent looking at sculptures and back end room objects. To account for the potential influence of these engagements, *Normalised Total Dwell Time* (i.e. the percentage spent fixating) on non-analysed objects was also quantified and included in the detailed statistical analysis.

Those participants whose eye-tracking data could not be used were unaware of the equipment failure and therefore still performed all tasks in the same context as other participants. For this reason, the data obtained from their memory test solutions was used in all analyses where it did not require matching with eye-tracking data.

21.4 RECOGNITION MEMORY TEST (REACTION TIMES)

The Recognition Memory Test was presented via OpenSesame software (Mathôt et al., 2012) on a 17″ laptop with two keys ('A' and 'L') labeled 'YES' and 'NO' by black-and-white printed stickers. The entire task was identical to the one used in Experiment 1 and 2 except the stimuli used were different. The task consisted of 3 training stimuli, followed by a randomly ordered projection of 19 correct artworks and 19 new objects (Appendix E contains all stimuli used in this task). The keyboard keys used for providing the response (left-right; yes-no) were swapped after participant number 8.

21.5 SPATIAL MEMORY TEST (MINIATURE TASK)

The Miniature Task was similar to the one used in Experiment 2 and it was presented on the same laptop used for the Recognition Memory Test. An empty PowerPoint slide was displayed in the full-screen mode, only with a single toolbar and the menubar visible. The layout of the gallery was displayed as a non-clickable, non-movable slide background. The layout contained the locations of all walls and of the entrance in black. It did not contain the pre-specified picture locations. Nineteen Picture Miniatures were randomly scattered outside the layout. Figure 21.7 a sample starting set up of the task, and Figure 21.8 its sample solution.

The miniatures were manipulated by 'drag and drop' action performed with a mouse. Participants were instructed to take as much time as they need to complete the task, but it was suggested that they place all pictures on their presumed location (Appendix D.4). The entire solution was recorded with CamStudio 2.7 software. All participants produced valid solutions. If ambiguities in miniature placement were present, the researcher asked to specify which exact wall location is considered as correct and to adjust the position of the miniature.

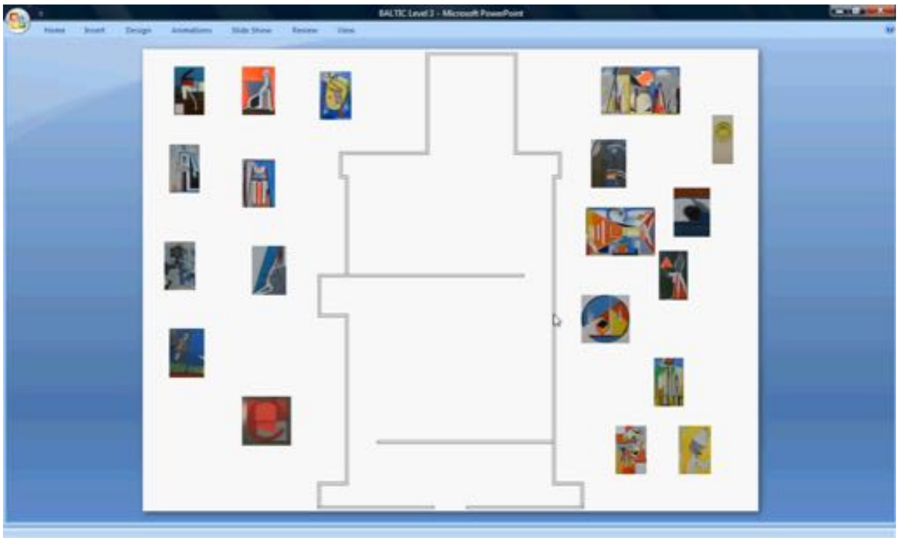


Figure 21.7.: Sample set up at the beginning of the Miniature Task.

Figure 21.8.: Sample solution of the Miniature Task.

21.6 REVISED HYPOTHESES (F)

The fact that each picture was 'fixed' to its location permanently and that all participants viewed the same works hanging at the same positions limits the possibility of distinguishing the influence of space from the influence of individual pictures. However, this limitation is primarily linked to the random-effect component of the prediction: compared to the analyses described in Experiments 1 and 2, here, the by-picture and by-location random effect cannot be separated. We will use the term 'by-item' random-effect to refer to it jointly. Nevertheless, using the knowledge combined from the analyses of two earlier studies, we can expect spatial factors to have a strong influence on visitors' visual attention and memory. Primarily *VCA*, *Targeted Co-Visibility*, and *Isovist Jaggedness* have been shown to affect eye movement (note that in Experiment 2 the effect of Co-Visibility has often been described by the influence of Experimental Condition, which by design differed primarily by this particular measure). In the layout of the BALTIC exhibition, the difference between individual locations' *Isovist Jaggedness* were minimal, and thus are expected to have no significant impact on psychological measures (Hypothesis F1).

Since no independent Salience Study was conducted, painting's size was the main variable describing content properties of individual pictures. Informal observations of individual visitors additionally suggest that this might have a linear impact, as larger pictures may attract more attention (Hypothesis F2).

As this study had no exploratory character, but was a real-life testing setting for the findings previously established in Experiments 1 and 2, the procedure of model construction was modified. For each outcome variable, first an attempt was made to build a model only from the data potentially available to the curator at the stage of exhibition design (i.e. without the demographic data and information about the time spent inside by each visitor) and only using predictors which have proven to be significant in the analysis performed in Experiments 1 and 2. Picture's size will also be used by default, as it has sound theoretical reasons to have an impact on both visual attention and memory. We will call these models 'practical' predictions. Only then the stepwise model selection procedure will be repeated, as it was employed previously to establish the actual best-fit model for the gathered data or the significance of individual factors verified. It is almost certain that the best-fit models will in all cases outperform 'practical' predictions, but the aim of this research exercise—as well as the reason standing behind the entire BALTIC Case Study—is to discuss the practical implications of the findings presented in the thesis. Such 'practical' prediction formulas could be potentially used to build software aiding the exhibition design process. Note, that a committed curator or artist could make further attempts to improve the fit of such simple models by modelling the distribution of visitors' age and of their times spent inside the gallery based on the records of visitors' activities held by the hosting museum institution. The spatial factors have shown to have a limited influence when considered in separation to these basic demographic aspects, especially for some of the considered outcome variables. Thus, these models and their R^2 values will not rightfully represent what 'the effect of space alone' is, but should only be treated as coarse estimations given the data potentially available to a curator.

22

DATA ANALYSIS

22.1 SPATIAL CHARACTERISTIC OF THE GALLERY (SPACE SYNTAX AND ISOVIST ANALYSIS)

Measures described in Section 11 were calculated for each picture location using Depthmap (Turner, 2001). Figure 22.1 below presents a sample visualisation). Each layout location was linked with a numeric measure, describing its relation to all other artwork locations in the layout.

The arrangement of artworks was specific in one particular respect: there was a small group of noticeably isolated paintings (low *VCAs* and low *Co-Visibility*) and a much larger group (10 out of 19 analysed artworks) which had both *Potential Co-Visibility* and *VCAs* high above the mean. This relation was not true for pictures of lower *Co-Visibility*, as their *VCAs* varied across the spectrum. Figure 22.2 presents these relations.

Figure 22.2 also demonstrates what can be considered a 'natural spatial order' of this (and possibly many others) art exhibition. Groups of paintings which have high *Co-Visibility* and low *VCAs* (or vice-versa) are underrepresented. Exposure is either limited entirely, or maximised with the use of both spatial factors. This seems to be the case for the majority of commonly encountered modern art exhibitions and could possibly be linked with the need of optimising floor surface usage. Any artwork can be isolated completely within a relatively small gallery sub-space, but maximising its *VCA* without increasing *Co-Visibility* would require leaving a lot of gallery space unused. On the contrary, maximising the painting's *Co-Visibility* while keeping its *VCA* small, would require hanging multiple paintings on the walls surrounding a very restricted sub-space, which can be unpractical (for instance due to potential crowding). This observation requires to differentiate BALTIC Case Study from the two artificially-arranged experiments, where *Co-Visibility* and *VCAs* were not highly correlated. Perhaps, by doing so, they did not represent a setting which would be preferred if organised by a curator. And yet, most certainly museum institutions differ in the spatial strategies employed (Tzortzi, 2007) and thus no far-reaching statement can be drawn from an observation made in the single BALTIC Case Study. In respect to the measures considered by the current thesis, however, it is an

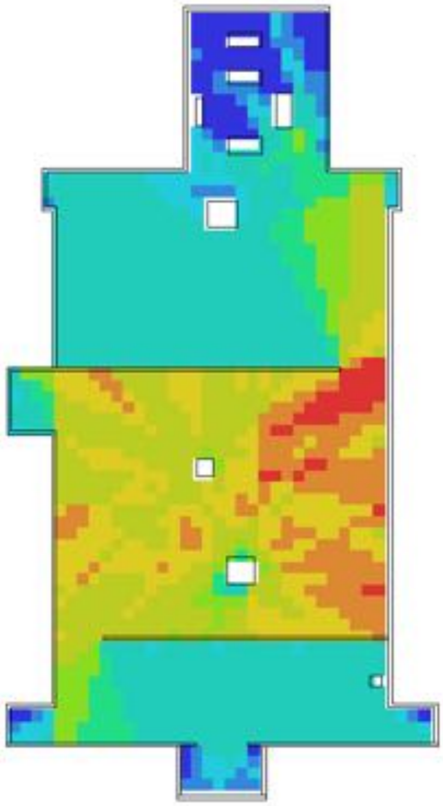


Figure 22.1.: Visibility Graph Analysis of the BALTIC Case Study layout.

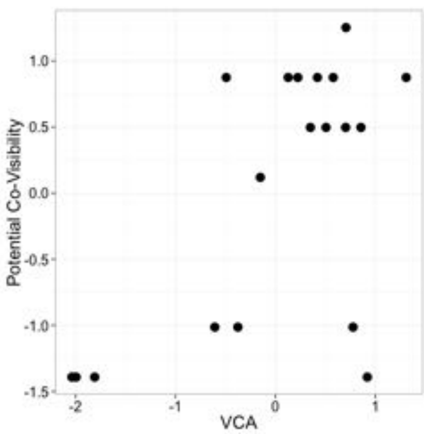


Figure 22.2.: Relation between Potential Co-Visibility and VCA with a visible gap in the middle part of the spectrum.

important feature. Spatial analysis of the setting revealed, that *Potential Co-Visibility* and *VCA* variables correlated at the level of $\rho = 0.4$. In this case, including both measures in a single statistical model is legitimate, given collinearity of the entire model is not violated[1].

Potential Co-Visibility was calculated according to the same procedure as in Experiment 1. The distribution of this variable was clearly binomial, with 12 out of 19 pictures having *Potential Co-Visibility* equal to 7 or higher and 10 was the maximum value in this particular spatial set-up. The remaining 7 pictures had *Potential Co-Visibility* of either 3 or 4. For this reason, the variable was coded as a categorical factor, distinguishing between 'low' and 'high' *Potential Co-Visibility*. This will make it possible to investigate interaction effects, i.e. whether the influence of individual factors on human cognition was different for these pictures with 'low' and 'high' co-visibility. Due to the fact that almost all highly *Co-Visible* pictures had relatively large *VCAs*, the interaction effect of *Potential Co-Visibility* and *VCA* will not be considered.

Additionally, pictures were classified for their *type*, differentiating between a traditional canvas painting, a flat sculpture hanging as-if-it-was-a-painting, or a flat artwork framed behind a glossy glass surface.

22.2 EYE-TRACKING MEASURES

Eye-tracking measures described in Section 8.2 were obtained from the coding logs, using a custom-built R script.

The recordings were also used to calculate total *time spent inside* the gallery by each participant (from the moment of the first dwell through the glass entrance to the gallery to the moment of passing the door when exiting). Where eye-tracking recording was unsuccessful, the researcher's manual timing estimation was used to assess *time spent inside*.

Out of 16 eye-movement recordings, only 9 were valid (i.e. with correct calibration and containing minimum 60% of valid fixations). Only this subset is considered in the analysis of eye-tracking recordings.

22.3 RECOGNITION MEMORY TEST (REACTION TIMES)

For Reaction Time data analysis, only the correct 'yes' answers were taken into consideration (90% of the entire dataset). However, the accuracy was also recorded and will be referred to as *RT accuracy*. Observations lying below 250 ms or further than two standard deviations from the mean of the whole dataset were removed (Ratcliff,

1 A recommendation given by (Tabachnick & Fidell, 2001) is to avoid including in a single model two variables which correlate at the value of 0.7 or higher.

1993; Whelan, 2008), leaving 84% of all responses valid (i.e. correct and within two Standard Deviations).

22.4 SPATIAL MEMORY TEST (MINIATURE TASK)

From all three measures available (Section 10.1), String Matching technique did not seem suitable for the analysis of this Miniature Task due to the presence of uninterrupted wall surfaces spreading over two large rooms. Classifying it as a holder of a single string would thus be self-contradictory. This, similarly to Experiment 1, shows that this method is unsuitable for non-trivial spatial environments, where objects of interest can be located at diverse set of angles in relation to each other.

Bidimensional Regression was also shown to be unreliable in spatial conditions providing higher chance of 'wrong-side-of-the-right-wall' error type (Section 14.5). The layout of the gallery analysed here provided multiple possibilities for such mistakes.

For this reason, less detailed but more robust *Back-to-the-Wall* score was again used to assess participants' spatial memory performance.

22.5 MINIATURE TASK SEQUENCE ANALYSIS

Similarly to the procedure described in Section 18.6, screen recordings of participants' 'drag and drop' actions were recorded while they were solving the Miniature Task. Each time a miniature was moved from an already taken position to a new one, its order in the sequence was changed to the last. All movement which was not affecting the sequence of the artworks along a given wall was ignored (e.g. when participant needed to adjust the positioning to make space for other artworks). Rearranging artworks within the same wall was also ignored (as the exact sequence is ignored by *Back-to-the-Wall* measure).

Two participants left out two miniatures outside the layout. In each case those two miniatures were added to the end of the sequence (at random order) and considered as false answers in *Back-to-the-Wall* measure.

There was no correlation between *ET sequence* and *Miniature Task* sequence (ρ = 0.05).

23

RESULTS

As the reader could have noticed, Experiment 2 yielded generally higher Conditional R^2 values in the case of all models in comparison to their equivalents in Experiment 1. This phenomenon might be the result of the feature which has already been named as a potential limitation of the current study - its restricted ecological validity. The space arranged in Experiment 2 was simpler, and as a consequence it afforded much more uniform type of behaviour (demonstrated e.g. by the measure of *viewing sequence similarity*). In a more restrictive laboratory-based situation, it is easier to control for factors potentially affecting the measured psychological processes. Consequently, it is expected that the next study—taking place during working hours of a public centre for contemporary arts—will allow the author to explain much lower proportion of the variance. A space deprived of any distractions, designed purely for the purpose of an experiment increases the chances of a researcher for observing repeatable patterns of human behaviour. When the research site contains noticeable amount of external 'noise' (i.e. factors not being the subject of the study and staying beyond the control of the researcher), the fit of any statistical model will decrease.

23.1 TIME SPENT INSIDE

Participants spent M = 1388 seconds (or about 23 minutes; SD = 596 sec.) inside the gallery. One person has used the entire time limit allowed for the exploration (40 minutes). The person who has been inside for the shortest period of time, left the gallery within 4.5 minutes of entering. As the available eye-tracking recordings revealed, at least one person have not complied to the instruction and explored an additional floor of the exhibition, although briefly.

23.2 VISUAL ATTENTION

Out of 9 participants whose eye-tracking data was analysed, all participants fixated at least once at every picture with a single exception of one person missing the work titled 'GP 171'. Informal observation of the eye-tracking recordings revealed that the painting appeared in the

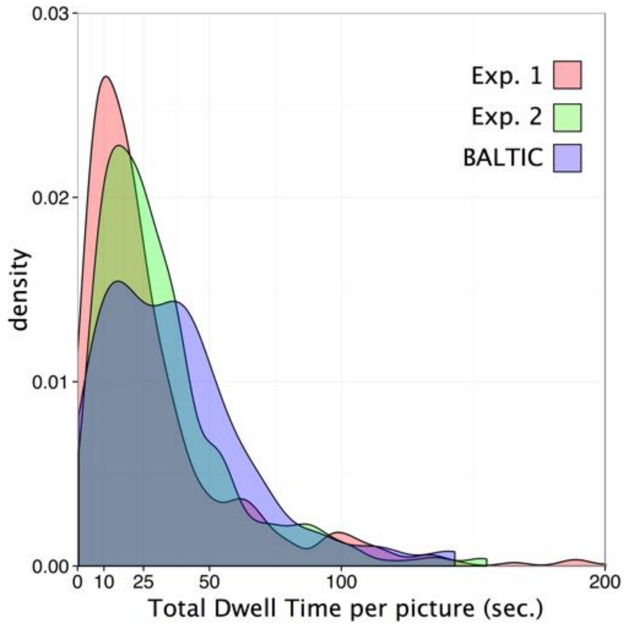


Figure 23.1.: Visualisation of *Total Dwell Time per picture* distributions across the 3 studies.

participant's periphery field of view on multiple occasions, therefore the visitor's data was not excluded from the analysis.

Participants performed 202 dwells on the analysed paintings on average during a single visit and the proportion of the total time inside spent looking at those pictures (or average *Engagement Ratio*) was 0.53 (SD = .07). Much lower *Engagement Ratio* compared to Experiments 1 and 2 comes from the fact that it does not account for the time participants spent fixating on sculptures, installations, and pictures excluded from the analysis. These instances account for 31% of all dwell times on average. Visitors have also spent an average of M = 2.86 sec. (SD = 2.34) per participant fixating on other people present inside the gallery.

Average *Total Dwell Time* per picture was M = 36.91 sec (SD = 28.72), however, due to the large variation in these values, the mean was not significantly higher from this obtained in Experiment 1 (Wilcoxon signed rank test against 31.71: V = 7948, p=.359). As presented in Figure 23.1, Total Dwell Time distributions were largely similar between Experiments 1 and 2 compared to the BALTIC Case Study, where some pictures attracted considerably longer viewing.

Average participant fixated on the first picture within 1.64 sec. from entering the gallery area[1], most likely due to the fact that some artworks were visible through the glass door and such fixations were coded on an equal basis. Each picture was a subject of M = 10.62

1 The gallery area for the purpose of the analysis is defined as containing the rectangular area in front of the glass door entrance to the gallery.

dwells during a single gallery visit. Mean *Long-to-Short Dwell Ratio* was 0.24, meaning that an average participant performed about three times as many *short dwells* (i.e. dwells lasting for less than 2 sec.) than *long dwells* (>2 sec.). As indicated by *Picture-Switching* metric equal to 0.49, almost half of all dwells were followed by an immediate fixation on another picture (or sculpture, or installation).

Mean *dwell length* was 3.48 sec., the cumulative distribution of dwell lengths was heavy-tailed (in line with Hypothesis A1), and the longest single dwell on a separate picture lasted for 88 seconds.

Since artworks were fixed to a specific spatial location, the by-picture and by-location descriptors are equivalent. Similarly to the analyses performed for Experiments 1 and 2, *Number of Dwells* and *Total Dwell Time* were modelled as two key variables describing the quantitative aspect of the visitors' eye movement.

Firstly, however, we will make an attempt to predict *Number of Dwells* falling onto each picture as it could be done by an interested exhibition designer based on the spatial factors only, without considering demographic data unavailable to a museum prior to the launch of the art show. These predictions will be called 'practical', as they could potentially deliver a simple, straightforward estimation of the cognitive impact spatial layout might have on the gallery visitors. The following 'practical' version of the model was proposed for *Number of Dwells* variable[2]:

```
log(No. of Dwells) ~ VCA + Iso. Jaggedness +
Potential CoVis. +  size(cm2) + (1|loc) + (1|ids)
```

Model 23.1: Logarithm of Number of Dwells ('practical')

Normality and homoscedasticity of residuals of the final model were not violated, although a single residual outlier was present. *Kappa* score was 2.6, *Marginal* R^2 was 0.05 and *Conditional* R^2 was 0.48. Table 23.1 describes the model in detail. Only *size* demonstrated a significant fixed effect.

To determine the best-fit model, the following predictors were entered:

```
log(No. of Dwells) ~ Potential CoVis. + VCA +
Iso. Jaggedness + size + type + LUX +
View. Sequence + Time Spent Inside +
age + gender + (1|ids) + (1|loc) +
Potential CoVis.(factorial)*(Iso. Jaggedness +
```

2 Note that the random-effect of by-participant variability was kept even in the 'practical' prediction. Running a model which is not 'aware' of the by-participant similarities in the data at hand would violate the statistical assumption of the test. Simultaneously, a random-effect factor does not contribute to the model's Marginal R^2 value, thus allowing us to establish the effect size being of interest. It represents the fact, that the curator would not know the exact by-participant differences, but can expect the dwell numbers to be correlated within subjects.

	Model 23.1
(Intercept)	2.34 (0.11)***
VCA	−0.06 (0.06)
Iso. Jaggedness	0.08 (0.06)
Potential CoVis.	0.01 (0.06)
size	0.11 (0.05)*
AIC	209.66
BIC	234.79
Log Likelihood	-96.83
Num. obs.	171
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.02
Variance: ids.(Intercept)	0.09
Variance: Residual	0.13

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.1.: Model 23.1.

```
size + LUX + Time Spent Inside + age)
```

Model 23.2: Logarithm of Number of Dwells (preliminary formula)

subject to the stepwise model selection, the following formula was established as the one best-predicting the logarithmic (+1) transformation of *Number of Dwells*:

```
log(No. of Dwells) ~ size + View. Sequence +
Time Spent Inside +gender + (1|ids) + (1|loc) +
Potential CoVis.(factorial)*age
```

Model 23.3: Logarithm of Number of Dwells (preliminary formula)

However, despite the significant impact of *VCA* adding *Isovist Area* as a predictor improved the general fit of the model (indicated by a lower AIC value). Lastly, accounting for the percentage of the time inside the gallery that each visitor spent fixating on other people further increased the fit of the model[3]. This resulted in the following best-fit formula:

```
log(No. of Dwells) ~ Isovist.Area + size +
View. Sequence + Time Spent Inside +
gender + Norm. Time on People +
(1|ids) + (1|loc) + Potential CoVis.(factorial)*age
```

Model 23.4: Logarithm of Number of Dwells

3 *Normalised Time on People* variable is calculated based on the cumulative number of seconds one has spent fixating on other people divided by the total time spent inside the gallery by this participant.

	Model 23.4
(Intercept)	2.53 (0.10)***
Isovist.Area	0.16 (0.07)*
Potential CoVis.(factorial)low	0.17 (0.13)
View. Sequence	−0.02 (0.01)**
Time Spent Inside	0.26 (0.03)***
gender-male	0.17 (0.08)*
Norm. Time on People	−0.03 (0.02)*
age	−0.04 (0.04)
Potential CoVis.f-low:age	−0.12 (0.05)*
AIC	170.83
BIC	208.46
Log Likelihood	-73.41
Num. obs.	170
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.02
Variance: ids.(Intercept)	0.00
Variance: Residual	0.10

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.2.: Model 23.4.

Distribution of residuals was normal and homoscedastic. *Kappa* score was 5.4, *Marginal* R^2 was 0.50. *Conditional* R^2 was 0.54, meaning that very little by-subject and by-location random variance remained unexplained. Similar Marginal and Conditional R^2 values suggest that the fixed-effect factors would perform similarly well in a simple linear model, without considering the random component. Table 23.2 describes the model in detail and Figure 23.2 visualises the effects.

For the 'practical' prediction of *Total Dwell Time*, the following model was proposed based on the results of the previous studies:

```
log(Total Dwell Time) ~ VCA + Iso. Jaggedness+
Potential CoVis. + size + (1|loc) + (1|ids)
```

Model 23.5: Logarithm of Total Dwell Time ('practical')

It had a normal and homoscedastic distribution of residuals, *Kappa* value of 2.6, *Marginal* R^2 of 0.11 and *Conditional* R^2 equal to 0.54. Table 23.3 describes the model. *Potential Co-Visibility* was the only non-significant fixed effect.

