Perceptual Dynamics of Daylight in Architecture

THÈSE N^{O} 7677 (2017)

PRÉSENTÉE LE 12 JUIN 2017

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE INTERDISCIPLINAIRE DE PERFORMANCE INTÉGRÉE AU PROJET
PROGRAMME DOCTORAL EN ARCHITECTURE ET SCIENCES DE LA VILLE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Siobhan Francois ROCKCASTLE

acceptée sur proposition du jury:

Prof. D. Dietz, président du jury
Prof. M. Andersen, directrice de thèse
Prof. K. Steemers, rapporteur
Prof. M. Inanici, rapporteuse
Prof. M. Pauly, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2017

ABSTRACT

In our experience of daylit architecture, our visual perception is greatly impacted by the ephemeral and inherently dynamic conditions of the surrounding environment. Driven by changes in sky type, time-of-day, and time-of-year, these variable conditions can alter our impressions and appraisal of indoor space. Daylight, including both direct sunlight and indirect skylight, drives a powerful range of perceptual phenomena, which transform structured geometry into a time lapse of slow and smooth or fast and harsh effects. Between the disciplines of architectural design and building engineering, there are many approaches to evaluating daylight performance, from qualitative considerations such as texture, color, composition, and ambiance to compliancy targets regarding task performance, energy, and visual discomfort. While some of these considerations, like visual discomfort, are both quantitative in prediction method and qualitative in subjective evaluation, there are few quantitative measurements developed to predict the positive perceptual impacts of light on human perception. Unlike other environmental factors like temperature, air quality, and sound, daylight creates direct impacts on the appearance of a space and cannot be divorced from the simultaneously aesthetic implications it has on architectural design. Research in daylight is therefore necessarily interdisciplinary, as it integrates physics, with the psychology of emotion, the bio-mechanics of perception, and the aesthetics of architectural design.

To integrate perceptual, aesthetic, and emotional considerations into lighting performance evaluation, the aim of this thesis is to determine whether objective, quantifiable characteristics of luminous composition within an architectural scene can be linked to subjective evaluations of visual interest (like contrast, excitement, pleasantness, etc.) and whether these characteristics are sensitive to temporal dynamics. This thesis will begin with a review of existing quantitative measures for predicting contrast perception in daylit scenes (both real and digital) and will present a comparison of these measures using a catalogue of rendered scenes.

Through a pair of experiments designed to induce visual effects and record subjective responses – an online survey using 2D renderings and an immersive 3D study done in Virtual Reality – this thesis will introduce a method for predicting those perceptual responses using image-based algorithms and a proportional odds model. Using the algorithms and model developed from experimental data, a visually immersive, simulation-based approach will be adopted to evaluate attributes of visual interest through space and over time. By selecting a series of architectural spaces to exemplify this approach, dynamic predictions of daylight-induced excitement across an array of eye-level view positions will show the highly variable nature of perceptual performance

and its capacity to impact occupant appraisals of space. The novelty of the proposed measures, prediction model, and simulation-based approach opens an exciting new frontier in daylight performance evaluation, demonstrating the importance of occupant perception alongside existing task, energy, and comfort considerations.

Keywords: Daylight performance, daylight simulation, perception, contrast, human-centric lighting, visual interest, architecture, experimental, virtual reality.

RÉSUMÉ

Lors de l'interaction entre éclairage naturel et architecture, notre perception visuelle est grandement affectée par les conditions éphémères et intrinsèquement dynamiques du milieu environnant. Dictées par les changements de type de ciel, de l'heure du jour et de l'époque de l'année, ces conditions variables peuvent altérer nos impressions et notre appréciation d'un espace intérieur. La lumière du jour, composée de la lumière directe du soleil et indirecte du ciel, engendre un puissant spectre de phénomènes perceptuels qui transforment une géométrie structurée en une transition temporelle d'effets lents et fluides ou rapides et contrasté. Entre les domaines de l'ingénierie du bâtiment et de la conception architecturale, de nombreuses approches existent pour évaluer la performance en éclairage naturel, allant de considérations qualitatives comme la texture, la couleur, la profondeur de champ et l'ambiance, à des objectifs normatives liées au rendement, l'efficience énergétique ou encore l'inconfort visuel. Alors que certaines de ces considérations, telles que l'inconfort visuel, sont à la fois de nature quantitative dans une méthode de prédiction et qualitative dans une évaluation subjective, il n'existe que peu de mesures quantitatives développées pour prédire l'impact positif de la lumière sur la perception humaine. Contrairement à d'autres facteurs environnementaux tels que la température, la qualité de l'air et le son, la lumière du jour a un impact direct sur l'apparence d'un espace et ne peut être séparée des implications esthétiques qu'elle a sur la conception architecturale. C'est pourquoi l'étude de la lumière naturelle doit nécessairement être interdisciplinaire car elle intègre la physique, la psychologie de l'émotion, la biomécanique de la perception, ainsi que l'esthétique de la conception architecturale.

Afin d'intégrer les considération perceptuelles, esthétiques et émotionnelles au sein de l'évaluation de la performance en lumière, cette thèse a pour but de déterminer si des caractéristiques objectives et quantifiables de la composition du spectre lumineux d'une scène architecturale peuvent être liées à une évaluation subjective des intérêts visuels (comme le contraste, l'enthousiasme, l'attractivité, etc.), et si ces caractéristiques sont sensibles à une dynamique temporelle. Cette thèse débutera par une revue des mesures quantitatives existantes pour prédire la perception de contraste dans les scènes éclairées naturellement (réelles et digitales), et présentera une comparaison de ces mesures en utilisant un catalogue de rendus photo-réalistes.

A travers deux expériences conçues pour induire des effets visuels et enregistrer les réponses subjectives - un sondage en ligne utilisant des rendus 2D et une étude immersive 3D effectuée en réalité virtuelle -, cette thèse va introduire une méthode de prédiction des réponses perceptuelles en utilisant des algorithmes basés sur des images et un modèle de probabilité proportionnelle. À

l'aide des algorithmes et du modèle développés à partir des données expérimentales, une approche visuellement immersive, basée sur des simulations, sera adoptée pour évaluer et visualiser les prédictions de l'intérêt visuel et de l'enthousiasme à travers l'espace et le temps. En sélectionnant une série d'espaces architecturaux pour illustrer cette approche, les prédictions dynamiques de l'enthousiasme stimulé par la lumière naturelle à divers points de vue situés à la hauteur de l'œil va démontrer la nature extrêmement variable de la performance perceptive et sa capacité à affecter l'évaluation de l'espace faite par l'occupant. L'aspect innovateur des mesures proposées, du modèle prédictif et de l'approche basée sur la simulation, définit une nouvelle et passionnante frontière dans l'évaluation de la performance en éclairage naturel, démontrant l'importance de la perception de l'occupant en parallèle aux considérations de rendement, de confort et d'énergie.

Mots Clés: performance de la lumière du jour, simulation de la lumière du jour, perception, contraste, l'eclairage avec une approache centrée