The best-fit model established for logarithmic (+1) transformation of *Total Dwell Time* was:

```
log(Total Dwell Time) ~ VCA + View. Sequence +
Iso. Jaggedness + size(cm2) + Time Spent Inside +
```

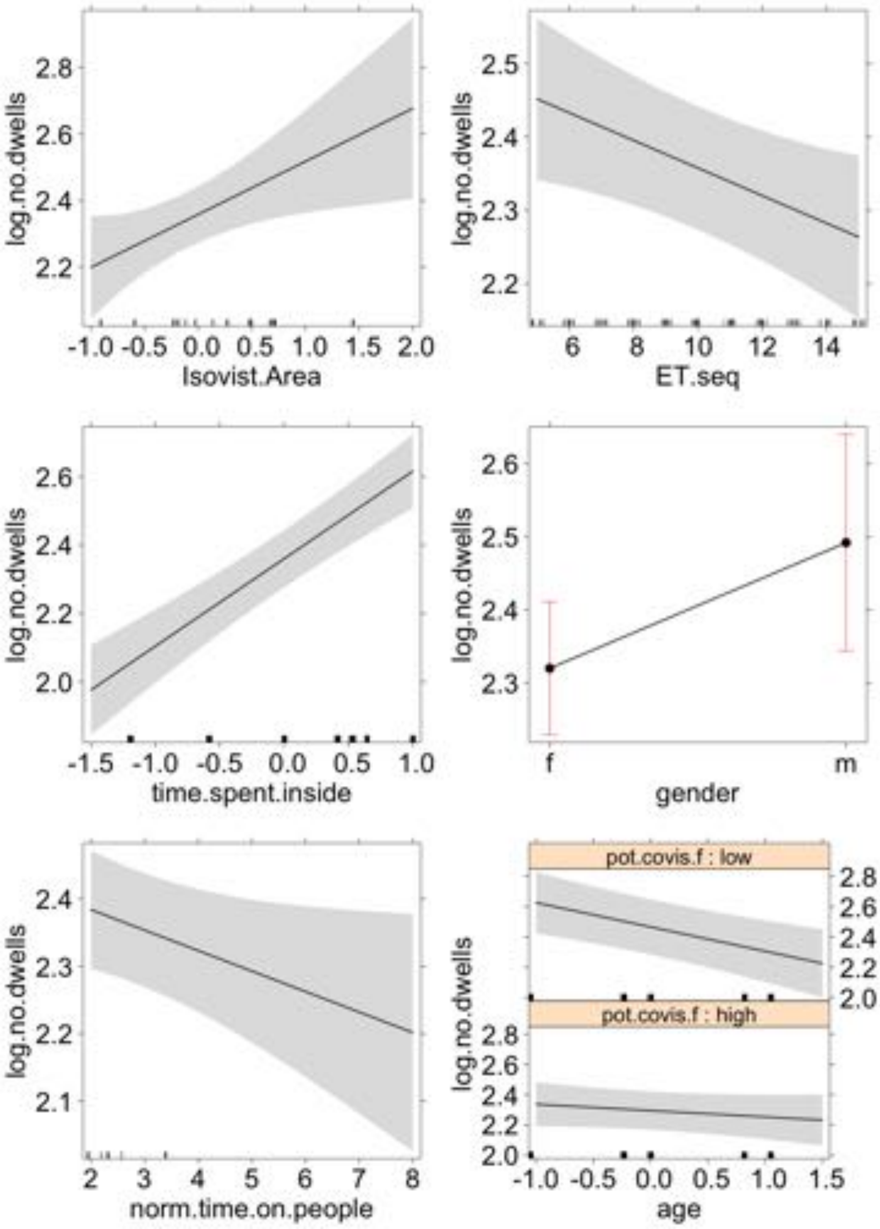


Figure 23.2.: Visualisation of the effects for Model 23.4.

	Model 23.5
(Intercept)	$3.30\ (0.21)^{***}$
VCA	$-0.20\ (0.08)^{*}$
Iso. Jaggedness	$0.21\ (0.08)^{**}$
Potential CoVis.	$-0.04\ (0.08)$
size	$0.32\ (0.06)^{***}$
AIC	392.95
BIC	418.08
Log Likelihood	-188.47
Num. obs.	171
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.02
Variance: ids.(Intercept)	0.37
Variance: Residual	0.41

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.3.: Model 23.5.

```
(1 | ids) + Potential CoVis.(factorial)*age
```

Model 23.6: Logarithm of Total Dwell Time

Homoscedasticity assumption was not violated, although it was heavy-tailed due to 2 outlying data points[4]. The effect size of *ET Sequence* was negligible but excluding it would significantly decrease the overall fit of the model. Substituting *VCA* with *Isovist Area* did not significantly improve the fit of the model. *Kappa* score was 5.1, Marginal R^2 was 0.53 and Conditional R^2 was 0.59. Considering fixations on other people did not improve the fit. Consult Table 23.4 for the statistics and Figure 23.3 for visualisation of the significant effects.

Just like in Experiments 1 and 2, *Picture-Switching* and *Long-to-Short Dwell Ratio* were analysed as key variables describing the dynamics of eye movement. Firstly an attempt was made to make a 'practical' prediction of the *Picture-Switching ratio*. Based on the previous findings, the following model was proposed:

```
Picture-Switching ~ Potential CoVis. + VCA +
(1|ids) + (1|loc)
```

Model 23.7: Picture-Switching ratio ('practical')

Residuals' distribution was homoscedastic. *Kappa* score was 2.1, Marginal R^2 was 0.1 and Conditional R^2 was 0.51. Both fixed-effect factors were significant or approaching significance. Table 23.5 presents detailed statistics.

The best-fit model for the data available was:

4 Linear mixed-effect models have been shown to be robust under this violation (Winter, 2013).

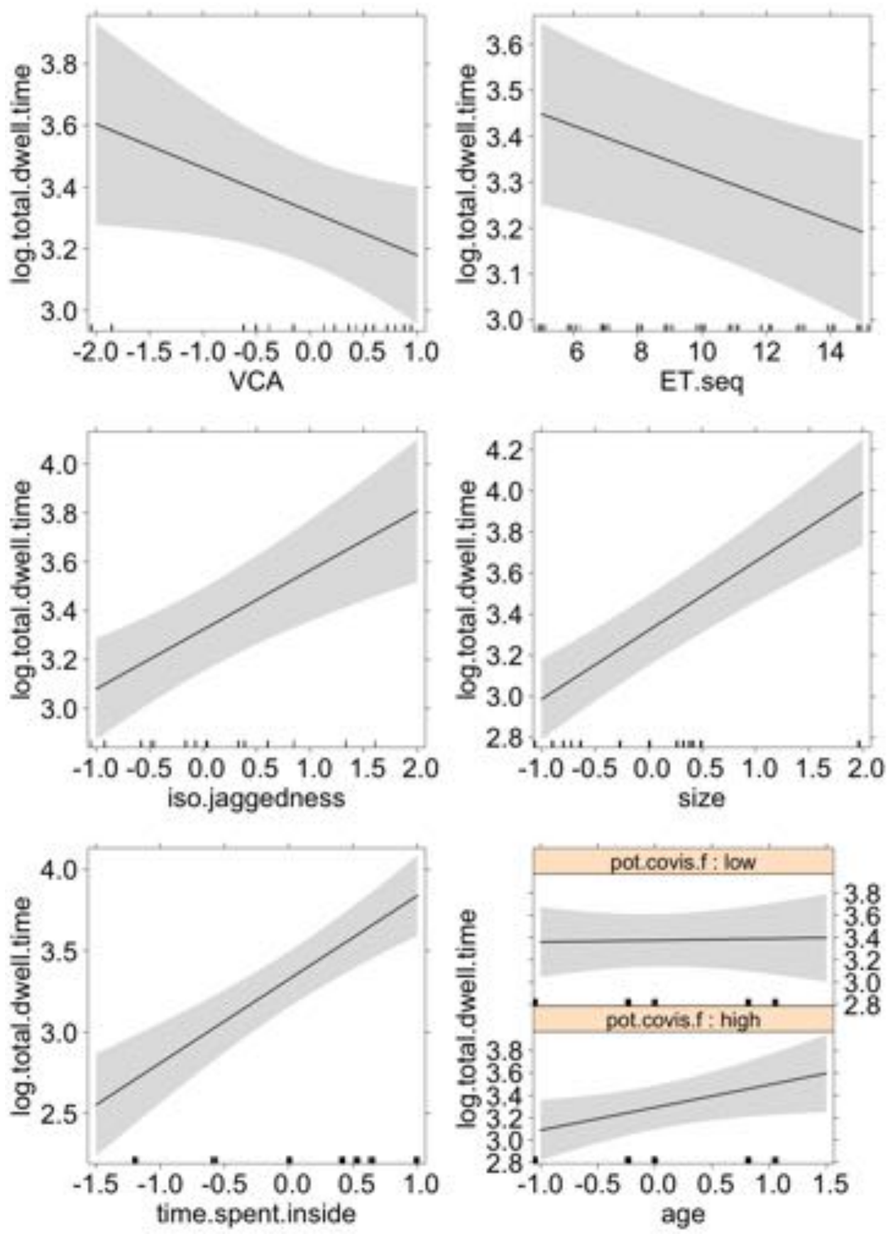


Figure 23.3.: Visualisation of the effects for Model 23.6.

	Model 23.6
(Intercept)	$3.55\ (0.13)^{***}$
VCA	$-0.14\ (0.07)^{*}$
View. Sequence	$-0.03\ (0.01)^{*}$
Iso. Jaggedness	$0.24\ (0.06)^{***}$
size	$0.34\ (0.05)^{***}$
Time Spent Inside	$0.51\ (0.09)^{***}$
Potential CoVis.(factorial)low	$0.08\ (0.13)$
age	$0.20\ (0.10)^{*}$
Potential CoVis.(factorial)low:age	$-0.19\ (0.09)^{*}$
AIC	354.70
BIC	389.19
Log Likelihood	-166.35
Num. obs.	170
Num. groups: ids	9
Variance: ids.(Intercept)	0.05
Variance: Residual	0.33

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.4.: Model 23.6.

	Model 23.7
(Intercept)	$0.49\ (0.04)^{***}$
Potential CoVis.	$0.08\ (0.03)^{*}$
VCA	$-0.07\ (0.03)^{*}$
AIC	-95.21
BIC	-76.40
Log Likelihood	53.61
Num. obs.	170
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.01
Variance: ids.(Intercept)	0.01
Variance: Residual	0.02

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.5.: Model 23.7.

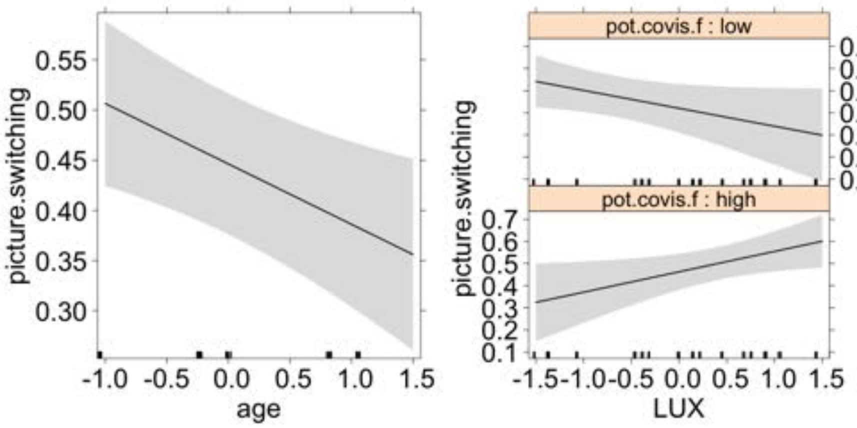


Figure 23.4.: Visualisation of the effects for Model 23.8.

```
Picture-Switching ~ Potential CoVis.(factorial)+
LUX + age +(1 | ids) + (1 | loc) +
Potential CoVis.(factorial)*LUX
```

Model 23.8: Picture-Switching Ratio

The distribution of residuals was homoscedastic, *Kappa* score was 4.3, Marginal R^2 was 0.21 and Conditional R^2 was 0.52. Consult Table 23.6 and Figure 23.4 for the results. Considering *fixations on other people* or *Isovist Area* did not improve the overall fit.

The attempt to make a 'practical' prediction of *Long-to-Short Dwell ratio* was made with the following model based on the previous findings:

```
Long-Short Dwell Ratio ~ VCA + Iso. Jaggedness +
Potential CoVis. + size(cm2) + (1|ids) + (1|loc)
```

Model 23.9: Long-to-Short Dwell Ratio ('practical')

resulting in a *Kappa* score of 2.6, Marginal R^2 of 0.05 and Conditional R^2 of 0.19. The distribution of residuals was homoscedastic. Only the fixed-effect of *size* was significant.

The best-fit model was a simplified version of the one above and described by the formula:

```
Long-Short Dwell Ratio ~ Potential CoVis.(factorial) +
size(cm2) + (1|ids)
```

Model 23.10: Long-to-Short Dwell Ratio

for which *Kappa* score was 2.3, Marginal R^2 was 0.05 and Conditional R^2 was 0.19. The residuals' distribution was homoscedastic. Considering *fixations on other visitors* and *Isovist Area* did not improve the overall fit. Please consult Table 23.7 and Figure 23.5 for the detailed results and effect sizes.

	Model 23.8
(Intercept)	$0.46\ (0.04)^{***}$
Potential CoVis.(factorial)low	$-0.04\ (0.06)$
LUX	$0.09\ (0.04)^{*}$
age	$-0.06\ (0.02)^{**}$
Potential CoVis.(factorial)low:LUX	$-0.17\ (0.06)^{**}$
AIC	-90.79
BIC	-65.70
Log Likelihood	53.39
Num. obs.	170
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.01
Variance: ids.(Intercept)	0.00
Variance: Residual	0.02

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.6.: Model 23.8.

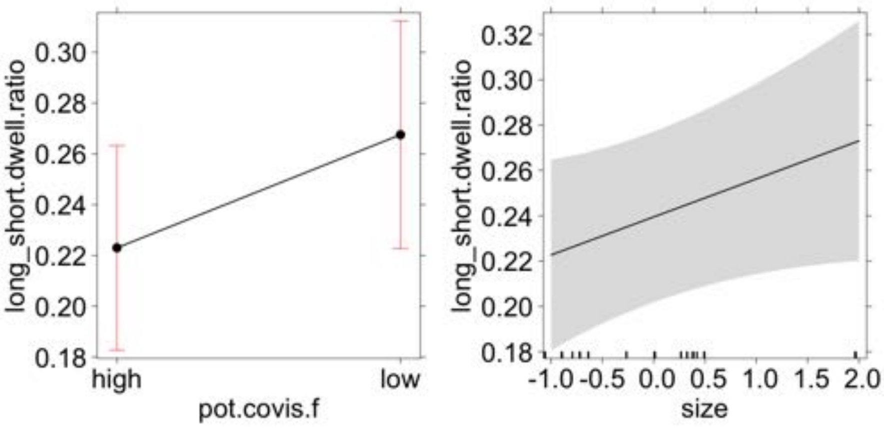


Figure 23.5.: Visualisation of the effects for Model 23.10.

	Model 23.10
(Intercept)	0.22 (0.02)***
Potential CoVis.(factorial)low	0.04 (0.02)*
size	0.02 (0.01)·
AIC	-193.54
BIC	-177.86
Log Likelihood	101.77
Num. obs.	170
Num. groups: ids	9
Variance: ids.(Intercept)	0.00
Variance: Residual	0.01

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, ·$p < 0.1$

Table 23.7.: Model 23.10.

23.3 RECOGNITION MEMORY

Following the analysis presented in the previous section, we will first make an attempt to make a 'practical' prediction of the memory performance relevant to each picture, as it could be done by an interested curator. This means, we will only consider spatio-visual predictors which could be available to a curator during the design process and we will only use factors which have been shown to have a significant influence on *RT* in the previous analyses. The following formula can be derived:

```
RT(inv) ~ Time Spent Inside + VCA + size(cm2) +
RTtrial + proceedingRT + proceedingRTacc+
(1|ids) + (1|loc)
```

Model 23.11: Recognition Times (inverted / 'practical')

Kappa score was 8.3, Marginal R^2 was 0.05 and Conditional R^2 was 0.4. Table 23.8 describes the model in detail. No factors other than *Proceeding RT* reached significance. The majority of variance was explained by random effects of location and participant.

About 90% of all answers given in the Recognition Memory test were correct (possibly due to large differences between 'correct' and 'false' stimuli; see Appendix E). As the *RT accuracy* measure was already considered unreliable in Experiment 2 (where the ratio of correct responses was lower), its analysis here is omitted. It is very likely that the incorrect 10% of answers represent outlying cases and not the true estimate of participants' Recognition Memory.

The 'practical' prediction however can be expanded on a limited subset of participants whose eye-tracking recordings were available. The following model was established as most accurate, based on a step-wise model selection procedure and AIC-criterion:

	Model 23.11
(Intercept)	-0.65 $(0.28)^{*}$
Time Spent Inside	-0.02 (0.13)
VCA	-0.03 (0.10)
size	-0.14 (0.10)
RTtrial	0.07 (0.05)
proceedingRT	-0.02 (0.07)
proceedingRTacc-TRUE	0.63 $(0.25)^{**}$
AIC	706.63
BIC	742.04
Log Likelihood	-343.31
Num. obs.	255
Num. groups: loc	19
Num. groups: ids	16
Variance: loc.(Intercept)	0.15
Variance: ids.(Intercept)	0.21
Variance: Residual	0.67

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.8.: Model 23.11.

```
RT(inv) ~ View. Sequence + Targeted CoVis.factor +
type + Time Spent Inside + Norm. Time on People+
RTtrial + proceedingRT + (1|ids) + (1|loc)
```

Model 23.12: Reaction Times (inverted) with respect to Eye Movement

Its *Kappa* score was 5.3, Marginal R^2 0.23 and Conditional R^2 equalled 0.35. Table 23.9 describes the model and Figure 23.6 presents the results visually.

The strongest effect was caused by type of the artwork - sculpture-like installations were remembered best, and those enclosed behind a glossy surface, the worst. The second strongest effect was caused by *ET sequence* although there was no clear 'primacy effect' visible (Figure 23.7).

Including *Isovist Area* or substituting it for *VCA* did not improve the fit of any of the models.

23.4 SPATIAL MEMORY

Participants have correctly placed 43% of all artworks back on their correct walls in the Miniature Task. Two logit mixed-effect models were built to explain this result. Firstly, a 'practical' model based on spatio-visual factors only was fit into the data:

```
BttW ~ Targeted CoVis..factor+ VCA + size +
```

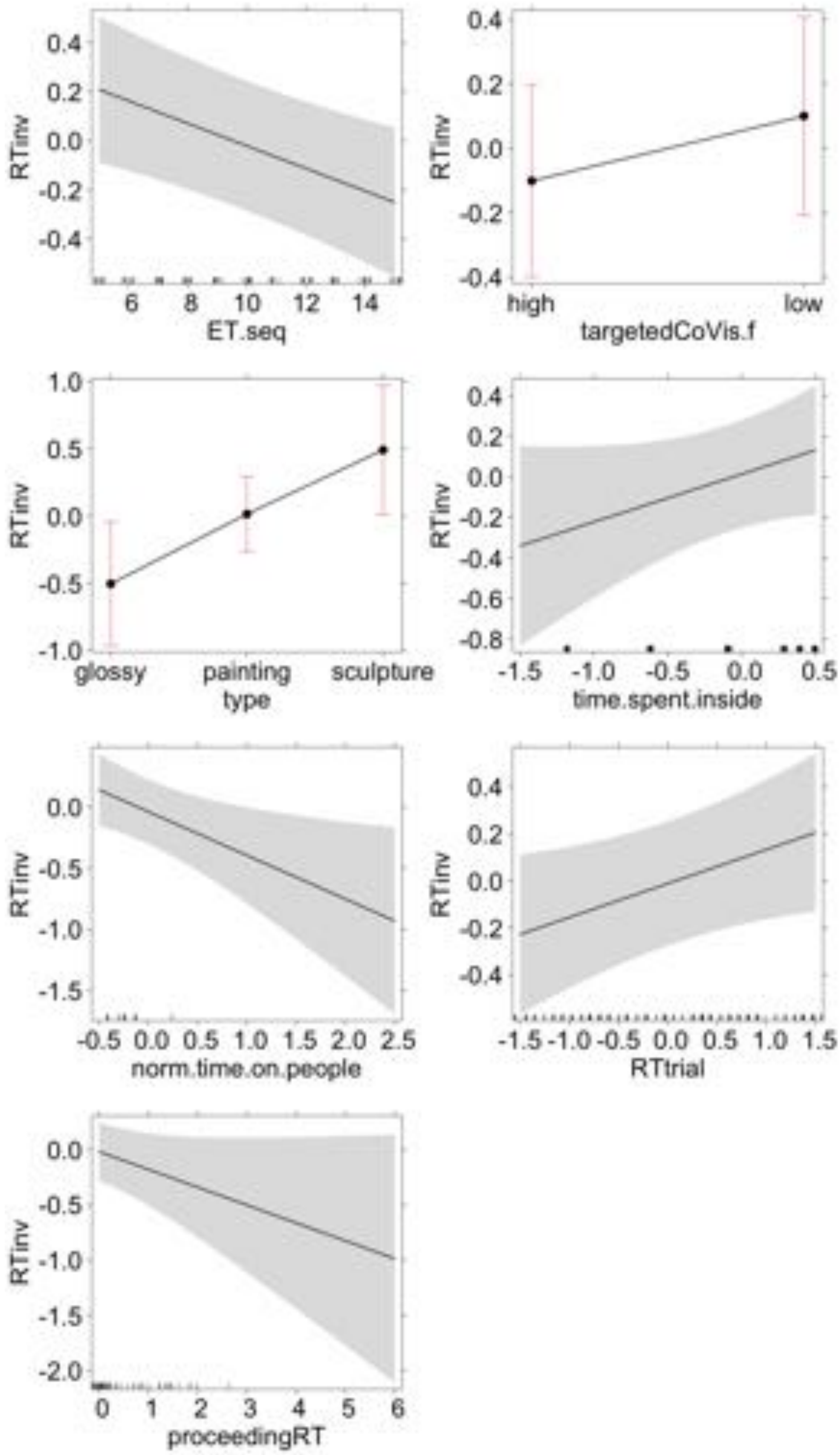


Figure 23.6.: Visualisation of the effects for Model 23.12.

	Model 23.12
(Intercept)	-0.18 (0.25)
View. Sequence	-0.05 (0.01)**
Targeted CoVis..flow	0.20 (0.15)
typepainting	0.52 (0.22)*
typesculpture	0.99 (0.30)**
Time Spent Inside	0.24 (0.15)
Norm. Time on People	-0.36 (0.14)*
RTtrial	0.14 (0.07)*
proceedingRT	-0.16 (0.09)$^{\cdot}$
AIC	422.32
BIC	458.45
Log Likelihood	-199.16
Num. obs.	150
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.01
Variance: ids.(Intercept)	0.11
Variance: Residual	0.69

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.9.: Model 23.12.

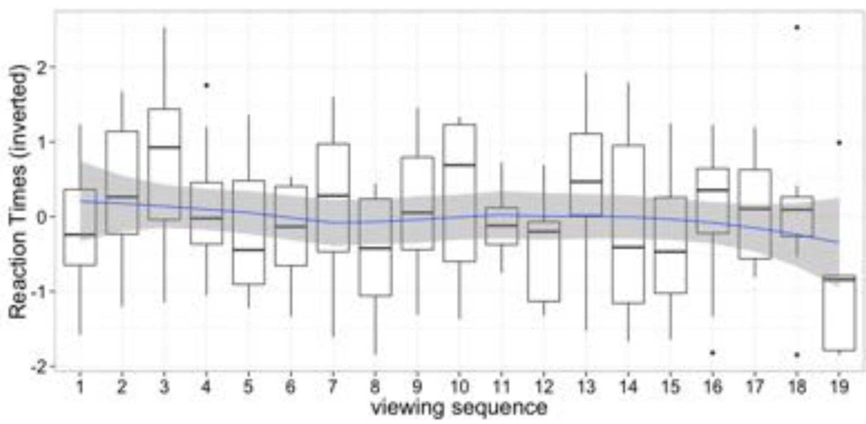


Figure 23.7.: Recognition Memory performance plotted against viewing sequence.

	Model 23.13
(Intercept)	$-0.87\ (0.34)^{*}$
Targeted CoVis..flow	$1.06\ (0.43)^{*}$
VCA	$0.36\ (0.22)^{\cdot}$
size	$-0.06\ (0.18)$
Time Spent Inside	$0.33\ (0.25)$
AIC	394.49
BIC	420.51
Log Likelihood	-190.24
Num. obs.	304
Num. groups: loc	19
Num. groups: ids	16
Variance: loc.(Intercept)	0.27
Variance: ids.(Intercept)	0.69
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.10.: Model 23.13.

```
Time Spent Inside + (1|ids) + (1|loc)
```

Model 23.13: Back-to-the-Wall Ratio ('practical')

Its Marginal R^2 was 0.07 and ˆConditional R^2 0.28. Two factors were significant (or approaching significance): *Targeted Co-Visibility* (measures as 'low' or 'high') and *VCA*. Table 23.10 presents the results in detail.

Further, another model was constructed in order to account for the eye-tracking data available for a subset of all participants and shed more light on the results. The model was described by the following formula:

```
BttW ∼ log(Quantity of ET) + Long-Short Dwell Ratio +
Picture-Switching + View. Sequence +VCA + gender +
Iso. Jaggedness + Targeted CoVis..factor+ size(cm2) +
age + Time Spent Inside + Norm. Time on Sculptures +
Norm. Time on People + (1|ids) + (1|loc)
```

Model 23.14: Back-to-the-Wall Ratio with respect to Eye Tracking

It achieved Marginal R^2 of 0.30 and Conditional R^2 of 0.36. Of factors of interest, *Quantity ET* reached significance, meaning that the more participants looked at a particular picture, the more likely they were to place it back on the correct wall in the subsequence *Miniature Task*. Table 23.11 describes the results.

Analysis of Miniature Task video sequences revealed no significant predictors.

	Model 23.14
(Intercept)	-6.51 (2.45)**
log(Quantity of ET)	0.70 (0.27)**
Long-Short Dwell Ratio	1.83 (1.60)
Picture-Switching	-1.64 (1.13)
View. Sequence	0.01 (0.05)
VCA	0.12 (0.41)
Iso. Jaggedness	0.08 (0.30)
Targeted CoVis..flow	0.26 (0.60)
size	-0.14 (0.25)
gender-male	1.21 (0.67)$^{\cdot}$
age	0.40 (0.34)
Time Spent Inside	-0.32 (0.36)
Norm. Time on Sculptures	0.08 (0.05)$^{\cdot}$
Norm. Time on People	0.04 (0.13)
AIC	227.08
BIC	277.25
Log Likelihood	-97.54
Num. obs.	170
Num. groups: loc	19
Num. groups: ids	9
Variance: loc.(Intercept)	0.09
Variance: ids.(Intercept)	0.20
Variance: Residual	1.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$

Table 23.11.: Model 23.14

24

DISCUSSION

This section will review the results of the BALTIC Case Study in relation to the hypotheses stated previously in Sections 12 and 17.6, as well as to the results of Experiments 1 and 2. The review is based on the structure introduced in the previous discussion sections and—where possible—generic interpretations will be made in consideration to all three studies introduced in the thesis.

It is important to once again emphasise that the Case Study did not provide the opportunity to randomise the picture sequences. Therefore the ability to differentiate between the influence of space and the influence of pictures is limited and only the effects most prominent statistically can be analysed with confidence. This is especially the case when to consider the differences in the size and content of artworks displayed within the exhibition. It has been shown previously how these impact viewing times (e.g. Brieber et al., 2014). However, building on the findings from Experiment 1 and 2, the impact of selected spatial factors can be nevertheless expected.

24.1 VISUAL ATTENTION

Similarly to Experiments 1 and 2, the distribution of dwell lengths confirmed the division suggested in Hypothesis (A1) between a quick scanning of an artwork and its longer, presumably more diligent examination. As it has been shown in Section 4.4, such distinction is not new and has been made previously both in laboratory-based eye-tracking experiments in aesthetics, as well as observational in-the-wild studies of museum visitors. Most recently, this distinction has been revisited and formally quantified by Krejtz et al. (2014). The authors build on the seminal work by Velichkovsky et al. (2005), who distinguished between *ambient* and *focal* mode of oculomotor behaviour corresponding to the neural correlates of information processing. In this classification, *ambient* mode is characterised by short fixations and long saccades, and *focal* mode by longer fixations and shorter saccades. Krejtz et al. (2014) provide a metric which distinguishes between these modes of viewing based on mean values of fixation and saccade lengths. Velichkovsky et al. (2005) also linked *focal* viewing mode with enhanced Recognition Memory. The results presented in the current thesis are in line with this classification, even though they are not built around the units of 'fixations' and 'saccades', but *dwells*, *dwell lengths*, and *picture-switching ratio*. Each *dwell* consists of multiple fixations and those can belong to either *focal* or *ambient* viewing mode. And yet, on the more generic level, human oculomotor behaviour seems to exhibit similar patterns. For instance, less dynamic viewing behaviour has been correlated with better spatial memory in Experiment 2. The technological limitations of Mobile Eye-Tracking technology employed in this thesis does not make it possible, however, to directly compare it to the work of Velichkovsky et al. (2005). Nevertheless, this experimental set-up provides a more robust real-life confirmation of the concept than observational visitor studies (J. K. Smith & Smith, 2001) and preliminarily suggests that low-level distinction between viewing modes translates to more generic one.

Even under radically different viewing conditions: a larger floor area, larger number of more diverse stimuli, and curated, deliberate artistic intentions underlying the artworks' layout compared to Experiments 1 and 2, there was not enough evidence to reject Hypothesis (A2). It stated—based on the works of J. K. Smith and Smith (2001) and Locher et al. (2007)—that despite high variability, the average viewing time of an artwork in a gallery typically is close to 30 seconds. Due to noticeable cross-individual differences in this pattern, it would be an overstatement to say that half a minute is a period after which each painting is typically considered 'viewed'. These periods vary within population and—as we have demonstrated already—can be affected by the spatial positioning of artworks. However, these influences seem to balance each other out and reach an overall average close to 30 seconds. Most interestingly, the content of individual pictures as well as an underlying curatorial intent in their positioning seemed to have little impact on how much time visitors are willing to devote to individual works. Considering the results of Brieber et al. (2014) and Tröndle, Greenwood, Bitterli and van den Berg (2014) regarding the influence of the museum context on increased viewing times, the findings of the current thesis do not confirm their generalised interpretations. Brieber et al. (2014) suggested that longer viewing inside a gallery compared to a computer screen-based virtual exhibition is the effect of stronger 'museum context'. Such interpretation would imply that as the viewing situation becomes more 'museum-like', the average viewing times increase. Three studies presented in the current thesis can be interpreted along this category. For example, artworks in Experiment 2 were exhibited in space which clearly is a psychological laboratory. Participants were required to visit a university building, marked as a research facility. The gallery was constructed from portable wall segments, it contained multiple security cameras, and a one-way mirror. Following the study, some participants made comments about the limited feeling of 'museum-ness' they experienced inside. It might not be an exaggeration to say that this space had as little 'museum context' in it, as it is possible for a large, rectangular room containing art hanging on otherwise empty walls.

On the contrary, Experiment 1 must have ranked higher on that comparison, as the space there used was a 'training' art gallery, used by Fine Art students as studio space and exhibition showroom for their yearly degree show. The building was located at a university campus but was related—by name and function—to the building occupied by Fine Arts department. The corridors were filled with flyers advertising older art events, and the white concrete walls of the gallery had some degree of 'museum-ness' in it, even if it could be considered 'improvised', or 'amateur'. This might be the reason for which no comments on the space's suitability were noted during debriefing. And finally, the last study took place at a world-class contemporary art facility, leaving no doubt to the extent of 'museum context' projected by the host institution. Participants knew they are in a well-known art gallery occupying one of the most prestigious buildings in the city. And yet, contrary to Hypothesis (D2), differences in average viewing times failed to reach statistical significance (compare Figure 23.1, p. 286), even despite the fact that BALTIC contained multiple large artworks, which typically require longer viewing (Brown & Farha, 1966). Additionally, if to consider the fact that apart from spatial layout and the content of art, its number in each study was also different, Hypothesis (B3) stands. It was based on the statement made by Robinson (1928), who suggested that the cumulative viewing times do not decrease proportionally to the number of artworks present in the gallery. Even though the experiments presented in the current thesis cannot be directly compared, *Total Dwell Times* seemed to be little affected by the changing number of artworks requiring viewing. The reader should consider however, that this is still a speculative statement, as too many factors differed individual viewing situations. Most importantly, even if the observed difference in means was true, it would fail to reach statistical significance due to the little sample size and large standard error, especially in the BALTIC Case Study. Hence, all can be said with certainty is that comparing three studies described in this thesis did not show a convincing evidence that average viewing times per picture differ across museum-like environments.

In relation to the results described by Brieber et al. (2014), it seems that what increased viewing times in their experiment might not be exclusively the 'museum context'. One aspect that was not considered by the authors is the spatial nature of that experience. Compared to a computer-based viewing, engaging with art inside a gallery requires locomotion, head movement, and orientation (wayfinding) inside it.

During a computer-based viewing, bodily movement is severely limited, and the amount of required navigation—at least in the computer application used in the aforementioned study—reduced. Not only three-dimensional exploration activates qualitatively different spatial cognition mechanisms, but also results in more diverse visual engagements (in terms of the angles and distances at which they occur). One could say that such engagements are 'less efficient' compared to the straight-on viewing of a computer screen from a chair located in front of it, and that this lack of efficiency possibly increases the time required to process the work to the satisfactory level. However, this cannot be stated without considering wider curatorial interpretation of the problem, namely the fact that diverse viewing and the spatiality of the museum experience is a crucial part of that aesthetic experience (Stavroulaki & Peponis, 2003; Newhouse, 2005).

We almost never come to art galleries with the aim of processing the hosted artworks as fast as it is possible. Quite the opposite - art galleries remain the place one visits to contemplate with little pressure of time. Tröndle, Greenwood, Bitterli and van den Berg (2014) suggests that the viewing times might be the result of the 'aesthetic viewing mode' one enters by passing the gallery door. If true, this statement would suggest that the 'aesthetic viewing mode' was equally-well achieved in all three studies presented in this thesis. Considering that some of them (especially Condition 1 in Experiment 2) carried very limited associations with a real museum institution, it can be concluded that the 'aesthetic viewing mode' is a matter of attitude. Seeing pictures hanging on empty walls, and willing to engage with them is enough to trigger this generic pattern of viewing behaviour. And yet, it is questionable whether similar attitude can be achieved in front of a computer screen (Brieber et al., 2014) where no spatial context is present. An interesting depth to this issue could be provided by a study allowing participants to walk freely in a virtual art gallery whilst wearing a head-mounted display. Does the spatial context of viewing and the limited feeling of virtual presence suffice to simulate a real art exhibition? With the raising popularity of augmented reality technology inside museum institutions (Wojciechowski, Walczak, White & Cellary, 2004) 'virtual home theatres' hosting personally-tailored world-class art exhibitions at little cost could enormously improve the accessibility and reach of the arts. One emerging responsibility of researchers in the field of experimental aesthetics is to investigate if (and how) this process can occur without depreciating the cognitive depth of the experience. The recently revised issue of multisensorial character of the museum experience (Levent & Pascual-Leone, 2013b) can be one of the main challenges.

Hypothesis (B1), referring to the task-specific oculomotor behaviour occurring naturally inside a gallery could not be reliably confirmed with the method previously used in Experiments 1 and 2. The reason for this is the presence of sculptures inside the BALTIC exhibition which constitute a legitimate target for the visitors' visual attention inside the gallery, even though they are not located on flat wall surfaces. The eye-tracking technology used in the study does not make it possible to reliably estimate where eye fixations *would* fall in an empty three-dimensional space *if* there were any sculptures located 'on its way'. The coding procedure employed in this thesis only makes it possible to estimate the 'ending point' of each eye gaze - i.e. a wall surface or a solid object.

While it would be a much biased approach to exclude sculptures from the analysis of task-specific oculomotor behaviour, including them in a reliable manner was impossible with the use of the methods available. The estimations performed in Experiments 1 and 2 involved making predictions of where an eye gaze *should* fall if this fact was determined by visibility and size of individual wall segment only. To perform such a comparison with relation to sculptures, visibility of all grid cells within the navigable floor surface area would have to be considered (and not only of those adjacent to wall surfaces). Further, comparing these visibility values with the actually observed eye movement would only be possible if the eye movement data contained information about grid cells through which the eye sight 'travelled' to reach its final fixation point. This information would describe how probable it was to fixate on a three-dimensional sculpture if the eye gaze was deployed at random directions. One possible solution to this problems are 'sight vectors' proposed by Müller-Feldmeth et al. (2014), although at the current stage of development the method remains unfeasible for larger-scale studies, as it lacks automatisation. Having discussed these technological limitations we should still consider, however, that fixations on empty wall segments or people (i.e. non-artworks) constituted only 15% of all dwell time recorded in the BALTIC study. Considering that pictures jointly occupied 24% of all walls' surfaces and all sculptures and display cabinets only 2% of the total floor area, there is little doubt that participants of the BALTIC study engaged in a task-specific visual behaviour of viewing artworks, as it was predicted by Hypothesis (B1).

Turning to the measures describing the quantitive aspects of visitors' eye movement, Model 23.4

```
log(No. of Dwells) ~ Isovist.Area + size + View. Sequence + t
ime.spent.inside+ gender + Norm. Time on People +
(1|ids) + (1|loc) + Potential CoVis.(factorial)*age
```

best explained *Number of Dwells* deployed to individual pictures. As it shows, females performed less dwells on average and older participants performed less dwells on pictures of lower *Co-Visibility*. Interpersonal differences, however, should be interpreted with caution due to a limited number and variability of participants for whom the eye-tracking data was valid. Unsurprisingly, *Number of Dwells* rose with *Time Spent Inside*. Participants who spent larger *proportion* of their time inside fixating on other people, performed smaller *Number of Dwells* per artwork which suggests, that looking at other visitors was not performed 'in addition' to regular viewing, but that it has 'stolen' dwells which would otherwise most likely be devoted to artworks.

The lack of significant influence of *Potential Co-Visibility* is consistent with the findings from Experiment 1, where *Co-Visibility* did not affect the quantity but only the dynamics of the oculomotor behaviour.

Pictures larger in size and having larger *Isovist Areas* attracted more dwells. The fact that *Isovist Area* was a predictor performing much better than the artworks' *VCA* is contrary to the results found in Exp. 1 and has been suggested as a potential explanation of the results of Experiment 2. As we noted in the relevant sections, the data observed seems to suggest that participants in Experiment 1 consciously directed their eye movement towards pictures of larger *VCAs*, potentially avoiding less comfortable visual interactions at more oblique angles. In Experiment 2 higher *Number of Dwells* could be linked either to larger *Isovist Areas* or *Co-Visibility*, but it was more the restrictive characteristic of space containing wall partitions that *limited* the occurring *Number of Dwells*, rather than some factors in the alternative condition that *increased* them to an unusual level. Inside the BALTIC, *Isovist Areas* are a much better predictor of this eye-tracking measure than *Visibility Catchment Areas*. One suggested explanation of this phenomenon is that the space in the gallery here analysed exhibited a very specific spatial pattern. Namely, it consisted of three main sub-spaces of varied size and regular rectangular shape. As a result, visibility properties of all paintings located inside a single sub-space were very similar while average visibility properties of artworks located in separate sub-spaces were distinct. This low within-subspace variability and high across-subspace variability has a profound influence on the visual choice potentially available to a visitor exploring the gallery.

Since a walking person spends the majority of the time in a single subspace only (and relatively short periods of time in the transition zones between them), the possibility of directing one's eye-movement towards pictures of larger *VCA*—as it seemed to be the case in Experiment 1—is limited. The visual choice cannot be conducted based on this criterion, as all artworks visually available in the surrounding space have almost identical *VCAs*. This means that even if the explanation suggested following the Experiment 1 is correct, it could not be verified in the space arranged by the curator for the analysed BALTIC exhibition. This fact is supplemented by the specific spatial arrangement involving two narrow passages which are the only transition zones between the mentioned sub-spaces. As a result, the visitor is 'forced' to move along the side wall where he or she is more likely to engage at oblique angles with pictures having the largest *Isovist Areas* (i.e. the ones hung at or near the transition zones). We can summarise this by noting that visibility of artworks plays a major role in how many dwells are directed towards them. What exact measures of visibility are a better predictor of this behaviour, depends however on the generic context of the experience and the physical possibilities given by the space to the visitor freely exploring the space. When larger diversity between available visual targets exists, and the space puts limited restrictions on the patterns of exploration (e.g. by making it possible to approach transition zones at multiple angles), in the context of an art gallery visit, 'top-down' decision making seems to guide the number of dwells towards more comfortably exposed pictures. This property of space can, however, be restricted by the designer (purposefully or not) on the level of not local visibility properties, but global relations between separate sub-spaces and consequently imposed walking trajectories.

Model 23.6:

```
log(Total Dwell Time) ~ VCA + View. Sequence + Iso. Jaggedness +
size + Time Spent Inside + (1 | ids) +
Potential CoVis.(factorial)*age
```

best described factors significantly impacting *Total Dwell Time*. Larger *Isovist Jaggedness* (despite low variability, thus against Hypothesis F1), and larger *Picture Size*, as well as *Time Spent Inside* the gallery all had positive impact on the cumulative dwell lengths devoted to individual artworks. Additionally, larger *VCAs* were associated with a decrease in *Total Dwell Time* per picture. Possibly, more comfortable viewing conditions provided by large *VCA* and small *Jaggedness* decreased the time required to process the artwork to a satisfactory level.

Moreover, older participants spent more time looking at highly co-visible pictures while there was no interaction with age for less co-visible ones. In line with the previous expectations described in Hypothesis (F2), *Picture Size* again had the highest impact on the oculomotor behaviour. At least two explanations are possible: that it can (a) suggest special importance of an artwork, and/or (b) simply require more viewing for satisfactory cognitive processing to take part.

As none of the pictures were isolated 'completely', Hypothesis (B4) based on the statement made by Robinson (1928) could not have been tested.

In relation to *Total Dwell Time*, negligible remaining random by-item variability demonstrated that *size* together with spatial measures included in the model comprehensively described the differences between individual locations. Selecting measures based on the exploratory Experiment 1 has proven beneficial in the context of a working gallery's case study. This does not mean that there are no other differences in content and location of artworks which would affect visual attention. It means however, that those differences which remain unaccounted for, do not generalise well across individual experiences. It must be noted, that the sample size of this case study was relatively small, and a significant random by-item variability would be expected in larger sample sizes. Yet, it appears that isovist-derived measures and size of the artworks were able to capture the most noticeable reasons for this variation.

Similarly to the previous two studies, *ET Sequence* had a statistically significant, but negligible negative effect on the quantity of visual attention devoted to each location. The 'museum fatigue effect' therefore seems to be stable, but much less impactful than it could be expected based on the existing literature. As the measure was used for controlling purposes only and does not contribute to path-independent spatial factors being in the centre of this thesis' focus, its further role will not be investigated in detail.

Hypothesis (B2) stated that pictures of larger *VCA* and *Isovist Area* will gather more dwells and longer cumulative dwell times due to the 'top-down' influence of perceived importance of those artworks. This statement can only partially be confirmed. As it has been shown above, it seems that the oculomotor behaviour inside a real, 'messy' art gallery is a mixture of the 'top-down' and 'bottom-up' influences.

To explain how the dynamics of the oculomotor behaviour was affected, *Picture-Switching ratio* and *Long-to-Short Dwell ratio* were investigated. Model 23.8

```
Picture-Switching ~ Potential CoVis.(factorial)+ LUX + age +
(1 | ids) + (1 | loc) + Potential CoVis.(factorial)*(LUX+age)
```

showed that *Picture-Switching Ratio* decreased with age and increased with lighting quality, but only for highly co-visible artworks. More noticeable impact of spatial factors on this outcome measure was expected based on the results of Experiments 1 and 2. The lack of effect might be the result of the spatial layout of this particular BALTIC exhibition. As noted earlier, relatively similar within-subspace visibility properties, as well as regular, uniform hanging distribution utilising all wall surfaces with a similar density might have caused *Picture-Switching ratio* to be much less predictable based on the analysed spatial characteristics.

Long-to-Short Dwell ratio was best described by the Model 23.10:

```
Long-Short Dwell Ratio ~ Potential CoVis.(factorial)+
size + (1|ids)
```

Although its fixed-effect factors could only account for 5% of the total variance, the random by-item effect was not significant and the random by-participant effect increased the percentage of variance explained only to 18%. This result may be due to the fact that the general mean of *Long-to-Short Dwell ratio* in this study sample was low and of a limited variation (M = .24; SD = .13). Still, participants tended to use a higher proportion of long dwells to investigate those pictures which were bigger and more isolated. The reason most likely being that they might simply require longer dwells in order for their entire content to be viewed (Brown & Farha, 1966) and lower potential distraction makes it possible. As the main effect of *Potential Co-Visibility* showed, higher *Co-Visibility* resulted in lower proportion of *long dwells*, as it was predicted by Hypothesis (B5) and the results of the previous studies.

The spatially-determined occurrence of 'spotty-focused' attention in line with Hypothesis (B6) cannot be confirmed with full confidence in the context of the BALTIC Case Study. Participant's varied in the dynamics of their oculomotor behaviour, but varied inconsistently in respect to spatial factors proven to be impactful in the previous studies. *Picture-Switching ratio* was almost unpredictable and *Long-to-Short Dwell ratio* was generally lower (0.24), compared to Exp. 1 (0.33) and Exp. 2 (0.46). As the experiments are not directly comparable, little can be said about this difference, except that whatever caused it might have also affected participants' memory and other aspects of the experience. A large, open-plan exhibition containing diverse artworks arranged by a curator to generate a complex aesthetic experience was explored through shorter visual interactions than a set of randomly positioned, relatively similar artworks on a smaller floor surface area in Experiments 1 and 2. So large differences across the studies perhaps suggest, that the dynamics of the visual attention is generally not very well predictable when the setting introduces variation only in the location of artworks, but also in their size and content.

'Practical' predictions in all cases yielded lower Marginal R^2 values compared to the best-fit models, explaining at most 11% of the variance (in case of Model 23.5). It thus seems that effective modelling of the influence of space on visitors' visual attention cannot be performed without estimating visit times and basic demographic characteristics of the museum patrons, such as their age and visit lengths. Much lower R^2 values in models lacking these predictors can be explained by the nature of this statistical technique - a good statistical model explains more than 'sum of its parts', as it considers multiple factors having an effect on the outcome variable while being able to correct for their influence on each other. For this reason, low R^2 values of 'practical' predictions do not necessarily mean that the effect of space on human eye movement inside art galleries is that small, but that it is different for individual participants, and additional data about their visit must support those estimations. In fact, this can be less of an issue for a practitioner than it at first appears. Many museum institutions typically monitor (in a more or less formal manner) the numbers and general characteristics of their visitors, including their average visit times. Some institutions, using an electronic ticketing systems gather this data in real-time. Moreover, reliable estimations of visit lengths can be based on the known works relating the average viewing times per picture (J. K. Smith & Smith, 2001; Brieber et al., 2014) depending on the exhibition size (Serrell, 1997) and walking speeds (Yoshimura et al., 2012). Combining this with the known number of pictures present in the gallery might be able to deliver reliable estimates of the distribution of visit durations for the given exhibition. The technical details remain beyond the scope of this thesis, but it is important to note how difficult it is to predict the influence of space on human oculomotor behaviour without considering the most important agents of any visitor-artwork interaction: the humans themselves.

24.2 RECOGNITION MEMORY

The 'practical' prediction based only on spatial factors did not reach significance in the Reaction Time data analysis. Large proportion of the variance however (about 40%) was explained by the by-item and by-participant variance. Available eye-tracking recording shed more light on the individual differences in viewing strategies and made it possible to establish clear relations between the predictors.

In Model 23.12 (p. 297), 5 factors reached significance. Firstly, *type* of the artwork played a significant role in how well the work was remembered. If to consider *type* as a measure of 'salience', this result could be related to *Perceived Salience* factor which was predicted to have a significant impact on Recognition Memory (Hypothesis C2) but had none in the previous two experiments. It is difficult, however, to generalise the influence of the artwork's *type* beyond the particular cases included in this exhibition. This particular effect might instead occur due to two factors: light being reflected on glossy surfaces might have jeopardised visual interaction with the artwork; and a relatively small number of distinctive sculpture-like objects might have created stronger memories among much larger number of 'traditional' paintings.

ET sequence also played a significant role in line with the ideas of the 'museum fatigue effect'. In general, pictures viewed later during the visit were remembered worse. This finding contributes to the debate on the justification for designing linear exploration paths throughout the exhibitions where all participants are equally likely to encounter some objects later in their visit. A gallery offering circular movement would potentially distribute this probabilities in a more balanced (although not random) manner throughout the space.

The more participants looked at other visitors, the worse their memory of artworks was. This explanation is in line with the one presented in the previous section, where it was suggested that the presence of other people 'steals' fixations from the artworks. It appears that as they 'steal' attention, the quality of cognitive processing also suffers and artworks are worse memorised spontaneously. Those findings also point to the importance of considering the social aspect of a gallery visit. People influence other visitors and one is never able to fully engage with an artwork as s/he would in an empty room. Since art galleries are public spaces, the significance of considering the influence of other people during the design process should certainly be greater.

Spatial factors did not reach significance in this model. This includes *Targeted Co-Visibility*, which in Hypothesis C1 was predicted to decrease Recognition Memory performance but was shown to have little influence under the presence of equal *VCAs* in Experiment 2. The BALTIC Case Study results stay in line with those findings.

The non-significant impact of other spatial factors must be discussed with consideration to the research set-up employed in this case study. Unlike in the two previous experiments, artworks inside the BALTIC gallery were fixed to their individual locations. For this reason, the effect of space on visitors' memory cannot be separated from the influence of individual pictures. And this by-picture impact can manifest itself through two sets of very influential characteristics: the 'salience' of an artwork (arising from its size, content, and colouring relative to the nearby paintings) and the personal interest generated by separate artworks in individual participants. To reach significance under this research setting, spatial characteristics (such as *VCA*; Hypothesis C6) would need to have an effect strong enough even under the presence of such 'noise' in the prediction derived from the by-painting differences. This is clearly not the case. The influence of space on visitors' Recognition Memory is therefore relatively subtle - it can only be generalised to a limited extent. Therefore the role of space in the creation of the visitor's experience—as previously mentioned in Section 16.2—is indirect. Space makes it easier or more difficult to create and sustain various types of visual engagement but does not—on its own—'guarantee' stronger memories for all visitors. *Art* and *personal preference* still seem to be the most important factors affecting our art gallery experience.

24.3 SPATIAL MEMORY

The fact that Miniature Task video sequences could not have been predicted based on the available dataset puts in doubt the reliability of this method in less restricted experimental conditions. Compared to Experiment 2 (where *Miniature Task sequence* was linked with *RT* and

Total Dwell Time), the Miniature Task presented to the participants of the BALTIC Case Study contained no pre-indicated picture locations. The number of the artworks and complexity of the layout (resulting in more possible combinations) was also higher. As a result, it seems that participants solved the Miniature Task in a less structured manner, with no apparent logic to the sequence of the moves.

Spatial Memory performance measured by *Back-to-the-Wall* score however was linked to spatial factors more than it was observed in the previous 2 experiments. It was higher for those pictures which had lower *Targeted Co-Visibility*, larger *VCA*, and were viewed more (i.e. for longer and more often).

The inverted relation between Spatial Memory performance and *Co-Visibility* contradicts Hypothesis C3 which predicted a positive impact of this factor due to increased perceived spatial importance of more *co-visible* paintings. As no evidence was found to support this hypothesis in the previous experiments, in can be rejected. As it has been discussed in Experiment 2, it seems that instead of increasing Spatial Memory performance, highly *co-visible* exhibition set-up might have a negative influence on this aspect, especially to a subgroup of visitors performing less effectively in such an environment. The effect was even stronger in a real-life exhibition set-up involving highly diverse set of artworks. How decreased Spatial Memory performance affects the holistic experience of an art gallery visit is an issue which should be addressed in separate studies (see e.g. Lu & Peponis, 2014).

Participants who spent larger percentage of their time investigating other sculptures had higher scores on average, potentially indicating that more interested visitors left with better spatial memories in general.

In a larger exhibition space of the BALTIC, the dynamics of oculomotor behaviour had smaller impact on memory than it was the case in more confined experimental spaces of the previous studies. Memory for picture's location was better if a painting was less *co-visible*, although in the BALTIC Case Study this is also correlated with earlier presence in the visitors' viewing sequence. Such an interpretation would remain in line with the results of Experiments 1 and 2 which did not show a linearly negative influence of *Co-Visibility* on Spatial Memory. In fact, in some cases the opposite was true. The true significance of this effect is even more doubtful when to consider the analysis of spatial impact on visual attention in this study (Section 24.1). In it, we have concluded that the setting of the gallery did not clearly caused a 'spotty-focused' attention, and therefore a presumed influence of *co-visibility* would have to affect Spatial Memory without affecting the dynamics of visual attention.

Positive impact of *VCA* however has been observed previously in relation to Recognition Memory and linked with eased cognitive processing of paintings located within larger *VCAs*. Its impact on Spatial Memory is an interesting finding, especially considering that *Quantity ET* also had a significant effect (Hypothesis D4). As the analysis of visual attention in Section 24.1 has shown, due to spatial arrangement of the gallery *VCA* was not a good predictor of the *total dwell length* falling on each artwork. A large proportion of viewing periods therefore occurred at oblique angles, outside of 'areas of comfortable viewing' described by *VCAs*. It appears that even those non-ideal interactions increased the chance of correctly recalling the location of separate paintings in the Miniature Task, while doing so within its *VCA* could have contributed even further to strengthening this spatial memory.

Part VII

SUMMARY AND CONCLUSION

25

SUMMARY

25.1 CONTRIBUTION OF THIS THESIS

It has been mentioned on multiple occasions earlier in this thesis that the two primary functions of spatial layout that can be considered in a user-centric exhibition design is to strengthen the visitors' focus on artworks and to empower them to seek for linkages between separate artworks. In this section we will turn back to the previously stated hypotheses created around these questions and review the degree to which they could have been confirmed in each study. Not all hypotheses were tested in each experiment. Especially when the initial interpretation was fully convincing or where the study's experimental design was not fit-for-purpose. The revision is presented in a table, with uppercase words indicating high level of certainty and lowercase font symbolising room for verification in further work. The 'u/t' symbol stands for 'untested' and indicates instances at which the given hypothesis could not be verified for the reasons mentioned above. The reader should consider that due to the nature of the Null Hypothesis Significance Testing theory, the lack of significant finding is not statistically equivalent to the confirmation of the lack of such an effect. In the table below therefore, those fields indicating 'no evidence' by definition are much more uncertain than the ones where statistically significant evidence has been found. Bold font marks these hypotheses which were previously indicated as key original (as opposed to confirmatory) contributions of the thesis. Following the hypothesis summary in Table 25.1 further sub-sections will briefly review the major limitations of this work.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
A1 - Eye movement on artworks can be distinguished between (a) getting a quick 'gist', and (b) diligent examination of individual pictures.	YES	YES	YES	The division was clearly visible. Some spatial conditions (e.g. Cond. 1 in Exp. 2) distinctively promoted longer dwells compared to others.
A2 - *Total Dwell Time* per picture averages to about 30 seconds.	YES	YES	YES	Similarity of mean viewing times with those suggested by unobtrusive visitor studies demonstrates that the research procedure closely simulated a 'real' art gallery visit.
A3 - The by-picture variance in *Total Dwell Times* will be lower than the by-location variance.	**YES**	**YES**	**u/t**	**Space, and not the content of artworks, had the most profound influence on the measure considered in curatorial studies as a key indicator of the visitor's engagement.**
A4 - Similarity of 'exhibition visit scripts' undertaken by separate visitors varies across spatial set-ups.	YES	YES	u/t	
A5 - *Number of Dwells* and *Total Dwell Time* should predict participant's *Spatial Memory* of individual objects, but not their *Recognition Memory*.	no	no	no	With the exception of Spatial Memory in Experiment 1, both memory measures were predicted by the quantity of visual engagement. As the hypothesis was derived from a study where participants spent only 5 seconds investigating the scene, the exact explanation attached to this hypothesis seems not to be applicable to the thesis' findings.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
B1 - Visitors inside art galleries naturally employ a distinctive viewing mode, under which they fixate on artworks more, than it could be predicted from the plain visibility and surface area of those stimuli.	YES	YES	yes	
B2 - More prominent locations (operationalised as those having larger *Isovist Areas*, and larger *VCA*) attract longer *Total Dwell Lengths* and higher *Number of Dwells*.	**YES**	**u/t**	**yes**	**The reason for this effect seems to be a mixture of the 'top-down' and 'bottom-up' strategies and could have not been fully explained within the scope of this thesis.**
B3 - Average *Total Dwell Time* per picture does not decrease proportionally to the increasing number of pictures inside the gallery.	YES	YES	YES	Only non-significant variations were observed across the experiments.
B4 - 'Complete' isolation of a picture increases its viewing times.	u/t	no	u/t	The lack of effect in Experiment 2 might arise from the lack of within-condition variation in the isolation of separate pictures. The hypothesis can still be true in more diversified spatial settings.
B5 - Higher *Co-Visibility* will have a negative impact on the proportion of *long dwells*.	YES	YES	YES	
B6 - Higher *Co-Visibility* will cause higher *Picture-Switching* ('spotty-focused' attention).	YES	YES	no	The dynamics of oculomotor behaviour was generally less predictable (i.e. consistently dependent on the fixed factors) in the BALTIC Case Study. Moreover, the occurrence of 'spotty-focused' attention was not directly linked with the decreased memory performance.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
B7 - Space has an equal influence on *long dwells* and on *short dwells*.	**yes**	**NO**	**u/t**	**Radically different spatial arrangements in Experiment 2 did not affect the number of *long dwells*. More diligent visual engagements are therefore to a smaller extent guided by spatial factors.**
B8 - Higher number of adjacent pictures reduces the number of *long dwells*.	u/t	no	u/t	In all cases the variation in the predictor was too small to verify the hypothesis.
C1- Increased *Co-Visibility* decreases Recognition Memory performance.	yes	NO	no	In Exp. 1 the effect found can be due to the existing correlation between co-visibility of paintings and their VCA. The set-up of Exp. 2 ensured no such correlation.
C2 - Higher *Perceived Salience* of an artwork is correlated with its higher *Recognition Memory*.	no	NO	u/t	It must be noted, that in BALTIC Case Study no Salience Rating was included but two alternative measures (*size* and *type*) were considered instead. *Type* of artwork had a significant impact on Recognition Memory yet it is difficult to generalise this effect beyond the studied exhibition.
C3 - Higher *Co-Visibility* increases Spatial Memory of artworks (possibly due to their perceived spatial importance).	no	no	NO	The opposite effect was observed. In the experimental settings the effect was not clearly destructive, but affected some group of worse performing participants. In the Case Study containing more diverse set of artworks the effect was stronger and generalisable across all participants.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
C4 - Higher *Isovist Jaggedness* results in better *Recognition Memory*.	u/t	u/t	NO	The hypothesis was not in the central interest of this thesis and cannot be fully verified in experimental settings which varied in so many other spatio-visual measures (which are deemed more important in curatorial literature).
C5 - Larger *Isovist Areas* causes better *Recognition Memory*.	u/t	NO	no	VCA has been a better predictor of Recognition Memory performance.
C6 - Larger *VCA* contributes to better Recognition Memory performance.	**yes**	**YES**	**no**	**In the Case Study setting, the effect of VCA was not strong enough to remain observable under high diversity of individual pictures fixed to constant locations inside the gallery. The predictor was also correlated with other visual features (such as co-visibility and position in the viewed sequence).**
C7 - According to the 'serial position effect' pictures seen as first are remembered better. (RT /Spatial Mem.)	no /yes	no /no	no /no	
C8 (Section 16.1) - Longer *Total Dwell Times* will result in deeper cognitive processing and contribute to better *Recognition Memory*.	YES	YES	no	Compare to A5.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
D1 - Memory performance decreases with age. (RT /Spatial Mem.).	NO /YES	NO /NO	NO	Only under specific spatial circumstances older participants seem not to be able to compensate for decreasing cognitive performance with their viewing strategy. Strategy-related factors (such as *Time Spent Inside*) were much stronger predictors.
D2 - BALTIC Case Study will result in longer average *Total Dwell Time* per picture due to its stronger 'museum context'.	-	-	NO	The differences were not significant. The arrangement of Exp. 1 and 2 seem to have created a sufficient feeling of immersion into the art gallery context.
D3 - Impact of space is higher on those participant who stay in the gallery for longer.	YES	u/t	u/t	Strong evidence of such relation on memory in Exp. 1. The hypothesis cannot be confirmed in Exp. 2 due to very low by-location variability within conditions. Cannot be verified in the BALTIC due to low sample size.
D4 - Longer viewing times positively influence memory performance.	YES	YES	YES	
E1 - In Experiment 2, the dynamics of visual attention will be higher in 'high co-visibility' condition, but the quantity of the attention deployed will not be affected.	-	**YES**	-	**Contrary to the previous literature (using less precise measurements) *co-visibility* has been shown to affect *how* we look at pictures, but not *for how long*.**
E2 - *Time spent inside* will remain a strongly influential factor in both radically different conditions of Exp. 2.	-	YES	-	
E3 - *Perceived Salience* will have a stronger influence in Exp. 2, where pictures were hung interchangeably.	-	NO	-	See comment to C2.

Hypothesis	Exp. 1	Exp. 2	BALTIC	Comment
E4 - When *VCAs* are kept equal in Exp. 2, no difference in Recognition Memory performance will be observed.	-	**YES**	-	**See comment to C6.**
F1 - *Isovist Jaggedness* will have no influence on eye movement under its low variability in the BALTIC.	-	-	NO	*Isovist Jaggedness* consistently had an impact on eye-movement patterns, meaning its effect is strong and consistent.
F2 - Larger paintings in the BALTIC Case Study will attract more attention.	-	-	YES	Also see comment to C2.

Table 25.1.: Summary of all hypotheses stated in the thesis.

At the beginning of this thesis we have identified two key functions associated with the spatial layout of museum exhibitions (see Section 5.1 on page 79):

1. to facilitate the curatorial narrative of the exhibition, or communicate a predefined message;
2. to strengthen the viewers' focus on the artworks, or his/her 'aesthetic experience'.

We have also identified *Visual Attention*, *Recognition Memory* and *Spatial Memory* as processes potentially indicative of such impact.

The empirical work presented in the thesis demonstrated that the selected cognitive processes indeed were susceptible to modifications of spatial arrangement of the same set of artworks. Within this framework it has been demonstrated that:

- *Recognition Memory* and *Spatial Memory* are sensitive to distinct aspects of the environment. This shows that two separate cognitive processes contribute to the creation of the overall visitor experience and that these two separate cognitive processes can be influenced by the curator using distinct spatial characteristics of the designed environment (Hypotheses C1, C6).
- The influence of space on memory and attention can be larger than the influence of artworks' content (Hypotheses A3, C2).
- *Visual Attention* has also been classified within two separate sets of measures: describing its 'quantity' and 'dynamics'. It has been demonstrated that, contrary to the existing literature, those two aspects are not always identically influenced by spatial factors and that their impact on memory performance is non-trivial (Hypotheses A1, E1).
- In relation to the above point, more looking in general results in better memory of artworks (Hypotheses A5, C8, D4), although closer analysis demonstrated that it is rather 'not enough looking' that jeopardises it.
- Furthermore, more dynamic (or 'spotty focused') attention need not be linked with generally worse visual experience, as visitors compensate in spaces encouraging them to engage with artworks in a very dynamic manner by investigating them repeatedly, until satisfactory cumulative viewing time is achieved (Hypotheses B3, B7, E1).
- In result, more 'distractive' space has not been shown to decrease the understanding of the exhibition (Hypotheses B3, C1).

- Further analysis of this linkage between memory results and spatio-visual analysis confirms the previously predicted role of *Visibility Catchment Areas* (Stavroulaki & Peponis, 2003) in supporting deeper cognitive processing of artworks in the museum context (Hypotheses C5, C6, E4). This shows the importance of considering context-specific visibility measures in architecture.

- On the more generic level, the theoretical framework of 'Spatially Guided Movement' (Wineman & Peponis, 2010) can be used to explain the relations between cross-environmental and cross-individual variability in the data describing the aforementioned cognitive processes (Hypotheses A4).

- Consistently with this framework, space affected those staying inside it for longer more. This means that it is the more interested, longer engaging visitor who will, in the end, develop a less personalised, less individual experience, and remain under a stronger influence of the curator who had designed the space (Hypothesis D3).

25.2 ON GENERALISABILITY OF BEHAVIOURAL AND COGNITIVE PATTERNS

As the reader might have observed, the R^2 values describing the proportion of variance in the dependent variable explained by the considered predictors highly varied throughout the analysed studies, ranging from .05 to .70[1]. Table 25.2 below lists the ranges of these effect sizes aggregated by study and dependent variables.

	Exp. 1	Exp. 2	BALTIC
Visual Attention	.21–.62	.06–.70	.05–.53
Recognition Memory	.11–.67	.05–.16	.05–.23
Spatial Memory	.12–.14	—	.07–.30

Table 25.2.: Range of effect sizes (smallest and largest *Marginal* R^2) observed across the studies conducted within the current thesis.

1 In many cases the Conditional R^2 values were calculated with a statistically significant impact of *experimental condition*. In itself, this factor does not provide the precise understanding of all spatio-visual factors influencing the existing difference and thus somewhat artificially increases the numerical statistic describing our understanding of the underlying processes. In result, the 'true' effect sizes might therefore be even lower than as they are by the estimated Conditional R^2 values. For these reason, here we focus on Marginal R^2 values

There is a number of factors which limit the predictability of visual attention and memory performance inside art galleries[2]. Looking across the studies, the primary one seems to be the studied setting - BALTIC Case Study unsurprisingly generated lower effect sizes than the more controlled experimental environments. It therefore appears that in a setting in which curators and artists typically exhibit their work, their influence on the visitor experience through the spatio-visual means is the most limited[3].

One reason for this is certainly the presence of other people inside the galleries. This thesis purposefully omitted an in-depth analysis of the interaction between other people present in the gallery and human engagement with an art exhibition. The dwells falling on other visitors inside the BALTIC gallery were controlled for statistical purposes, but were not subject to an in-depth analysis. While the problem can be viewed as separate from the *influence of space* (which this thesis is primarily concerned about), in reality, facilitating social aspect of the visit is a large part of any museum's function, closely related to its spatial layout (Peponis & Hedin, 1982; Wineman & Peponis, 2010). The main principle being that spatially-defined patterns of social encounter are rarely random and they partially define the experience specific to any given building. Also researchers in Human-Computer Interaction for long have been interested in the touching point between the two effects (Akpan, Marshall, Bird & Harrison, 2013). For example, Dalton, Collins and Marshall (2015) measured the percentage of dwells falling on other people during a shopping task in a commercial mall and Brignull and Rogers (2003) described 'the honey pot effect' - the ability of a group of people gathered around a single public display to draw attention of others. Informal observations of crowds gathering in narrow corridors of some art galleries for no apparent reason, or of visiting groups following a guide throughout the exhibition leave no doubt of the importance of other people to our experience of space. Social behaviour within the architectural context therefore remains highly relevant to curators designing the art experience for a larger audience. With the decreasing cost of research technology (such as mobile eye-trackers, virtual reality, or position tracking), it will become more viable to study larger groups of people interacting in public architectural spaces. The study of cognitive processes in a social context, even though deemed important (Hutchins, 1995), is thus still awaiting its full revival.

The predictability of the results differed not only across the galleries but also across the dependent variables. This means that separate

2 Although R^2 values above 0.26 are generally considered 'high' in psychological science (Cohen, 1988), rarely researchers have the possibility to control for so many relevant factors.

3 This statement should be however interpreted with caution as BALTIC Case Study also involved lower number of participants and a relatively high percentage of the collected data must have been removed during the analysis.

parts of the gallery experience can be explained to a different extent. The factors that seem to play a crucial role in this respect are subjective art appreciation (e.g. Brieber et al., 2014), individual memory skills (e.g. Mulligan, 2013) and spatial skills (e.g. Hegarty et al., 2006). This is also exemplified by the fact that removing up to 3 most influential data points[4] from the models was often able to improve the R^2 values by 0.1 - 0.15, which is a notable difference. What would be a justifiable statistical procedure, however, is purposefully not reported in this thesis. Art galleries and the context of art contemplation is where 'outlying' behaviour occurs naturally, often being even implicitly encouraged by the context. Art galleries are places were individual viewing behaviour is allowed, or even desired. Perhaps, the art experience simply by definition can be modelled to a very limited extent and does not allow the statistical procedures to fully describe the depth and variability of the experience. And yet, it was able to prove the little repeatable influence of spatial factors, often outweighing this of individual artworks. The influence of space is therefore limited partially because a visitor who is free to explore the exhibition is equally free not to attend to it if he or she so wishes (Screven, 1976).

Finally, an interaction of these two groups of effects can be considered: not only some aspects of the individual experiences are highly varied (e.g. art appreciation, uncontrolled in this thesis), and some spatial settings (e.g. real-world, busy art galleries) decrease the predictability of the cognitive interactions. But also some places might further nurture, promote, and embrace unpredictable experiences through which subjective, individual and often unique understanding develops. By definition, art galleries are such places.

25.3 CAN EXPERIENCE EVER BE PRE-DETERMINED BY SPATIAL LAYOUT?

From the viewpoint of architectural analysis, this thesis utilises path-independent measures. Their potential influence derives from the assumption that a person randomly exploring the environment has higher (or lower) chances of falling under the influence zone of a particular picture, depending on its spatial properties. As these measures do not require the path of the visitor as an input, the amount of variance they are likely to explain decreases, but their practical usefulness for the purpose of simulating distinct design alternatives rises. High variance of human behaviour is by definition expected in relation to these spatial factors. In fact, by recommending unrestricted exploration, providing generous time limits, and no pre-defined viewing agendas, this thesis purposefully maximised the random as-

4 Data points are here understood as all aggregated interactions between one participant and one artwork.

pect of human spatial behaviour. This fact reflects the role of the built environment in our everyday lives: the space can be viewed merely as a background to our primary actions and behaviours, but at the same time it is always present and affects all of it, anytime.

The concept of *spatially guided movement* (Wineman & Peponis, 2010) suggests that human movement inside art galleries is not determined, although it can be restricted to a particular sequence of rooms, so that its variability is limited (Choi, 1999). It is therefore *determined* by the physical boundaries, but *guided* by spatio-visual relationships between artworks, and possibly *random* within these limitations. This corresponds well with the curatorial theory which supports empowering visitors in seeking their own interpretations of the exhibitions (L. C. Roberts, 2004) but does not argue for completely unrestricted (un-curated) art galleries. The *guidance* of visitors' movement and attention can therefore take place through modifying the probability that certain isovists and co-visible arrangements are experienced more than others. As the majority of this potential outcome nevertheless depends on other factors (such as personal preference, the exact path taken, or the freedom to ignore various exhibits), it is assumed that the curator will only ever be able to aid, or inhibit this cognitive outcome to a limited extent.

From the analytical viewpoint these probabilities of various engagements could potentially be adjusted based on the known behavioural patterns of specific museum's visitors. A plethora of additional demographic information on local audience is often available to the museum institutions. It would be a misinterpretation however, to assume that the methodology presented in this thesis can ever aim at reaching full explanation of the pattern of visitor-artwork interactions.

25.4 GENERALISING TO OTHER BUILDING TYPES

We have discussed how space modifies viewing behaviour of art gallery visitors even under the presence of large individual differences in viewing styles. The spatial factors which caused the observed effects were sometimes more radical (Experiment 2), and rather subtle in other cases (BALTIC). For this reason, the probability of a visitor falling under the semi-deterministic influence of space differed across the environments. Space can therefore have an enormous impact on our behaviour in some situations, and less so in others, even when the context (art gallery visit), our goals (viewing art), and the items of interest (artworks) are identical. When space has been purposefully designed to have a noticeable impact (e.g. Experiment 2), we could still imagine a situation when a visitor attempts to impose a personally preferred viewing behaviour despite the limitations of space. This could be done perhaps by running between pictures, or at least spending prolonged viewing periods in front of the wall partitions from where two artworks could be observed simultaneously. The fact that none of these situations was a regular occurrence leads to a belief that the influence of spatial layout on human behaviour inside an art gallery is tied to a set of norms typical for this particular situational context. A curator designing an exhibition makes use of these norms. However, if we consider other building types—or the built environment in general—those norms of exploration vary tremendously. As a result, the effect which space has on the behaviour and cognition of its users is also likely to vary. Methods employed in this thesis and the analytical focus on separate visual interactions can therefore be potentially applied only to those situations where the norms of exploration and engagement are similar. This might include cases where users do not rush through space and spontaneously engage with distinct visual stimuli. Potential extensions of this work's methodology can for instance be foreseen in the context of (non-emergency) signage.

26

LIMITATIONS

In this section we will review major limitations of the work presented in this thesis and suggest how these challenges can be tackled in the future studies.

26.1 CRITICISM OF SPACE SYNTAX APPROACH

In order to quantify relevant spatio-visual characteristics of an art gallery layout, this work has relayed on methods and techniques developed within Space Syntax studies. This theory however bears a certain set of assumptions and simplifications which have been criticised over the years by multiple scholars.

One of the common criticism of Space Syntax techniques is that they rarely consider metric distance and focus on the topological understanding of space (Montello, 2007). In the context of this thesis, this is a viable concern. It could be the case that despite the fact humans process paintings at large spectrum of distances (as described in Section 3.3), the quality or type of this processing differs[1]. A possible extension of the measures used in this thesis would therefore include weighting of the artwork's isovist and Visibility Catchment Area based on its distance from the viewer. Some preliminary developments have been made in this respect on the methodological ground (Dalton, Marshall & Conroy Dalton, 2013), although their relevance lacks empirical confirmation.

Some other criticism of the Space Syntax approach is the fact that it underplays the impact of interpersonal differences (Montello, 2007) on the spatial experience. As the large spread of Spatial Memory performance showed in this work, personal abilities and strategies indeed might have had a crucial impact on the human comprehension of the spatial relationships between the paintings. And yet, the methods applied to relate spatial factors to this experience lack the

1 Determining 'threshold distances' at which the processing is affected would require a more focused, laboratory-based experiments, and most likely would only remain applicable to the particular set of artworks being studied. A similar work has been proposed on signage (Xie et al., 2007), but there the ability to read a particular font size is a much better benchmark of the viewing quality than what can be established for pictorial artworks.

capabilities of accounting for diverse engagement patterns (e.g. by considering interactions on a conditional basis 'if x then y'). It seems clear that the subject deserves further attention from the researchers, although its extensiveness would require a research procedure fully dedicated to these factors.

One more disadvantage of Space Syntax pointed out by Montello (2007) is that it remains ignorant of the overall look of the environment. This criticism can also be related to the current thesis, as no aesthetic aspect—–except for lighting—was taken into account. In the art gallery context these factors are often critical, especially when the 'white cube' atmosphere is desired (O'Doherty, 1986). For instance the issue of label placement has been a separate topic in museum studies (Bitgood & Patterson, 1993). Contemporary museum institutions often spend large proportions of their budget on detailed elements of the interior influencing the overall aesthetic atmosphere e.g. by replacing the whole floor surface to complete the 'pure' look of the space (Dorsett, 2013; personal communication). Such details were not considered in Experiments 1 or 2. Since the rooms used for Experiment 1 are otherwise used by art students as studio and exhibition space, their similarity to a 'real' art gallery is relatively acceptable with the only major compromising factors being its workshop-like character and university campus location. The space used for Experiment 2 however, was a typical psychological laboratory located inside a known research building. This, and the temporary character of the wall partitions could have contributed to a different perception of this space by its visitors. This atmosphere detached from 'real' galleries was even verbally reported by some participants at the end of their visit.

Concurrently, it must be noted that all compromising factors remained similar for all participants of this experiment, across both conditions. Thus, while it can be a factor decreasing the ecological validity of this particular study, the within- and across-conditional comparisons remain valid. It is also less likely that while the general perception of that space was altered, the specific patterns of individual visual interactions changed significantly. The alone fact that participants are aware they participate in a psychological study does not mean they will deliberately execute any other behavioural habit than the one natural for the situation they are at. In fact, it would be very unnatural to focus on anything else then typical 'picture viewing mode' in a space where empty walls are used to display artworks. The visual attractiveness of these walls remains therefore of secondary importance when configurational aspects of space are measured. This factor is much more likely to be damaging for studies interested in aesthetic appraisal of art, or the general appreciation of the exhibition setting. In this respect, also the criticism of Space Syntax approach can be reviewed - no theory is able to represents the world exactly as it is. Key aspect in applied studies of spatial cognition is therefore to select such theoretical concepts which are likely to correspond to those aspects of human behaviour and cognition which are important for the given research question. Being able to verify the majority of hypotheses stated at the beginning of this work is an evidence, that relevant concepts from Space Syntax theory were selected for this work.

26.2 DISTANCE BETWEEN THE VIEWER AND THE PICTURE

Turning back to the argument on ignorance of the metric distance in spatial measures we must also consider the 'cognitive facet' of the same problem. If spatial measures weighting the possible interactions by distance were used in the empirical studies presented in this thesis, theoretical explanations of their potential influence would require a dataset which includes viewing distances for each visitor-artwork interaction. This was not possible within the limitations of the available technology. With only raw video files being available, Simultaneous Localisation and Mapping (SLAM) has been proposed as a method for tackling this issue. The technique allows the software to estimate the location from which a gaze was performed and map both the location and the target of the gaze onto an ad-hoc constructed 3-D model of the environment. Up to date, no working software implementation is publicly available, despite a few prototypes proposed (e.g. Santner, Fritz, Paletta & Mayer, 2013). A method for manually encoding distance to objects from Tobii Glasses videos has been also proposed by Dalton et al. (2015) but is only viable for estimations at a specific point of the interaction. In case of the mentioned authors, this was the moment of the first engagement with an object. In the context of the current study and following the suggestions made in museum studies literature, it was predicted that this one point in time is not as important as understanding the holistic dynamics of the interaction throughout the visit. Nevertheless, informal observations of the eye-tracking recordings suggest that high correlation could be expected between short glimpses and larger distances as well as long dwells and shorter distances. For these reasons, distance to object was not explicitly coded in the dataset, and a dwell-based analysis was proposed instead.

It can be noted however, that the availability of a more advanced technology would open a variety of new research avenues, some of which could remain relevant within the art gallery context. Müller-Feldmeth et al. (2014) have demonstrated (again based on a manually encoded subset of eye-tracking data) how the availability of the viewing angle and distance can reveal behavioural patterns highly relevant for the studies of signage efficiency. The potential contribution of such work to the isovist theory, and its linkage with human cognition is immense and even the dataset studied in this thesis can be used again for such purpose, once software facilitating its automated analysis becomes available.

26.3 ILLUSIONS OF THE PATH-(IN)DEPENDENT APPROACH

This thesis has purposefully avoided utilising path-dependent approaches to the analysis of human interactions with the exhibition. It focused on measures which can be understood and quantified without the knowledge of the exact path a visitor has or will take through the exhibition. This has been done despite the fact that such path-dependent approaches had been proven beneficial to the exploratory power of environmental characteristics on human cognition in diverse contexts (Lu & Peponis, 2014; Meilinger, Franz & Bülthoff, 2012). The major reason for such an omission is its low practical value - path-dependent measures allow the researcher to make post-dictions, but not predictions of the layout's influence on human cognition. Making predictions within this framework would require to firstly predict (model) potential pathways inside the building only to calculate visibility and co-visibility characteristics of the exhibitions afterwards - based on these modelled pathways.

In reality, aggregate human movement inside buildings cannot be predicted perfectly well - the most prominent findings inside art galleries explained 86% of the variance in human movement (in terms of its room-by-room distribution, i.e. on a relatively coarse resolution; Hillier et al., 1996). This in itself is quite impressive, but it is difficult to expect that the immense variation in human walking preferences and decision making can be predicted any better any time soon. It also means that about 14% of that variance remains unexplained. For this reason path-dependent approach to making curatorial predictions would offer an 'illusion of accuracy'. It would offer better understanding at the finer detail of some interactions, but by no means would let us understand the wider spectrum of the visitor behaviour. The increased accuracy of prediction would be achieved within the most repetitive patterns of movement and interaction (which are, most likely, explained quite well in the path-independent approach) while the less predictable cases of interactions would still remain underrepresented within the explained variance.

It is the intention of this thesis to prioritise a more robust way of thinking about architectural cognitive engineering over presenting the largest possible effect sizes. This is due to the potential applicability of such robust methods in practically-oriented architectural and curatorial design workflows. This is not equivalent to saying that path-dependent approaches are less sound. Quite the opposite - their exploratory power can be greater and thus their scientific usefulness is often justified in cases where the main focus is on explaining the past processes, as opposed to predicting the future cognitive outcome of an exhibition.

However, even the path-independent approach bears its own 'illusion' of robustness. Creating and analysing measures which are 'path-independent' (such as *Visibility Catchment Area* or *Potential Co-Visibility*) assumes that human movement and human gaze is equally distributed throughout the entire space. This is not true. Some parts of the analysed space (e.g. room corners) are traversed through less then others (e.g. room centres). In giving this equal value to all parts of the spatial layout, this thesis prioritises the acknowledgment of uncertainty of the human movement over its non-perfect explanation biased towards more popular route choices.

26.4 issues with 'measuring' cognition

Measuring cognition can be problematic on multiple levels. One of them is that we are never really able to measure the actual cognitive process, but only its secondary outcomes, manifested in responses to abstract tasks. Each of such tasks carries its own measurement error which can influence the generalisability of the findings. For instance, Recognition Task employed in the study promotes a head-on view of the pictures. According to some findings (Tarr, 1995; Meilinger et al., 2012, 2011; Adamou, Avraamides & Kelly, 2014) this could result in faster reaction times not necessarily to pictures which were *remembered better*, but to those which were rather *viewed up-front* more. In such case, the influence of Visibility Catchment Area would be simply to increase participants scores in this particular task, and not to increase chances of *deeper processing* of pictures. Few theoretical assumptions serve as counter arguments to this concern. Firstly, pictures were shown to be processed relatively well from a very wide range of angles, especially when the information on background surface orientation is available (Section 3.3) - and this condition could not be fulfilled better than it is in a white-cube art gallery. Secondly, each picture was encountered multiple times, from many different angles. Object files theory (Section 3.1.4) suggests that such encounters can contribute to a single mental representation of an object, most likely represented up-front. This fact justifies the use of the test in the given context. The aspect here describes is likely to bear more significant influence if visual encounters were shorter or if only a single path tied to a limited range of viewing angles was available to participants, which is a limitation to always be considered in similar studies.

Another concern can be considered in regard to the Miniature Task. Like in any other sketch-map type of a task, some proportion of participants' error can potentially derive from the difficulty occurring during the transformation of the 3-dimensional experience into a 2-

dimensional frame of reference (Montello, 1998; Iwanowska & Voyer, 2013)

Another methodological limitation of this thesis is linked to the limited accuracy of the eye-tracking device. As described in Appendix A, the process of coding the eye-tracking videos is highly interpretative. In contrary to traditional observatory studies of museum visitors however, there is no systematic omission of visual targets encountered by the viewer. Especially during walking, observational studies were not able to account for this factor, and therefore are biased towards more longer and obvious viewing periods, concentrated in direct adjacency to the artwork. This bias is unlikely to be present under the method employed in this thesis. And yet, as has been previously mentioned, the analysis of eye-tracking recordings should not be taken for the equivalent of the visual attention as such.

Lastly, a less technical concern considers the limited nature of any set of measures applied to measure human cognitive experience. Like any other situation in life, museum visits are created by multiple sensorial experiences (Levent & Pascual-Leone, 2013a, 2013b). Visual memories are not the sole source of informational input, as their bounding in space is for instance strengthened by bodily movement (Zisch et al., 2013). For this reason, like any other study of cognition, this work by no means described anything more than very narrowly defined, indirect side-effects of how actually the space impacted its visitors.

26.5 DESCRIBING THE 'EXPERIENCE'

Coming from the positivist perspective, this thesis made the assumption that it is possible to define, measure, describe, and compare the human (aesthetic and visual) 'experience'. This, in itself, is a highly controversial notion with a long history of philosophical debate. Additionally, the work here presented fully (yet consciously) ignores previous experiences of the studied individuals. It also does not study how the experience of the exhibition develops and evolves over time after the visit, focusing only on the initial phase of the holistic experience of the exhibition.

It is therefore important to note that even though studied in fine detail, the 'experience' which this thesis describes is only a small part of a larger set of presumptions, attitudes, and memories which are affected and modified by each single art gallery visit - no matter how dynamic or intense the visual interaction between the visitor and the artworks it consists of.

26.6 A NOTE ON POWER

Statistical power (denoted as β) is the property of a statistical test describing the probability with which the test can detect a statistically significant effect (i.e. reject the null hypothesis) given the true effect exists in the population (Cohen, 1992; Descoteaux, 2007). The power of the test is dependent on its sample size (N), true population effect size (ES), and the accepted Type I error probability (denoted as α and in psychological science typically fixed at 0.05). Mathematically, the value of each of these four components can be calculated if the other three are known or estimated (Cohen, 1992).

All of the main three studies presented in this thesis investigated the same (or very similar) effects, although within distinct experimental designs. It is thus important to note the difference these experimental designs might have on statistical power of the conducted analyses and its potential consequences for the interpretation of their results. This is particularly important in the context of lower effect sizes resulting from the analysis of the BALTIC Case Study.

As mentioned previously (e.g. in Section 25.2), a number of factors might have impacted the predictability of studied phenomena in the BALTIC setting, to name the lack of randomisation of artwork content and artwork location, as well as its linkage permanently pre-defined by the curator to name but a few. This section will look at the potentially insufficient power of that design and re-evaluate the sample size necessary for increasing the chance of detecting potentially significant effects presented earlier in this thesis, under a real-world study conditions inside a public art gallery.

Since no well-established methods (other than simulations) exist to evaluate the statistical power of linear mixed-effect models (Castelloe & O'Brien, 2001), let us consider a simplified example referring to the subset of the previously conducted analyses. In particular, we will review the effect of *VCA* on *Recognition Memory* (quantified by participants' Reaction Times; Section 9, p. 125). This effect was well established in Experiments 1 and 2 but not in the BALTIC Case Study (see the comment to Hypothesis C6 in Table 25.1, p. 325). The sample size in Experiment 1, where the effect was found was $N = 31$ (compared to the BALTIC Case Study's $N = 16$). What would be the recommended sample size to guarantee a better visibility of this effect in the real-world gallery setting?

We previously concluded that the effect of *VCA* on *Recognition Memory* is well-demonstrated by the statistically significant *t-test* across the two experimental conditions employed in Exp. 1. The effect size estimated from this analysis is $d = 0.39$, in non-mathematical sense equivalent to small-to-medium effect size according to Cohen (1992). Post-hoc power analysis conducted in GPower software (S. Mayr, Erdfelder, Buchner & Faul, 2007) determines that the power for our test to detect such an effect size within this experimental design would be β = 0.18. It thus can be concluded that the researcher either got very lucky (having only an 18% chance to detect such an effect), or the true population effect is much larger. The upper limit of the 95% Confidence Interval around the calculated d = 0.39 was 0.78, suggesting that the true population effect is rather medium-to-large. And yet, in a real-world art gallery, under the presence of multiple uncontrolled factors introducing statistical noise to the measured processes (such as the permanent pre-defined linkage between the artwork location and its content), the effect size expected is even *smaller* than that demonstrated in Experiment 1. In order to exercise a more conservative approach, we can assume that we wish to detect d = 0.39 in a between-group design (assuming it would be possible to re-create the experimental design from Exp. 1 inside the BALTIC Centre for Contemporary Art). Aiming for the chance of detecting the main effect equal to β = 0.8, this would require two groups consisting of 105 participants each (or 210 in total) compared to 16 engaged in the here-described Case Study.

Statistically 'noisy' research settings (such as buildings in public use) call for much larger sample sizes in order to detect effects relating to cognitive processes impossible to observe and measure directly (such as memory, or the depth of cognitive processing it is indicative of). This is the primary reason for which highly controlled, laboratory-based studies will remain the key stage in research projects aimed at establishing environmental correlates of cognition. The vision of sensory-enriched buildings and cities, capturing large quantities of observational data in real-time, offers the possibility of easier, cheaper, and much more reliable verification of these effects in applied settings.

27

CONCLUSION

The main research question of this work was whether spatial arrangement of an art gallery has an effect on the visitor experience. The main hypothesis stated that the influence of artwork's spatial location is often larger than that of the artwork's content and that it can be used to partially predict the cognitive impact of an exhibition. Three distinct studies showed that the initial assumption is true. After controlling for other factors, many spatial characteristics had an impact on visual attention and memory of the gallery visitors. Some of those effects were strong enough to remain significant in a much less controllable real art gallery setting.

As a result, the contribution of this thesis to the current state-of-knowledge can be classified in two generic points:

1. The contribution to curatorial theory and observational visitor studies. Applying a relatively advanced methodology allowed the investigator to measure the cognitive (and not only observable/behavioural) outcome of the gallery visit. This outcome has then been linked to generalisable, quantifiable aspects of space and can be applied to potential simulations of yet-unbuilt exhibition layouts.

2. The contribution to the fields of Spatial Cognition and Space Syntax. Ever since Space Syntax was introduced, the interest in linking its measures with empirically proven aspects of human spatial cognition was central to the field. The potential usefulness of such a linkage for both disciplines has been widely acknowledged (Montello, 2007; Penn, 2003; Conroy Dalton et al., 2012) and some empirical studies followed in this domain (e.g. Conroy Dalton, 2003; Franz & Wiener, 2008). By creating a direct empirical connection between some aspects of human cognition relevant to the museum experience and spatial measures describing the gallery set-up, this work is making a further step in closing this gap, at least in a limited museum context.

Both of these points can rise a question whether (1) curatorial and (2) architectural theory need a similar kind of contribution and whether it does not ultimately lead to 'architectural determinism' (Franck,

Figure 27.1.: All 'star' objects of the imaginary exhibitions (all other artworks are ommited in this visualisation). Red dots: curators /architects /artists; Green dots: others.

1984) potentially restricting the variability and exciting novelty of our built environment. To some extent, this assumption can be tested. In a study supplementing the main empirical work presented in this thesis, a group of 26 participants (19 of whom were artists, curators, or architects) were asked to place 11 'artworks' on a miniature layout of an imaginary exhibition they were asked to 'curate'. The 'artworks' were represented as dots, 3 of which had a different colour and were specified as 'artworks of a special meaning', or the 'star objects' of the exhibition. The same task was conducted twice, each time on a different layout. Each participant performed the task individually, on separate copies of the layout. One of the layout was based on the setting employed in this thesis in Experiment 1, and the other one on the layout of the BALTIC Case Study. Figure 27.1 presents the summary of the results, where the positions of the dots were 'stacked' on top of each other when two or more participants placed the dot in the same location of their layouts.

The uniformity of the results is striking. As it can be noticed, the 'star' objects are not dispersed among all the walls and the similarity of choices is even greater when to ignore the contribution of the laymen (green dots). This preliminary data suggests that there is some conceptualisation of the 'importance' of various elements of the gallery layout which is common to many curators even when they are asked to think about it independently. Is it possible that those choices were based on some underlying formal spatio-visual characteristic of the chosen location? Indeed, when the mean isovist area sizes of the 'star' locations were compared with mean isovist area sizes of 'all other' artworks (not displayed here in the figure), the analysis revealed that 'star objects' were placed on more visible locations in Layout 1[1]. It thus seems that—while planning exhibitions—curators already have assumptions about what stretches of the gallery wall have a higher 'value'. It also seems that this 'value' can be perceived based on intuitively detected spatio-visual properties, and that the (abstract, imaginary) choices about the exhibition set-up are made in respect to this assumption.

So the curatorial practice (as demonstrated above), as well as curatorial theory (as exemplified by a long tradition of visitor studies), already follow some form of 'architectural determinism'. And this fact has not—at least so far—jeopardise the joy of novel, unexpected experience that we associate with a visit to an art gallery. The fact that these connections between space and cognition are sought and exploited does not predetermine the exact way they will be used. Even just to build a completely 'unpredictable' space, the designer needs to understand what its 'predictable' version would look like.

In regard to the wider architectural context, this thesis made a claim that buildings can be studied as other physical artefacts are in Cognitive Engineering, Usability, and User Experience studies. A claim which carries the assumption that spatial design can ultimately be better adjusted to our cognitive processes, or that the entire design process can in fact be centred around the desired cognitive experience. None of these ideas are new. Designing the experience by considering formal properties of space has been proposed in the seminal paper popularising isovists by Benedikt (1979) himself, and in many different forms extended later (e.g. Stan Allen, 1996) up to a point where software tools facilitating such a design process are becoming available for architectural (Schultz & Bhatt, 2013) and urban planning applications (Schneider & König, 2012). To what extent will this shift towards experience-oriented architectural design affect the flexibility and creativity of architects' visionary ideas? An excellent example of how this can be a mutually beneficial relation comes from the field of Human Computer-Interaction. To describe it, Nielsen (2013) has used an example of opera singers. Despite the fact they perform under narrow constrains of a screenplay which has remained unchanged for hundreds of years, no person in the audience has any doubt as to the creativity, skill, and talent involved in creating their unique performance each single time they go on stage. The same type of 'constrains' applies to usability and design. As he put it, the study of usability barely documents reality (Nielsen, 2013). Therefore just as much as an architect must consider physical laws which allow the building

1 Based on the analysis of Isovist Area sizes. Paired t-test: $t(26) = 2.46$, $p=.021$

to stand stable, cognitive issues which have an impact on our architectural experience need to be built into thinking about design at all stages of the process. The potential benefits to the quality of our built environment coming from such a shift in thinking is difficult to predict. But, as Montello (2014) wrote on the same issue (p. 79): "when even imperfect scientific understanding has modest implications for so many people in so many places for so many hours, its considerable value becomes obvious."

27.1 FUTURE WORK

Further work is planned towards the extension of the impact of the findings described in this thesis. More empirical work is required to refine the understanding of Visibility Catchment Areas in relation to cognitively-demanding spatial tasks. As it was mentioned early in this work, Visibility Catchment Areas were defined in this work based on a limited amount of empirical data. The simple fact that in many cases this modification of the classic Isovist concept had improved the understanding of the visual interactions with art, calls for more focused work in this respect. How Isovists can be refined to better reflect our natural strategies of visual interaction with the external environment is a questions which has recently gathered increasing interest, for instance with respect to the third dimension (Derix, Gamlesæter & Carranza, 2008). As it was mentioned previously, this will become an even more fruitful field of research when a more refined eye-tracking technology becomes available (Müller-Feldmeth et al., 2014).

With respect to the curatorial domain, a software tool is planned which would apply some formal 'rules' described in this thesis onto a newly planned exhibition layout. This is aimed to be a guide for curators interested in designing their exhibition with a pre-defined cognitive outcome in mind, although due to the limitations listed in Section 25.3, it is never expected to be a stand-alone tool providing deterministic answers to design challenges. Further work with curators and artists is planned to build the tool around their actual design needs, as well as to provide the output of the architectural analysis in a form intuitively understood to an interested layman. As our understanding of the cognitive impact of space expands, it is a justified research aim to make this knowledge inclusive for other interested professions. Curators empowered with a tool for architectural analysis might in fact be able to generate much more reliable predictions, given their expert knowledge of multiple other factors that might have an influence on the visitor experience but were not considered in this thesis (see Section 26.1).

Further, the project here described can be a point of departure for considering cognitive usability of other building types. On the most generic level, the set of methods employed in this work has been an exercise in defining, measuring, and interpreting those aspects of human cognition which might be relevant to the architectural experience in the pre-defined context. The number of buildings which could be improved in how they interact with our cognitive processes is immense and for each of them the set of relevant psychological and architectural methods is likely to differ. Our definition of what building usability is, how it can be measured and improved is lacking the level of understanding which would allow other researchers and practitioners to replicate and improve it in a reliable, comprehensive manner. It is my hope the this thesis improves our understanding of the issue in its narrow context, but also methodologically paves the way for similar investigations in a wider group of future applications.

As architectural settings become larger and more complex, the distinction between indoor and outdoor becomes less obvious (Kray et al., 2013). Our knowledge of similarities and differences in spatial cognitive processes employed inside and outside buildings, as well as the quickly progressing advancement in the available technology makes it possible to employ similar research techniques to studying buildings and cities. As our urban environments become denser occupied by visual stimuli of varied importance (such as public art, large- and small-scale advertisement, signage) the linkage between its positioning in space and the probability of its successful processing becomes an indicator of how well the urban governance is able to distribute its limited 'visual real estate' to balance out the interests of all parties. In comparison to this, curators enjoy an almost unlimited freedom in managing the empty area of the gallery's wall, while the outdoor urban context imposes many more spatial, social and political limitations on those decisions, making the choice between the accessibility and placement of different types of visual information potentially much more difficult.

Evidence-based decision-making can thus answer to the cognitive needs and preferences of the population on multiple planning scales - from building to city scale. The combination of increasingly user-centric thinking among the planners, more scientific data paving the way for the understanding of environment-cognition interaction, as well as 'big data' omnipresent in smart buildings and smart cities of the future; the combination of these factors creates the possibility for developing novel workflows for planning and managing our built environment. Workflows which will emphasise user-friendliness and responsibility for the citizens' cognitive wellbeing in response to their own needs and behaviour.

And yet, such a vision does not guarantee that buildings and cities of the future will be any more inclusive and more democratic than they are now. Just as much as a (more or less user-centric) city planner or an architect makes top-down decisions about the way our built environment is shaped today, so much such a decision will be shaped top-down in the future. Although identifying the moment when the actual decision is made might not be as easy as it is today because in semi-automated planning and management processes of the near future those decisions will be made by algorithms. Very complex algorithms, designed not by a single person, or even a single architectural office, but by large dispersed research teams. Those algorithms will be capable of reacting, often in real time, to the needs of the building/city users in response to their behavioural patterns. Public displays can automatically modify the type of the displayed information based on the characteristics of the user group currently passing by (Ballendat, Marquardt & Greenberg, 2010); building's heating systems might modify the temperature depending on the type of activity sensed in its different areas (Kauppinen, Litvinova & Kallenbach, 2014; Göçer, Hua & Göçer, 2015). What remains an often ignored, highly relevant, and very timely research question is how to adjust these algorithms so that they do not always aim for the most 'efficient' and 'computationally favourable' solution, but take into account cognitive biases of the citizens they serve: e.g. human preference for less distraction over information overload, or the feeling of control over one's own immediate environment. Only such an approach, pursuing which the author of this work is fully dedicated to, can ensure that our smart, 'cognitively engineered' buildings and cities become truly *user friendly*. With the rise of crowdsourced approaches to data gathering, and given a careful consideration of all user groups (currently often underrepresented in environmental research), those buildings and cities also finally present hope for becoming truly *democratic*.

Part VIII

APPENDIX

A

EYE-TRACKING: CODING PROCEDURE AND RELIABILITY

For each participant, Tobii Glasses provide a video clip recorded from the egocentric perspective via a camera mounted on the device frame. An independent infrared camera pointed towards the user's pupil records its movement. This is then superimposed on the egocentric video clip based on the coordinates obtained during the calibration process, which proceeded each participant's gallery visit.

During data analysis, each eye-tracking recording was manually coded for dwell occurrences, as no reliable automatised method suitable for the context of the study is available. The fact that the infrared camera tracks only a single eye creates the effect of *parallax*, meaning its estimation of where the participant was fixating is calculated independently of the actual distance to the target stimulus. As distance between the viewer and the pictures was constantly changing throughout the visit, the parallax effect could not be fully accounted for by the software. Instead, a single assumption of the the most likely viewing distance must have been provided to the Tobii software package for these estimations. In experiments here described, the 'most likely distance' was set as 'large (>150cm)' in Tobii Studio's 'parallax correction tool'. The choice was based on informal observations of museum visitors at an exhibition containing pictures of similar size. The resulting disadvantage affecting the data analysis is the fact that the eye-tracker's estimated fixation locations might be imprecise for any event when the distance to the target was notably different from 150 cm.

For this reason, the process of coding oculomotor dwells from Tobii Glasses recordings contains an interpretative component. Especially for oculomotor events occurring simultaneously with head movement, the coder's judgment was more likely to be accurate than the actual position of the dot visible on the screen. The next section of this Appendix describes the procedure incorporated to code and interpret the eye-tracking recordings. It is followed by a section describing how internal reliability of the procedure was assessed agains a subjective bias of the researcher performing the coding.

A.1 CODING PROCEDURE

Each video contained a recording from the front camera mounted on the eye-tracking device, with a superimposed red dot indicating the eye-tracker's estimation of where the person was fixating at any given moment. Saccades were indicated by lines connecting the dotes. Each dot was visible on the video between 250 and 500 ms (depending on the specific settings), after which it disappeared. Visibility of these dynamic indicators allowed the coder to estimate whether a change in fixated objects was not an artefact caused by the participant's rapid head movement.

CowLog software (Hänninen & Pastell, 2009) was used to code the recordings manually. Default .avi output from Tobii Studio was converted to .mp4 file format compatible with CowLog. Each video was played at 0.3 speed rate. This additionally allowed to take head movement into account, in opposition to frame-by-frame analysis. Due to the fact that videos are recorded by Tobii Glasses at a frame rate of 15 frames par second, this playing speed displayed approximately 5 frames per second. Pausing and rewinding the video clip was possible to ease the coding process during the most rapid series of dwell changes.

Each time the dot estimating a fixation entered the borders of a painting, the letter associated with this painting was pressed on the keyboard and saved to a text file by CowLog. Buffer area surrounding each picture, estimated at the size equal to 1 /3 of the paintings' width was also counted as a correct dwell. This was dictated by the observation from preliminary analyses suggesting that the parallax effect might account for many of such events. It is improbable that participants intended to investigate the wall surface in the nearest proximity to the picture for periods often longer than 10 seconds. Additionally, a valid 'dwell' was only counted if 2 dots estimating fixations appeared within the defined boundaries, as lighting reflections often caused the device to record false, single rapid fixations around the boundary of the viewing field.

Distant stimuli, often unrecognisable under poor lighting conditions and low resolution were interpreted by the coder on the basis of their known spatial location in the given gallery setting.

Entrance and exit were coded to calculate the exact time spent inside. Unless reported otherwise, this was counted from the moment each participant opened the door and made his /her first fixation inside the gallery space (or outside of it, in case of exiting).

Each valid dwell (i.e. longer than 2 dot occurrences) happening outside of any picture's boundary was coded by the letter 'x'. This made it possible to calculate the cumulative time spent looking at empty wall surfaces, ceiling, floor, or minor distractors present in the space (light switches being one of the main attractors).

In Experiment 2, each fixation on the mirror was coded with the letter 'q'. Looking at pictures' reflections in the mirror was coded as a regular picture dwells.

In the BALTIC Case Study analysis, dwells on other people present in the gallery were coded with the letter 'q'. Dwells on labels could not be reliably distinguished from dwells on the walls due to low resolution of the recordings and white colour of the walls reflecting strong light. Since not all artworks present in the gallery were subject of further analyses (e.g. sculptures and artworks from the far end subspace of the gallery), fixations on those were jointly coded with a single letter 'v'.

A.2 INTERNAL RELIABILITY

The coding process was largely based on subjective estimates of dwell occurrences. This method was applied in order to correct for the parallax effect and low frequency of fixation estimates resulting in multiple artefacts recorded throughout the video. Manual coding of this type of data bears the risk of unreliable estimates, not only biased by the technological defects, but also by factors randomly affecting the coder's judgment. These factors might include tiredness or involuntary changes to the understanding of the coding protocol.

Interrater reliability testing is a method allowing the researcher to control for anticipated bias in subjective judgements. For each experiment, a subset of the video recordings was independently coded by a second coder—a graduate psychologist trained in the coding procedure described in the previous subsection. The training session was conducted prior to the coding and no further consultancy was available during the procedure.

To assess the coding reliability in a dataset consisting of multiple nominal categories provided by two coders, Cohen's Kappa (Cohen, 1960) is a broadly recommended (Kline, 2005; Huck, 2009) and most typically used approach. Stemler (2004) however notes frequent misuse of distinct interrater agreement techniques and divides the available methods into three categories based on their purpose: consensus estimates, consistency estimates, and measurement estimates. The major difference concerns consensus and consistency estimates. Consensus estimates (such as Cohen's Kappa) are relevant for a situation in which it is possible to come to an *exact* agreement about the nature of the event, which would indicate that both judges share *the same* interpretation of the construct they are rating. In the light of technological limitations imposed by the device used in this thesis, this is unlikely. Low video frame rate, low fixation sampling rate, and the parallax effect result in multiple imprecise estimates superimposed by the eye-tracking software on the egocentric video. For this reason, judgements of the oculomotor behaviour (especially during rapid head movement periods) is subjective - based on the number of fixations, their distribution, and the presence of stimuli most likely attracting eye movement. Due to the number of such events, achieving consensus by the judges would be highly time-consuming and despite the effort, most likely would never reflect the factual oculomotor event. For this reason, it is important to emphasise that all measures reported in this thesis are subjective estimates—although most likely they are more accurate from external observations (J. K. Smith & Smith, 2001)–especially in relation to shorter dwells–but by no means they represent the *exact* viewing times.

The use of consistency estimates is also recommended if there is a consistent bias likely to occur in the ratings of the individual judges (Stemler, 2004). In this research procedure, CowLog recorded key strokes with the accuracy of tens of milliseconds. The software was therefore sensitive to individual differences in reaction times across the coders[1]. The judgement of buffer area around the picture counting as a valid dwell could have also been affected by the interpersonal differences in distance estimation across the judges. Both of these factors are likely to cause individual judges' ratings to differ from identical. Consistency estimates of interrater agreement allow the researcher to assess if this bias was consistent, i.e. if each rater was able to follow subjectively conceptualised rating scheme. The oculomotor measures are only used in this thesis in relative comparison within each experiment. The exact number of seconds people look at artworks is not of interest, as this is likely to slightly differ across each museum space and each particular stimuli set. Subjectively biased estimates are therefore fully satisfactory for the purpose of this work, as long as they are internally reliable, i.e. biased randomly across the data set, so that each picture was likely to be estimated with equal amount of bias.

For the purpose presented above, Stemler (2004) recommends calculating correlation coefficient for scores of two judges. The resulting values grater than 0.7 are accepted as interrater agreement. Where the distribution of ratings violates the assumptions of Pearson's correlation coefficient, Spearman's rank coefficient can be used (Stemler, 2004). As a non-parametric measure, it is likely to present more conservative result.

For this procedure to be valid, the second rater only needs to assess a subset of the complete dataset. The exact percentage of full sample which should be verified by the second coder varies across the literature (Sauppé & Mutlu, 2014; Franchak, Kretch, Soska & Adolph, 2011). Specific recommendations suggest 10% of the whole data set (Lombard, Snyder-Duch & Bracken, 2004; Neuendorf, 2002), considering that verifying a larger number of data points does not usually bring any substantial changes to the score. For each of the three eye-

1 One coder might have been consistently quicker than the other.

tracking datasets presented in this thesis (derived from three studies), a random participant file was drawn for the second coder to rate. Files were drawn until their cumulative length reached or exceeded 10% of the total length of all files in the given dataset. Second rater performed the coding procedure on the selected sample of recordings. The data obtained from the coding was then used to derive the eye-tracking dependent measures described previously in Section 8.2. The number of data points therefore was equal to the number of keys used in the coding procedure multiplied by the number of sample video recordings. These data points were then correlated with the data obtained from the same recordings by the researcher. The same R script was used to derive the dependent measures from the data coded by the second rater.

Table A.1 below presents results of the analysis in respect to each measure controlled, for each dataset.

	Number of Dwells	Norm. No. Dwells	Total Dwell Time	Norm. Total Dwell Time	Time to First Fixation
Exp. 1 (df = 43)	*.93*	*.82*	*.99*	.98	*.58*
Exp. 2 (df = 54)	*.93*	*.67*	*.99*	.99	*.98*
BALTIC (df = 42)	*.99*	*.97*	*.99*	.99	*.99*

Table A.1.: Interrater reliability comparison. All results were significant at $p<.001$. Result in italics indicates that Spearman's rho was used instead of Pearson's r where non-parametric procedure was more suitable.

Lower correlation values on 'Time to First Fixation' were the subject of further investigation. The analysis of scatterplots revealed that the value was low due to 3 outlying data points. While one of the coders recorded a very early encounter with a given artwork, the other did not code it, until a much later moment in the video. This fact most likely reflects the ambiguity of dwell occurrences during quick scanning of the gallery at the beginning of participant's visit.

Nevertheless, the majority of the compared video codings remained scored highly on the interrater reliability measure. Despite being a subjective process, manual eye-tracking analysis conducted within a strict set of coding rules proven to be reliable method of data encoding.

B

CUSTOM-BUILT STRING MATCHING ALGORITHM FOR MINIATURE TASK ANALYSIS

In most of String Matching algorithms, the similarity of strings is scored according to the number of actions required to transform the answer into a reference string. Possible actions are *substitution* (e.g. substituting F for C), *insertion* (i.e. adding a character into the string), *deletion* (removing a character), and a *transposition/swap* (when two adjacent characters change places, e.g. 'A-C-F' $->$ 'A-F-C'. Separate algorithms differ by the exact actions they allow, how they score their occurrence, and what order the actions are performed in. In this thesis, String Matching will only be used for situations where the length of the participant's answer and the length of the reference string are equal. Thus, the only actions required to assess participant's memory performance are *substitution* and *swap*. The order and 'cost' of these actions is also important: *swaps* should be performed first, so that the participant is punished with smaller amount of negative points for situations where a correct picture was placed on the correct wall, but in a wrong place relative to the entire sequence. *Substitutions* should only be executed afterwards, and at a higher cost of negative points. This would reflect the assumption that spatial memory of those participants who placed pictures closer to their original location is higher from those who did not include them in the given wall section, or placed them further away in the sequence.

Classic String Matching algorithms contain two solutions similar to the required purpose: firstly, Damerau (1964) identified all four operations listed above in a single spellchecking algorithm, suggesting they account for over 80% of misspelling errors. However, this only accounted for string transformations which can be achieved with the use of a single action. The use of multiple edit actions was introduced by Levenshtein (1966), although it did not consider the *swap* action. Damerau-Levenshtein Distance is known as the combination of the two: it allows all four operations to be used multiple times, until the input string matches the reference string. In it, each operation is scored as 1, but in its weighted version (Weighted Damerau-Levenshtein Distance) discrete scores can be associated with each action type. This would allow to score swaps lower than substitutions, according to the need of Miniature Task analysis described in the former paragraph. However, Weighted Damerau-Levenshtein Distance allows each character to be swapped only once. If such operation does not match it with the desired location in the reference string, two substitutions are performed instead. In the context of the Miniature Task, this is not the desired operationalisation as the assumption we wish to implement is that the memory of a participant who placed a picture 2 locations from its correct location is better than the one who placed it 6 places away. Weighted Damerau-Levenshtein Distance would punish both of such cases on the equal basis.

C

FORMS

Below examples are based on the material used in Experiment 1. Material in Experiment 2 and the BALTIC Case Study was modified as little as it was possible.

C.1 INFORMED CONSENT FORM

School of the Built and Natural Environment

RESEARCH PARTICIPANT CONSENT FORM

Name of participant	
Organisation	**Northumbria University**
Researcher's name	**Jakub Krukar**
Title of research project/dissertation	
Researcher's name	
Programme of study *[Only if researcher is a student]*	**Full-Time PhD**
Supervisor's name *[Only if researcher is a student]*	**Prof. Ruth Conroy Dalton**

Standard statement of participant* consent (please tick as appropriate)

I confirm that:

I have been briefed about this research project and its purpose and agree to participate*
I have discussed any requirement for anonymity or confidentiality with the researcher**
I agree to being audio taped / videotaped during the interview

* *Participants under the age of 18 normally require parental consent to be involved in research.*

****Specific requirements for anonymity or confidentiality**

Signed ____________________________ **Date** ____________

Standard statement by researcher

I have provided information about the research to the research participant and believe that he/she understands what is involved.

Researcher's signature ……………………………………….

Date ……………………………………….

Thank you for your interest in this study which is being carried out by Jakub Krukar, a PhD student at Northumbria University. The aim of this study is to investigate associations between spatial layout of art galleries and human perception of the exhibition.

It is expected that this experiment will last no longer than an hour and in return for your time you will receive 6 pounds.

The data which we gather will be treated in the strictest of confidence. It will only be analysed in aggregate. Results of this research may be published or reported to government agencies, funding agencies, or scientific groups, but your name will not be associated in any way with any published results.

If you have any questions or require any further information, please don't hesitate to ask. You are free to withdraw from this study at any time whereby the data you have provided will be deleted. If you wish to withdraw please contact the researcher, Jakub Krukar, at **jakub.krukar@northumbria.ac.uk**.

This research has received ethical approval from School of the Built and Natural Environment at Northumbria University. If you require confirmation of this please contact Director of Postgraduate Research at eric.johansen@northumbria.ac.uk.

I understand the purpose of this research and agree to participate.

I understand I am free to withdraw from the study at any time.

....
(name)

....
(signature)

C.2 DEMOGRAPHIC INFORMATION FORM

UNIQUE ID:

....

ENTRY TIME:

....

EXIT TIME:

....

HAVE YOU BEEN IN THIS BUILDING BEFORE?

yes no

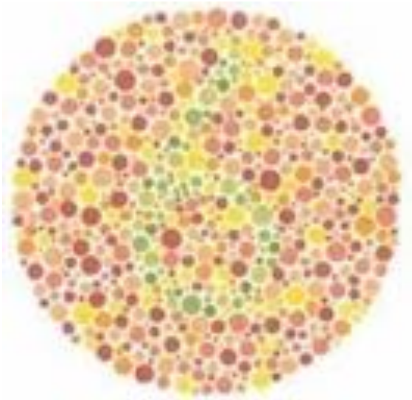
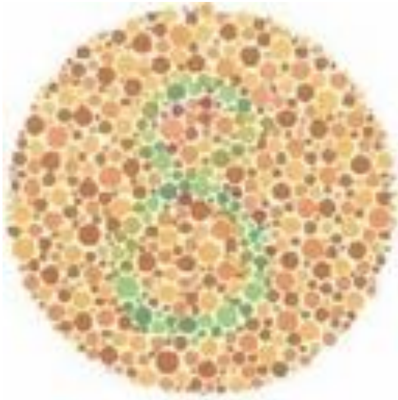
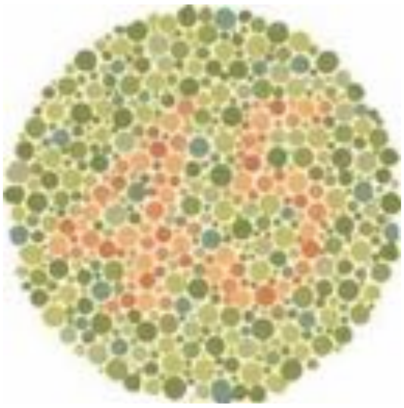
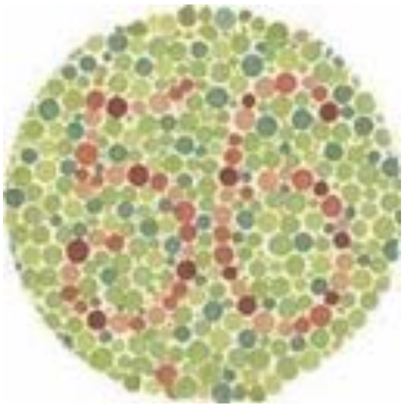
GENDER:

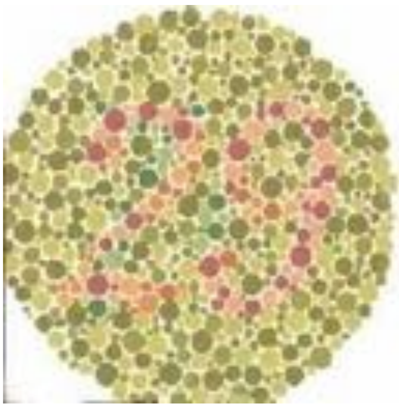
male female

AGE:

E-MAIL ADDRESS:

C.3 EXCERPT FROM COLOUR BLINDNESS TEST



D

INSTRUCTIONS

Below examples are based on the material used in Experiment 1. Material in Experiment 2 and the BALTIC Case Study was modified as little as it was possible.

D.1 PRE-MEETING INSTRUCTION (E-MAIL)

```
Dear [name],

Thank you very much for your interest in my research.

In this study you will be asked to spend some time in
a small art gallery with a portable eye-tracking device
(which in practice looks like a pair of large glasses).

There are just a few questions I need to ask you first:
1. Do you have normal colour vision?
2. Do you wear glasses or contact lenses? (it's fine if
you do but glasses frames can't be very large)
3. Are you a student in Fine Arts or Architecture
or are you a professional artist/architect?


You will also need to bring your bank account details with
you because the payment (6 pounds) will be transferred
by the University directly to your account (i.e. account
number, sort code, name of the bank, branch address).
To avoid disappointment, please note that all earnings
will be taxed at Basic Rate (unless you provide a P45
or send a P46 available on request from University's
Payroll Office on 0191 227 4767).
Also, please bring your National Insurance Number
(or send it via email to the researcher). If you do not
have a National Insurance Number (e.g. international
students) your payment will not be taxed.
[NOTE: Above paragraph was removed in further studies
where participants were paid in cash.]
```

```
The experiment will last no longer than an hour but you
shouldn't be late. It will take place at Northumbria
University's City Campus.
Could you meet me there at [time] ?
If not, feel free to suggest a different date.

I will send further directions upon your response.

Best regards,
Jakub Krukar
```

D.2 Briefing instructions (prior to entering the gallery)

```
"Now you will enter to an art gallery. Explore it as you
would usually do it in an art gallery. Make sure you
enter to each room and see all the pictures.
There's a door in the end. It's open. I'll be on the
other side of it. You have maximum 30 minutes but
you don't have to wait until the end of that time.
I will let you know when the time is running out."
```

D.3 Recognition test instructions (on-screen)

```
[spoken]
"Now I have another task for you.
Everything will be written on the screen.
Please use only the 2 labelled keys".

[on-screen]
Welcome to the experiment!

You will see a series of 28 pictures.

Half of them are EXACTLY the same as the ones
you've seen in the art gallery.
Another half is either completly new or a modified
version of pictures from the gallery.

Your task is to indicate whether you have seen the
picture in the gallery or not.

Press 'no' key if you think you have not seen it.
```

```
Press 'yes' key if you think you have seen it.

Try to be as fast and accurate as possible!

First 3 pictures are for training purposes only.
Place you fingers on the keys and press any of
them to continue.
```

D.4 MINIATURE TASK INSTRUCTIONS (SPOKEN)

```
"Ok, thanks for this. There's one last think
I will ask you to do. These are miniatures
of the pictures you have really seen inside.
And that's the layout of the gallery.
Can you please try to arrange the miniatures
by placing them where you think you saw
them in the gallery?"
```

D.5 SALIENCE STUDY INSTRUCTIONS (ONLINE; IN POLISH)

```
Welcome and thank you for your participantion.

This work is conducted within doctoral studies of Jakub
Krukar at Northumbria University in Newcastle, UK.
The project has received acceptance from the university
ethics comittee. All answers will remain anonymous and
will be only used in aggregate.

The study will take you only a few minutes. It consists
of four questions. In each of them I will ask you to
sort a pool of pictures according to how much they
draw your attention.

In case of any questions you can contact me at
jakub.krukar@northumbria.ac.uk.

If you agree to participate, please make sure you will
not be distracted at your computer for the next few
minutes and click 'NEXT'.

=======

On the next page you will see a list of paintings
```

```
arranged vertically.

Please sort them according to how much each of
them draws your attention.

On the top of the list (marked as number 1) place
the painting which draws your attention the most.
On the bottom of the list place the one which draws
it the least.

You can drag and drop the images using the mouse.

After completion click 'NEXT' at the bottom of the
web page in order to progress to the next screen.
You will see 4 screens in total, each containing
3-4 pictures.

Click 'NEXT' to continue.
```

D.6 SALIENCE STUDY SCREENSHOT (ONLINE)

Figure D.1 below presents a screenshot from the Salience Study. The only other elements visible on the screen would be the tool bars of the particular browser used by each participant.

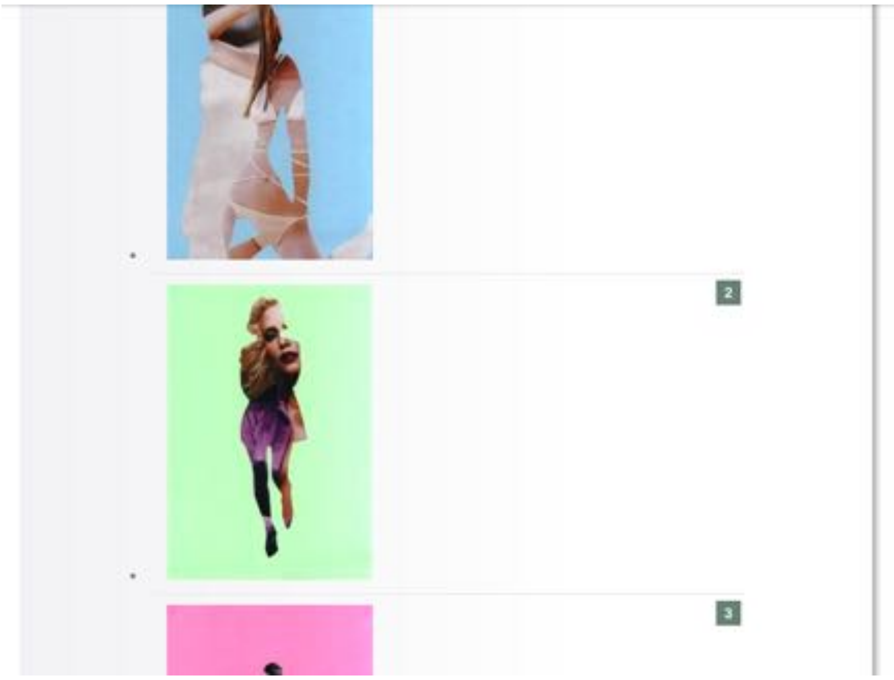


Figure D.1.: Screenshot of the Salience Study. Participants were able to drag and drop images to position them in the desired vertical order. After scrolling to the bottom of the screen they could see a button labelled 'NEXT' which progressed the study to the next set.

E

STIMULI USED IN RECOGNITION MEMORY TESTS

Figure below presents images used as distractors in the Recognition Memory Test ('correct' stimuli were already presented in section 13.2). Distractors used in the Recognition Memory Test in the BALTIC Case Study were a combination of 'correct' and 'distracting' stimuli from Experiments 1 and 2. Some of the distractors are distinct pictures coming from the same artistic set, while others are modifications of the 'correct' stimuli modified by manipulating colour filters in image editing software.

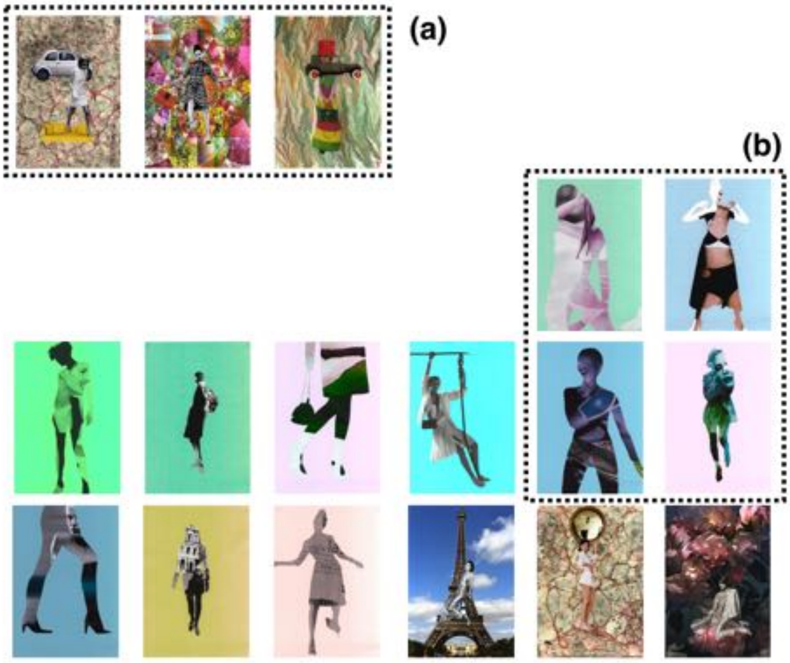


Figure E.1.: All distractors used in Recognition Memory Test in Experiment 1. Subset (a) was used in the training phase only. Subset (b) was excluded from Experiment 2 in order to keep the numerical balance between 'correct' and 'false' stimuli (two paintings originally present in Exp. 1 but excluded from Exp. 2 were also used there as distractors).

F

GLOSSARY

Back-to-the-Wall - a measure of *Spatial Memory* defined by the proportion of pictures which were placed back on their correct wall in the *Miniature Task*; Section 10, p. 129.

Bidimensional Regression - a measure of *Spatial Memory* defined by the correlation of two-dimensional coordinates between picture miniatures' placements and their reference locations; Section 10, p. 129.

Dwell - the time a person spends looking at a single object (e.g. an artwork) without glimpsing outside the boundaries of this object; Figure 3.1, p. 44.

Dynamic Viewing - the act of *engaging* with multiple objects (e.g. artworks) by multiple short *dwells* (as opposed to less longer *dwells*); Section 8.2, p. 113.

Ecological Validity - the degree to which the experimental set-up reflects the equivalent real-world context it aims to represent; mentioned e.g. in Section 3.3.

Engagement - the interaction between the viewer and the artwork defined by the act of looking (dwelling) on a single object, by a single participant; Section 3.1, p. 42.

External Validity - the degree to which the experimental set-up and measurement results generalise to other experimental set-ups aiming to study and describe the same phenomenon; mentioned e.g. in Section 4.3.

Eye-Tracker / Eye-Tracking - a device based on a pair of glasses with two cameras which measures the movement of the eye and estimates the gaze allocation; Section 8, p. 111.

Fixation - the period of suspended movement of the eye (lasting a fraction of a second), during which the visual information is perceived and processed; Section 3.1.2, p. 43.

Isovist - a geometrical shape describing the area on the floor plan from which a given point (e.g. an art object) is physically visible; Section 5.2, p. 83.

Miniature Task - a task designed to measure *Spatial Memory*, in which the participant is asked to place miniature representations of the artworks originally seen in the art gallery on a miniaturised representation of the spatial layout of that gallery; Section 10, p. 129.

Potential Co-Visibility - a measure describing the number of other pre-specified objects (e.g. artworks) visible from the location of the target object. Derived based on the total *Isovist* area calculated from the location of the target object; Section 11, p. 137.

Reaction Time - the number of milliseconds it takes the participant to recognise an object within the computer Recognition Memory test; Section 9, p. 125.

Recognition Memory - the ability to recall a previously encounter object during its later viewing, i.e. during a computer Recognition Memory test; Section 3.2.1, p. 48.

Saccade - the rapid movement of the eye between the fixations during which the visual information is not processed and does not contribute to the memory of that object; Section 3.1.2, p. 43.

Spatial Memory - the ability to recall the position of multiple objects after they have been encountered in space; Section 3.2.2, p. 51.

Statistical Model - an equation listing 'predictors' (or 'Independent Variables') which have an impact on the studied 'output' (or 'Dependent Variable'); see caption under Table 15.1, p. 174.

Statistical Power - the probability of detecting the studied effect by the particular statistical technique, in the particular experimental setup, given the true effect exists in the population; Section 26.6, p. 339.

String Matching - a mathematical technique of comparing the similarity of two letter strings based on the number of operation it takes (given a set of predefined rules) to transform one string into the other. In this thesis used for the analysis of Viewing Sequence Similarities (Section 8.3) as well as the Miniature Task (Section 10.1).

Targeted Co-Visibility - a measure describing the average number of other pre-specified objects (e.g. artworks) visible from the location of the target object's *Visibility Catchment Area* (i.e. within a pre-defined angular range). Derived based on the *Visibility Catchment Area* calculated from the location of the target object; Section 11, p. 137.

Visibility Catchment Area - the restricted *Isovist* area, limited only to a predefined angular range and describing the fraction of the floor plan from which a given point is visible at the predefined angle; Section 11, p. 137.

Visitor Experience - the memories of the art exhibition resulting from the individual interactions between the visitor and the art objects contained in that exhibition; see Section 2, p. 35 but compare with the limitations noted in Section 26.5, p. 338.

Visual Attention - the cognitive process responsible for acquiring and early processing of visual information from the environment; Section 3.1.1, p. 42.

Visual Interaction - see: *engagement*.

BIBLIOGRAPHY

Adamou, C., Avraamides, M. N. & Kelly, J. W. (2014). Integration of visuospatial information encoded from different viewpoints. *Psychonomic bulletin & review*, *21*(3), 659–665.

Akaike, H. (1992). Information theory and an extension of the maximum likelihood principle. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics* (pp. 610–624). New York: Springer.

Akdal, G., Hodgson, T. L., Hill, A. C., Mannan, S. & Kennard, C. (2002). Visual object memory and memory-guided saccades rely on shared mental representations. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, *143*(4), 509–14.

Akpan, I., Marshall, P., Bird, J. & Harrison, D. (2013). Exploring the effects of space and place on engagement with an interactive installation. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 2213–2222). New York: ACM.

Allen, S. [Stan]. (1996). Field Conditions. *Architectural Design*, *66*, 21.

Allen, S. [Sue]. (2004). Designs for learning: Studying science museum exhibits that do more than entertain. *Science Education*, *88*(S1), S17–S33.

Alpers, S. (1991). The Museums as a Way of Seeing. In I. Karp & S. D. Lavine (Eds.), *Exhibiting cultures: the poetics and politics of museum display* (pp. 25–32). Washington: Smithsonian Institution Press.

Andersen, G. (Ed.). (2004). *Reinventing the museum: Historical and contemporary perspectives on the paradigm shift*. Walnut Creek, CA: Rowman Altamira.

Andersson, R., Nyström, M. & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research*, *3*(3), 6.

Arnsdorf, I. (2010). The Museum Is Watching You: Galleries quietly study what people like, or skip, to decide what hangs where. *The Wall Street Journal*, *August*(18).

Ashcraft, M. H. & Radvansky, G. A. (2010). *Cognition* (5th). Upper Saddle River, NJ, USA: Pearson.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York: Cambridge University Press.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390–412.

Baayen, R. H. & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Ballendat, T., Marquardt, N. & Greenberg, S. (2010). Proxemic interaction: designing for a proximity and orientation-aware environment. In *Acm international conference on interactive tabletops and surfaces* (pp. 121–130). New York: ACM.

BALTIC Centre for Contemporary Art, M. (2013). Thomas Scheibitz: ONE—Time Pad. Vimeo.com.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bartoń, K. (2014). MuMIn: Multi-model inference. R package version 1.15.1.

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.

Baxandall, M. (1991). Exhibiting Intention: Some Preconditions of the Visual Display of Culturally Purposeful Objects. In I. Karp & S. D. Lavine (Eds.), *Exhibiting cultures: the poetics and politics of museum display* (pp. 33–41). Washington: Smithsonian Institution Press.

Belke, B., Leder, H., Strobach, T. & Carbon, C. C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 214.

Bellamy, S. (2012). Strike a Pose. http://www.susibellamy.com.

Benedikt, M. (1979). To take hold of space: isovists and isovist fields. *Environment and Planning B: Planning and Design*, *6*, 47–65.

Benjamin, W. (1968). The work of art in the age of mechanical reproduction. In H. Arendt (Ed.), *Illuminations* (pp. 214–218). London: Fontana.

Bitgood, S. (1991). Common Beliefs About Visitors: Do We Really Understand Our Visitors? *Visitor Behavior*, *6*(1), 6.

Bitgood, S. (2003). Environmental psychology in museums, zoos, and other exhibition centers. In R. B. Bechtel & A. Churchman (Eds.), *Handbook of environmental psychology* (pp. 461–480). New York: John Wiley and Sons.

Bitgood, S. (2009). Museum fatigue: A critical review. *Visitor Studies*, *12*(2), 93–111.

Bitgood, S. (2010). *An attention-value model of museum visitors*. Center for the Advancement of Informal Science Education.

Bitgood, S., McKerchar, T. L. & Dukes, S. (2013). Looking Back at Melton: Gallery Density and Visitor Attention. *Visitor Studies*, *16*(2), 217–225.

Bitgood, S. & Patterson, D. D. (1993). The effects of gallery changes on visitor reading and object viewing time. *Environment and Behavior*, *25*(6), 761–781.

Bollo, A. & Dal Pozzolo, L. (2005). Analysis of visitor behaviour inside the museum: An empirical study. In *Proceedings of the 8th international conference on arts and cultural management*. Montreal.

Bourdeau, L. & Chebat, J.-C. (2003). The Effects Of Signage And Location Of Works Of Art On Recall Of Titles And Paintings In Art Galleries. *Environment & Behavior*, *35*(2), 203–226.

Brieber, D., Nadal, M., Leder, H. & Rosenberg, R. (2014). Art in time and space: context modulates the relation between art experience and viewing time. *PloS one*, *9*(6), e99019.

Brignull, H. & Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In M. Rauterberg, M. Menozzi & J. Wesson (Eds.), *Proceedings of interact* (Vol. 3, pp. 17–24). Amsterdam: IFIP.

Brown, L. T. & Farha, W. (1966). Some physical determinants of viewing time under three instructional sets. *Perception & Psychophysics*, *1*(1), 2–4.

Buchner, A. & Jansen-Osmann, P. (2008). Is Route Learning More Than Serial Learning? *Spatial Cognition & Computation*, *8*(4), 289–305.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, *10*(12), 551–557.

Buswell, G. T. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Caduff, D. & Timpf, S. (2008). On the assessment of landmark salience for human navigation. *Cognitive processing*, *9*(4), 249–67.

Camfield, W. A. (1990). Marcel Duchamp's Fountain: Its History and Aesthetics in the Context of 1917. In R. E. Kuenzli & F. M. Naumann (Eds.), *Marcel duchamp: artist of the century*. Cambridge, MA: MIT Press.

Carbon, C.-C. (2013). BiDimRegression: Bidimensional Regression Modeling Using R. *Journal of Statistical Software*, *52*(1), 1–11.

Castelloe, J. M. & O'Brien, R. G. (2001). Power and sample size determination for linear models: A review of power concepts. In *Twenty-sixth annual sas users group international conference* (pp. 240–26). Cary, NC: SAS Institute Inc.

Chan, E., Baumann, O., Bellgrove, M. A. & Mattingley, J. B. (2012). From objects to landmarks : the function of visual location information in spatial navigation. *Frontiers in psychology*, *3*(August), 304.

Chey, A. (2013). Adrien-Kuba Distance. Unpublished Computer Program.

Choi, Y. K. (1999). The morphology of exploration and encounter in museum layouts. *Environment and Planning B: Planning and Design*, *26*(2), 241–250.

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407–9.

Chun, M. M. & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, *17*(2), 177–184.

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, *112*(1), 155–159.

Conroy Dalton, R. (2001). *Spatial navigation in immersive virtual environments* (Unpublished Doctoral Dissertation, University of London).

Conroy Dalton, R. (2003). The Secret Is To Follow Your Nose: Route Path Selection and Angularity. *Environment and Behavior*, *35*(1), 107–131.

Conroy Dalton, R., Hölscher, C. & Turner, A. (2012). Understanding space: the nascent synthesis of cognition and the syntax of spatial morphologies. *Environment and Planning B: Planning and Design*, *39*(1), 7–11.

Conroy Dalton, R., Kuliga, S. F. & Hölscher, C. (2013). POE 2.0: exploring the potential of social media for capturing unsolicited post-occupancy evaluations. *Intelligent Buildings International*, *5*(3), 162–180.

Craik, F. I. M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of experimental Psychology: general*, *104*(3), 268.

Dalton, N. S., Collins, E. & Marshall, P. (2015). Display Blindness? Looking Again at the Visibility of Situated Displays using Eye Tracking. In *33th international conference on human factors in computing systems, chi 2015*. New York: ACM.

Dalton, N. S. & Conroy Dalton, R. (2010). Solutions for visibility-accessibility and signage problems via layered-graphs. *The Journal of Space Syntax*, *1*(1), 164–176.

Dalton, N. S., Conroy Dalton, R., Hölscher, C. & Kuhnmünch, G. (2012). An iPad app for recording movement paths and associated spatial behaviors. In C. Stachniss, K. Schill & D. Uttal (Eds.), *Spatial cognition viii* (pp. 431–450). Berlin Heidelberg: Springer.

Dalton, N. S., Marshall, P. & Conroy Dalton, R. (2013). Extending architectural theories of space syntax to understand the effect of environment on the salience of situated displays. In *2nd acm international symposium on pervasive displays (perdis '13)* (pp. 73–78). Mountain View, CA: ACM.

Dalton, N. S., Marshall, P. & Conroy Dalton, R. (2010). Measuring Environments for Public Displays : A Space Syntax Approach. In *Chi '10 extended abstracts on human factors in computing systems* (pp. 3841–3846). New York: ACM.

Damerau, F. J. (1964). A Technique for Computer Detection and Correction of Spelling Errors. *Commun. ACM*, *7*(3), 171–176.

Davies, R. L. (1957). Deeper Knowledge: Better Design. *Architectural Record*, *121*.

Davis, R. L., Therrien, B. A. & West, B. T. (2008). Cue Conditions and Wayfinding in Older and Younger Women. *Research in Gerontological Nursing*, *1*(4), 252–263.

Derix, C., Gamlesæter, Å. & Carranza, P. M. (2008). 3d Isovists and Spatial Sensations: Two methods and a case study. In S. Haq, C. Hölscher & S. Torgrude (Eds.), *Movement and orientation in built environments: evaluating design rationale and user cognition* (pp. 67–72). Bremen: Universitat Bremen.

Descoteaux, J. (2007). Statistical power: An historical introduction. *Tutorials in Quantitative Methods for Psychology*, *3*(2), 28–34.

Diamond, J. (1999). *Practical evaluation guide: Tools for museums and other informal educational settings*. Walnut Creek, CA: Rowman Altamira.

Eghbal-Azar, K. & Widlok, T. (2013). Potentials and Limitations of Mobile Eye Tracking in Visitor Studies Evidence From Field Research at Two Museum Exhibitions in Germany. *Social Science Computer Review*, *31*(1), 103–118.

Emo, B. (2014). Seeing the Axial Line: Evidence from Wayfinding Experiments. *Behavioral Sciences*, *4*(3), 167–180.

Falk, J. H. (1993). Assessing the impact of exhibit arrangement on visitor behavior and learning. *Curator: The Museum Journal*, *36*(2), 133–146.

Falk, J. H. & Dierking, L. D. (2004). The contextual model of learning. In G. Anderson (Ed.), *Reinventing the museum: historical and contemporary perspectives on the paradigm shift* (pp. 139–142). Walnut Creek, CA: Rowman Altamira.

Fechner, G. T. (1871). *Zur experimentellen Aesthetik [Toward an experimental aesthetic]*. Leipzig: S. Hirzel.

Fisk, W. J. & Rosenfeld, A. H. (1997). Estimates of improved productivity and health from better indoor environments. *Indoor air*, *7*(3), 158–172.

Franchak, J. M., Kretch, K. S., Soska, K. C. & Adolph, K. E. (2011). Head-mounted eye tracking: a new method to describe infant looking. *Child development*, *82*(6), 1738–50.

Franck, K. A. (1984). Exorcising the ghost of physical determinism. *Environment and Behavior*, *16*(4), 411–435.

Franz, G. & Wiener, J. M. (2008). From space syntax to space semantics: a behaviorally and perceptually oriented methodology for the efficient description of the geometry and topology of environments. *Environment and Planning B: Planning and Design*, *35*(4), 574–592.

Friedman, A. & Kohler, B. (2003). Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and Other Two-Dimensional Data Sets. *Psychological Methods*, *8*(4), 468.

Gardony, A., Brunyé, T., Taylor, H. & Wolford, G. (2013). Streamlining Sketch Map Analysis: The Gardony Map Drawing Analyzer. In *Proceedings of the 54th annual meeting of the psychonomic society.* Toronto, Canada.

Gifford, R., Hine, D. W., Muller-Clemm, W. & Shaw, K. T. (2002). Why architects and laypersons judge buildings differently: Cognitive properties and physical bases. *Journal of Architectural and Planning Research*, *19*(2), 131–148.

Göçer, Ö., Hua, Y. & Göçer, K. (2015). Completing the missing link in building design process: Enhancing post-occupancy evaluation method for effective feedback for building performance. *Building and Environment*, *89*, 14–27.

Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of experimental psychology. Human perception and performance*, *13*(2), 256–66.

Gurian, E. H. (1991). Noodling Around with Exhibition Opportunities. In I. Karp & S. D. Lavine (Eds.), *Exhibiting cultures: the poetics and politics of museum display* (pp. 176–190). Washington: Smithsonian Institution Press.

Hänninen, L. & Pastell, M. (2009). CowLog: Open-source software for coding behaviors from digital video. *Behavior Research Methods*, *41*(2), 472–476.

Harrell Jr, F. E. & Dupont, M. C. (2006). Hmisc: Harrell Miscellaneous. R package version 3.16-0.

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T. & Lovelace, K. (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176.

Heidenreich, S. M. & Turano, K. A. (2010). What predicts where one will look when viewing artwork? *Journal of Vision*, *3*(9), 688–688.

Heidenreich, S. M. & Turano, K. A. (2011). Where does one look when viewing artwork in a museum? *Empirical Studies of the Arts*, *29*(1), 51–72.

Henderson, J. M. (2008). Eye Movements and Scene Memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 87–121). Oxford: Oxford University Press.

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

Hillier, B., Major, M., Desyllas, J., Karimi, K., Campos, B. & Stonor, T. (1996). *Tate Gallery, Millbank: a study of the existing layout and new masterplan proposal*. University College London.

Hollands, M., Patla, A. & Vickers, J. (2002). Look where you're going!: gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research*, *143*(2), 221–30.

Hollingworth, A. (2008). Visual Memory for Natural Scenes. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 123–162). Oxford: Oxford University Press.

Hollingworth, A. (2012). Task specificity and the influence of memory on visual search: Comment on Vo and Wolfe (2012). *Journal of experimental psychology Human perception and performance*, *38*(6), 1596–1603.

Hollingworth, A. & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(1), 113.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Huck, S. W. (2009). *Statistical misconceptions*. New York: Routledge.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Isham, E. A. & Geng, J. J. (2013). Looking Time Predicts Choice but Not Aesthetic Value. *PloS one*, *8*(8), e71698.

Ishihara, S. (1917). *Test for Coiour-Blindness*. Tokyo: Hongo Harukicho.

Itti, L. & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506.

Iwanowska, K. & Voyer, D. (2013). Dimensional transformation in tests of spatial and environmental cognition. *Memory & cognition*, *41*(8), 1122–1131.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434–446.

Janzen, G. (2006). Memory for object location and route direction in virtual large-scale space. *Quarterly Journal of Experimental Psychology*, *59*(3), 493–508.

Janzen, G. & van Turennout, M. (2004). Selective neural representation of objects relevant for navigation. *Nature neuroscience*, *7*(6), 673–7.

Johnson, A. (2011). *The Role of Goals and Attention on Memory for Distance in Real and Virtual Spaces* (Unpublished Doctoral Dissertation, Northumbria University).

Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, *5*(9), 944–946.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. *Attention and performance IX*, *9*, 187–203.

Judd, C. M., McClelland, G. H. & Ryan, C. S. (2011). *Data analysis: A model comparison approach*. London: Routledge.

Kahneman, D., Treisman, A. & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Kauppinen, T., Litvinova, E. & Kallenbach, J. (2014). Capturing and Linking Human Sensor Observations with YouSense. In *Proceedings of the iswc*. Seattle: ACM.

Kelly, J. & McNamara, T. P. (2009). Spatial memory. In B. Goldstein (Ed.), *Encyclopedia of perception* (pp. 916–918). Thousand Oaks, CA: SAGE Publications.

Kirasic, K. C. (2000). Age differences in adults' spatial abilities, learning environmental layout, and wayfinding behavior. *Spatial Cognition and Computation*, *2*(2), 117–134.

Kirchberg, V. & Tröndle, M. (2012). Experiencing exhibitions: A review of studies on visitor experiences in museums. *Curator: The Museum Journal*, *55*(4), 435–452.

Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Thousand Oaks, CA: Sage.

Kray, C., Fritze, H., Fechner, T., Schwering, A., Li, R. & Anacta, V. J. (2013). Transitional Spaces: Between Indoor and Outdoor Spaces. In *Spatial information theory* (pp. 14–32). Scarborough, UK: Springer.

Krejtz, K., Duchowski, A. T. & Çöltekin, A. (2014). High-Level Gaze Metrics From Map Viewing - Charting Ambient/Focal Visual Attention. In P. Kiefer, I. Giannopoulos, M. Raubal & A. Krüger (Eds.), *Proceedings of the 2nd international workshop on eye tracking for spatial research co-located with the 8th international conference on geographic information science (giscience 2014)* (pp. 37–41). Vienna, Austria.

Krukar, J. (2014a). Spontaneous Memory of Art Exhibitions: Explanations from Eye-Tracking. In C. Freksa, B. Nebel, M. Hegarty & T. Barkowsky (Eds.), *Spatial cognition 2014: poster presentations* (pp. 60–63). Bremen, Germany.

Krukar, J. (2014b). Walk, Look, Remember: The Influence of the Gallery's Spatial Layout on Human Memory for an Art Exhibition. *Behavioural Sciences*, *4*(3), 181–201.

Krukar, J. & Conroy Dalton, R. (2012). Towards a Unified Model of Building Usability. In *22nd international association people-environment studies (iaps) conference*. Glasgow, UK.

Krukar, J. & Conroy Dalton, R. (2013a). Spatial Predictors of Eye Movement in a Gallery Setting. In P. Kiefer, I. Giannopoulos, M. Raubal & M. Hegarty (Eds.), *Eye tracking for spatial research, proceedings of the 1st international workshop (in conjunction with cosit 2013)* (pp. 14–19). Scarborough, UK.

Krukar, J. & Conroy Dalton, R. (2013b). Walk, Look, Remember: Art Galleries as Spaces Facilitating Memory. In Y. O. Kim, H. T. Park

& K. W. Seo (Eds.), *Proceedings of ninth international space syntax symposium* (074:1–19). Seoul: Sejong University Press.

Kruskal, J. B. (1983). An overview of sequence comparison: Time warps, string edits, and macromolecules. *SIAM review*, *25*(2), 201–237.

Kuznetsova, A., Bruun Brockhoff, P. & Haubo Bojesen Christensen, R. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models. R package version 2.0-11.

Land, M. F. & Tatler, B. W. (2009). *Looking and acting: vision and eye movements in natural behaviour*. Oxford: Oxford University Press.

Larish, J. & Andersen, G. (1995). Active control in interrupted dynamic spatial orientation: The detection of orientation change. *Attention, Perception, & Psychophysics*, *57*(4), 533–545.

Leder, H., Belke, B., Oeberst, A. & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508.

Lefcheck, J. & Casallas, J. S. (2014). R-squared for generalized linear mixed effects models. GitHub Repository.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. In *Soviet physics doklady* (Vol. 10, p. 707).

Levent, N. & Pascual-Leone, A. (2013a). Introduction. In N. Levent & A. Pascual-Leone (Eds.), *The multisensory museum: cross-disciplinary perspectives on touch, sound, smell, memory, and space*. Lanham: Rowman & Littlefield Publishers.

Levent, N. & Pascual-Leone, A. (Eds.). (2013b). *The Multisensory Museum: Cross-Disciplinary Perspectives on Touch, Sound, Smell, Memory, and Space*. Lanham: Rowman & Littlefield Publishers.

Locher, P. (2011). Contemporary experimental aesthetics: State of the art technology. *i-Perception*, *2*(7), 697.

Locher, P., Krupinski, E. A., Mello-Thoms, C. & Nodine, C. F. (2007). Visual interest in pictorial art during an aesthetic experience. *Spatial Vision*, *21*(1-2), 55–77.

Loftus, G. R. (1972). Eye fixations and recognition memory for pictures. *Cognitive Psychology*, *3*(4), 525–551.

Lombard, M., Snyder-Duch, J. & Bracken, C. C. (2004). Practical resources for assessing and reporting intercoder reliability in content analysis research projects.

Lu, Y. & Peponis, J. (2014). Exhibition visitors are sensitive to patterns of display covisibility. *Environment and Planning B: Planning and Design*, *41*(1), 53–68.

Lu, Y. & Zimring, C. (2012). Can intensive care staff see their patients? An improved visibility analysis methodology. *Environment and Behavior*, *44*(6), 861–876.

Markus, T. A. (1993). *Buildings & power: Freedom and control in the origin of modern building types*. London: Routledge.

Marsh, W. E., Chardonnet, J.-R. & Merienne, F. (2014). Virtual Distance Estimation in a CAVE. In *Spatial cognition ix* (pp. 354–369). Springer.

Mather, M. & Nesmith, K. (2008). Arousal-Enhanced Location Memory for Pictures. *Journal of memory and language*, *58*(2), 449–464.

Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior research methods*, *44*(2), 314–24.

Mayr, E., Knipfer, K. & Wessel, D. (2009). In-sights into mobile learning: An exploration of mobile eye tracking methodology for learning in museums. In G. Vavoula, N. Pachler & A. Kukulska-Hulme (Eds.), *Researching mobile learning: frameworks, methods, and research designs* (pp. 189–204). Oxford: Peter Lang.

Mayr, S., Erdfelder, E., Buchner, A. & Faul, F. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology*, *3*(2), 51–59.

McCarley, J. S., Steelman, K. S. & Horrey, W. J. (2014). The View from the Driver's Seat: What Good Is Salience? *Applied Cognitive Psychology*, *28*(1), 47–54.

McDermott, K. B. & Arnold, K. M. (2013). The Testing Effect. In T. J. Perfect & D. S. Lindsay (Eds.), *The sage handbook of applied memory* (p. 183). Thousand Oaks, CA: SAGE.

McLean, K. (1993). *Planning for people in museum exhibitions*. Washington: Association of Science-Technology Centers.

McManus, P. M. (1993). Memories as indicators of the impact of museum visits. *Museum Management and Curatorship*, *12*(4), 367–380.

McNamara, T. P. (1986). Mental representations of spatial relations. *Cognitive psychology*, *18*(1), 87–121.

McNamara, T. P. (2013). Spatial Memory: Properties and Organization. In D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 173–190). Washington: American Psychological Association.

Meilinger, T., Berthoz, A. & Wiener, J. M. (2011). The integration of spatial information across different viewpoints. *Memory & cognition*, *39*(6), 1042–1054.

Meilinger, T., Franz, G. & Bülthoff, H. H. (2012). From isovists via mental representations to behaviour: first steps toward closing the causal chain. *Environment and Planning B: Planning and Design*, *39*(1), 48–62.

Melton, A. W. (1935). *Problems of installation in museums of art.* Washington: American Association of Museums.

Miller, J. & Carlson, L. (2011). Selecting landmarks in novel environments. *Psychonomic Bulletin & Review*, *18*(1), 184–191.

Miller, J., Carlson, L. & Hill, P. L. (2011). Selecting a reference object. *Journal of experimental psychology. Learning, memory, and cognition*, *37*(4), 840–50.

Miyahira, A., Morita, K., Yamaguchi, H., Nonaka, K. & Maeda, H. (2000). Gender differences of exploratory eye movements: a life span study. *Life sciences*, *68*(5), 569–577.

Montello, D. R. (1993). Scale and multiple psychologies of space. In A. U. Frank & I. Campari (Eds.), *Spatial information theory: a theoretical basis for gis* (Lecture No, pp. 312–321). Berlin: Springer-Verlag.

Montello, D. R. (1998). A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In M. Egenhofer (Ed.), *Spatial and temporal reasoning in geographic information systems* (pp. 143–154). New York: Oxford University Press.

Montello, D. R. (2007). The contribution of space syntax to a comprehensive theory of environmental psychology. In *Proceedings of the 6th international space syntax symposium* (iv:01–12). Istanbul.

Montello, D. R. (2014). Spatial Cognition and Architectural Space: Research Perspectives. *Architectural Design*, *84*(5), 74–79.

Müller-Feldmeth, D., Schwarzkopf, S., Büchner, S. J., Hölscher, C., Kallert, G., von Stülpnagel, R. & Konieczny, L. (2014). Location Dependent Fixation Analysis with Sight Vectors. Locomotion as a Challenge in Mobile Eye Tracking. In P. Kiefer, I. Giannopoulos, M. Raubal & A. Krüger (Eds.), *Proceedings of the 2nd international workshop on eye tracking for spatial research co-located with the 8th international conference on geographic information science (giscience 2014)* (pp. 67–71). Vienna, Austria.

Mulligan, N. W. (2013). Memory for Pictures and Actions. In T. J. Perfect & D. S. Lindsay (Eds.), *The sage handbook of applied memory* (p. 20). Thousand Oaks, CA: SAGE.

Münzer, S., Zimmer, H. D., Schwalm, M., Baus, J. & Aslan, I. (2006). Computer-assisted navigation and the acquisition of route and survey knowledge. *Journal of Environmental Psychology*, *26*(4), 300–308.

Nadel, L. (2013). Cognitive Maps. In D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 155–171). Washington: American Psychological Association.

Nakagawa, S. & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs: Prentice-Hall.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* New York: WH Freeman/Times Books/Henry Holt & Co.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.

Newhouse, V. (2005, May). *Art and the Power of Placement*. New York: The Monacelli Press.

Nielsen, J. (2013). Mobile Usability Futures. Talks at Google.

Norman, D. A. (1986). Cognitive engineering. Hillsdale, NJ: Lawrence Erlbaum.

Oakes, L. M. (2010). Infancy Guidelines for Publishing Eye-Tracking Data. *Infancy*, *15*(1), 1–5.

O'Doherty, B. (1986). *Inside the White Cube: The Ideology of the Gallery Space*. San Francisco: The Lapis Press.

O'Neill, M. J. (1992). Effects of familiarity and plan complexity on wayfinding in simulated buildings. *Journal of Environmental Psychology*, *12*(4), 319–327.

Packer, J. & Bond, N. (2010). Museums as restorative environments. *Curator: The Museum Journal*, *53*(4), 421–436.

Pallasmaa, J. (2013). Museum as an Embodied Experience. In N. Levent & A. Pascual-Leone (Eds.), *The multisensory museum: cross-disciplinary perspectives on touch, sound, smell, memory, and space* (pp. 239–249). Lanham: Rowman & Littlefield Publishers.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K. & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2*(2), 4.

Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Penn, A. (2003). Space Syntax And Spatial Cognition Or Why the Axial Line? *Environment and Behavior*, *35*(1), 30–65.

Peponis, J. & Hedin, J. (1982). The layout of theories in the Natural History Museum. *9H*, *3*, 21–25.

Peponis, J., Wineman, J. D., Rashid, M., Kim, S. & Bafna, S. (1997). On the Description of Shape and Spatial Configuration Inside Buildings. In M. D. Major (Ed.), *Space syntax. first international symposium.* (40:1–21). London: UCL.

Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of consumer research*, *28*(3), 450–461.

Pevsner, N. (1976). *A History of Building Types*. London: Thames and Hudson.

Philbeck, J. W. & Sargent, J. (2013). Perception of Spatial Relations During Self-Motion. In D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 99–115). Washington: American Psychological Association.

Pihko, E., Virtanen, A., Saarinen, V.-M., Pannasch, S., Hirvenkari, L., Tossavainen, T., . . . Hari, R. (2011). Experiencing art: the influence of expertise and painting abstraction level. *Frontiers in human neuroscience*, *5*.

Posner, M. I., Snyder, C. R. & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of experimental psychology: General*, *109*(2), 160.

Psarra, S. (2006). Spatial culture, way-finding and the educational message. In S. MacLeod (Ed.), *Reshaping museum space* (pp. 78–94). Oxon: Routledge.

Psarra, S. (2009). *Architecture and narrative: the formation of space and cultural meaning*. London: Routledge.

Qualtrics. (2013). Qualtrics. Provo, Utah, USA: Qualtrics.

R Core Team. (2013). R: A Language and Environment for Statistical Computing. Vienna, Austria.

Radvansky, G. A. & Copeland, D. E. (2006). Walking through doorways causes forgetting: Situation models and experienced space. *Memory & Cognition*, *34*(5), 1150–1156.

Radvansky, G. A., Krawietz, S. A. & Tamplin, A. K. (2011). Walking through doorways causes forgetting: Further explorations. *The Quarterly Journal of Experimental Psychology*, *64*(8), 1632–1645.

Radvansky, G. A., Tamplin, A. K. & Krawietz, S. A. (2010). Walking through doorways causes forgetting: Environmental integration. *Psychonomic bulletin & review*, *17*(6), 900–904.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological bulletin*, *114*(3), 510–32.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Reinitz, M. T. & Hannigan, S. L. (2001). Effects of simultaneous stimulus presentation and attention switching on memory conjunction errors. *Journal of Memory and Language*, *44*(2), 206–219.

Roberts, C. & Sibbald, B. (1998). Randomising groups of patients. *British Medical Journal*, *316*(7148), 1898–1900.

Roberts, L. C. (2004). Changing practices of interpretation. In G. Anderson (Ed.), *Reinventing the museum: historical and contemporary perspectives on the paradigm shift* (pp. 212–232). Walnut Creek, CA: Rowman Altamira.

Robinson, E. (1928). *The behaviour of the museum visitor*. The American Association of Museums. Washington.

Rudofsky, B. (1987). *Architecture Without Architects: A Short Introduction to Non-Pedigreed Architecture*. New Mexico: University of New Mexico Press.

Santner, K., Fritz, G., Paletta, L. & Mayer, H. (2013). Visual recovery of saliency maps from human attention in 3D environments. In *2013 ieee international conference on robotics and automation (icra)* (pp. 4297–4303). Karlsruhe, Germany: IEEE.

Sauppé, A. & Mutlu, B. (2014). How Social Cues Shape Task Coordination and Communication. In *Proceedings of the 17th acm conference* *on computer supported cooperative work & social computing*. New York: ACM.

Schneider, S. & König, R. (2012). Exploring the Generative Potential of Isovist Fields. In *30th international conference on education and research in computer aided architectural design in europe*. Prague: eCAADe.

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition*, *80*(1-2), 1–46.

Schultz, C. & Bhatt, M. (2013). InSpace3D: A middleware for built environment data access and analytics. *Procedia Computer Science*, *18*, 80–89.

Screven, C. G. (1976). Exhibit Evaluation-A Goal-Referenced Approach. *Curator: The Museum Journal*, *19*(4), 271–290.

Serrell, B. (1997). Paying attention: The duration and allocation of visitors' time in museum exhibitions. *Curator*, *40*(2), 108–125.

Serrell, B. (2011). *Paying More Attention to "Paying Attention"*. Center for advancement of informal science education.

Shinoda, H., Hayhoe, M. M. & Shrivastava, A. (2001). What controls attention in natural environments? *Vision Research*, *41*(25–26), 3535–3545.

Siegel, A. W. & White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in child development and behavior*, *10*, 9–55.

Sixsmith, M. (1999). *Designing galleries: The complete guide to developing and designing spaces and services for temporary exhibitions*. Arts Council of England.

Smith, C. S. (1989). Museums, Artefacts, and Meanings. In P. Vergo (Ed.), *The new museology* (pp. 6–21). London: Reaktion Books.

Smith, J. K. & Smith, L. F. (2001). Spending time on art. *Empirical Studies of the Arts*, *19*(2), 229–236.

Spiers, H. J. & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of environmental psychology*, *28*(3), 232–249.

Stavroulaki, G. & Peponis, J. (2003). The spatial construction of seeing at Castelvecchio. In *Proceedings of the 4th international space syntax* (66:1–13). London: UCL.

Steck, S. & Mallot, H. (2000). The role of global and local landmarks in virtual environment navigation. *Presence: Teleoperators & Virtual Environments*, *9*(1), 69–83.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, *9*(4), 66–78.

Stephen, G. A. (1992). *String Search*. University College of North Wales. Bangor, UK.

Tabachnick, B. G., Fidell, L. S. et al. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

Talmi, D., Schimmack, U., Paterson, T. & Moscovitch, M. (2007). The role of attention and relatedness in emotionally enhanced memory. *Emotion*, *7*(1), 89.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review*, *2*(1), 55–82.

Tatler, B. W. (2008). On the perception of natural scenes: an introduction to the special issue. *Journal of Eye Movement Research*, *2*(2), 1: 1–4.

Tatler, B. W., Baddeley, R. J. & Gilchrist, I. D. (2005). Visual correlates of fixation selection: Effects of scale and time. *Vision research*, *45*(5), 643–659.

Tatler, B. W., Gilchrist, I. D. & Land, M. F. (2005). Visual memory for objects in natural scenes: from fixations to object files. *The Quarterly Journal of Experimental Psychology*, *58A*(5), 931–60.

Tatler, B. W., Gilchrist, I. D. & Rusted, J. (2003). The time course of abstract visual representation. *Perception*, *32*(5), 579–92.

Tatler, B. W., Hayhoe, M. M., Land, M. F. & Ballard, D. H. (2011). Eye guidance in natural vision: reinterpreting salience. *Journal of vision*, *11*(5), 5:1–23.

Tatler, B. W. & Land, M. F. (2011). Vision and the representation of the surroundings in spatial memory. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1564), 596–610.

Tatler, B. W. & Tatler, S. L. (2013). The influence of instructions on object memory in a real-world setting. *Journal of vision*, *13*(2), 5.

Taylor, B. L. (2010). Reconsidering Digital Surrogates: Toward a viewer-oriented model of the gallery experience. In S. H. Dudley (Ed.), *Museum materialities: objects, engagements, interpretations* (pp. 175–184). London: Routledge.

Taylor, H., Naylor, S. J. & Chechile, N. a. (1999). Goal-specific influences on the representation of spatial perspective. *Memory & cognition*, *27*(2), 309–19.

Thompson, D. (1990). An Architectural View of the Visitor-Museum Experience. In S. Bitgood, A. Benefield & D. Patterson (Eds.), *Visitor studies: theory, research, and practice* (pp. 72–85). Jacksonville: Center for Social Design.

Tobler, W. R. (1965). Computation of the correspondence of geographical patterns. In *Papers of the regional science association* (Vol. 15, *1*, pp. 131–139). Springer.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.

Torralba, A., Oliva, A., Castelhano, M. S. & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological review*, *113*(4), 766.

Tröndle, M., Greenwood, S., Bitterli, K. & van den Berg, K. (2014). The effects of curatorial arrangements. *Museum Management and Curatorship*, *29*(2), 140–173.

Tröndle, M., Greenwood, S., Kirchberg, V. & Tschacher, W. (2014). An Integrative and Comprehensive Methodology for Studying Aesthetic Experience in the Field Merging Movement Tracking, Physiology, and Psychological Data. *Environment and Behavior*, *46*(1), 102–135.

Tröndle, M. & Tschacher, W. (2012). The physiology of phenomenology: The effects of artworks. *Empirical Studies of the Arts*, *30*(1), 75–113.

Tröndle, M., Wintzerith, S., Wäspe, R. & Tschacher, W. (2012). A museum for the twenty-first century: the influence of 'sociality'on art reception in museum space. *Museum Management and Curatorship*, *27*(5), 461–486.

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K. & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96–103.

Tufte, E. (1974). Interpretation of Regression Coeficients when the Variables are Re-expressed as Logarithms (with Five Examples). In *Data analysis for politics and policy* (pp. 108–134). New Jersey: Prentice-Hall Englewood Cliffs.

Turner, A. (2001). A program to perform visibility graph analysis. In J. Peponis, J. Wineman & S. Bafna (Eds.), *Proceedings of the 3rd space syntax symposium* (pp. 31.1–9). Atlanta: University of Michigan.

Turner, A., Doxa, M., O'Sullivan, D. & Penn, A. (2001). From isovists to visibility graphs: a methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design*, *28*(1), 103–121.

Turner, A. & Penn, A. (1999). Making isovists syntactic: isovist integration analysis. In *Proceedings of the 2nd international symposium on space syntax* (pp. 103–121). Brasilia.

Tzortzi, K. (2003). An approach of the microstructure of the gallery space: The case of the Sainsbury Wing. In *Proceedings of the 4th international space syntax symposium* (pp. 67.1–16). London.

Tzortzi, K. (2007). Museum building design and exhibition layout. In *Proceedings of the 6th international space syntax symposium* (072:01–16). Istanbul.

van der Loo, M. P. J. (2014). The stringdist Package for Approximate String Matching. R package version 0.9-3. R.

Velichkovsky, B. M., Joos, M., Helmert, J. R. & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In *Proceedings of the xxvii* *conference of the cognitive science society* (pp. 2283–2288). Mahwah, NJ: Lawrence Erlbaum Associates.

Vergo, P. (Ed.). (1989a). *The New Museology*. London: Reaktion Books.

Vergo, P. (1989b). The Reticent Object. In P. Vergo (Ed.), *The new museology* (pp. 41–59). London: Reaktion Books.

Vishwanath, D., Girshick, A. R. & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature neuroscience*, *8*(10), 1401–10.

von Stülpnagel, R. & Steffens, M. C. (2012). Can Active Navigation Be as Good as Driving? A Comparison of Spatial Memory in Drivers and Backseat Drivers. *Journal of Experimental Psychology-Applied*, *18*(2), 162.

Wagner, R. A. (1975). On the complexity of the extended string-to-string correction problem. In *Proceedings of seventh annual acm symposium on theory of computing* (pp. 218–223). New York: ACM.

Wai, J., Lubinski, D. & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, *101*(4), 817.

Waller, D., Loomis, J. M. & Steck, S. D. (2003). Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin & Review*, *10*(4), 987–993.

Waller, D. & Nadel, L. (2013). Introduction: Frameworks for Understanding Spatial Thought (or Wrapping Our Heads Around Space). In D. Waller & L. Nadel (Eds.), *Handbook of spatial cognition* (pp. 3–12). Washington: American Psychological Association.

Ward, J. (2013). Multisensory Memories: How Richer Experiences Facilitate Remembering. In N. Levent & A. Pascual-Leone (Eds.), *The multisensory museum: cross-disciplinary perspectives on touch, sound, smell, memory, and space* (pp. 273–284). Lanham: Rowman & Littlefield Publishers.

Weil, S. E. (2004). Rethinking the museum: An emerging new paradigm. In G. Anderson (Ed.), *Reinventing the museum: historical and contemporary perspectives on the paradigm shift* (pp. 74–79). Walnut Creek, CA: Rowman Altamira.

Wessel, D., Mayr, E. & Knipfer, K. (2007). Re-viewing the museum visitor's view. In *Workshop research methods in informal and mobile learning* (pp. 17–23). London: Institute of Education.

Westat, F. J. (2010). *The 2010 User-Friendly Handbook for Project Evaluation*. Directorate for Education.

Whelan, R. (2008). Effective analysis of reaction time data. *The Psychological Record*, *58*(3), 475–482.

Wickham, H. (2009). *Ggplot2: elegant graphics for data analysis*. New York: Springer.

Wiener, J. M., Franz, G., Rossmanith, N., Reichelt, A., Mallot, H. A. & Bülthoff, H. H. (2007). Isovist analysis captures properties of space relevant for locomotion and experience. *Perception*, *36*(7), 1066–1083.

Wiener, J. M., Hölscher, C., Büchner, S. & Konieczny, L. (2012). Gaze behaviour during space perception and spatial decision making. *Psychological research*, *76*(6), 713–729.

Wiener, J. M., Hölscher, C., Buechner, S. J. & Konieczny, L. (2009). How the Geometry of Space controls Visual Attention during Spatial Decision Making. In *Annual meeting of the cognitive science society (cogsci)* (pp. 2286–2291).

Wineman, J. D. & Peponis, J. (2010). Constructing Spatial Meaning Spatial Affordances in Museum Design. *Environment and Behavior*, *42*(1), 86–109.

Wineman, J. D., Peponis, J. & Conroy Dalton, R. (2006). Exploring, Engaging, Understanding in Museums. In C. Hoelscher, R. Conroy Dalton & P. Turner (Eds.), *Space syntax and spatial cognition workshop: spatial cognition '06* (pp. 33–51). Bremen: Universität Bremen.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv preprint arXiv:1308.5499*.

Witmer, L. (1893). *Zur experimentellen Aesthetik einfacher raeumlicher Formverhaeltnisse*. Leipzig: W. Engelmann.

Wojciechowski, R., Walczak, K., White, M. & Cellary, W. (2004). Building virtual and augmented reality museum exhibitions. In *Proceedings of the ninth international conference on 3d web technology* (pp. 135–144). New York: ACM.

Wright, P. (1989). The Quality of Visitors' Experiences in Art Museums. In P. Vergo (Ed.), *The new museology* (pp. 119–148). London: Reaktion Books.

Wright, R. D. & Ward, L. M. (2008). *Orienting of attention.* Oxford: Oxford University Press.

Wundt, W. M. (1874). *Grundzuege der physiologischen Psychologie*. W. Engelman.

Xie, H., Filippidis, L., Gwynne, S., Galea, E. R., Blackshields, D. & Lawrence, P. J. (2007). Signage Legibility Distances as a Function of Observation Angle. *Journal of Fire Protection Engineering*, *17*(1), 41–64.

Yalowitz, S. S. & Bronnenkant, K. (2009). Timing and tracking: Unlocking visitor behavior. *Visitor Studies*, *12*(1), 47–64.

Yang, T. & Kubovy, M. (1999). Weakening the robustness of perspective: evidence for a modified theory of compensation in picture perception. *Perception & psychophysics*, *61*(3), 456–67.

Yoshimura, Y., Girardin, F., Carrascal, J. P., Ratti, C. & Blat, J. (2012). New tools for studying visitor behaviours in museums: a case study at the Louvre. In *Information and communication technologies in tourism 2012* (pp. 391–402). Berlin Heidelberg: Springer.

Zamani, P. (2009). Architecture as Curatorial Device. In D. Koch, L. Marcus & J. Steen (Eds.), *Proceedings of the 7th international space syntax symposium* (134:1–12). Stockholm: KTH.

Zisch, F., Gage, S. & Spiers, H. (2013). Navigating the Museum. In N. Levent & A. Pascual-Leone (Eds.), *The multisensory museum: cross-disciplinary perspectives on touch, sound, smell, memory, and space* (pp. 215–237). Lanham: Rowman & Littlefield Publishers.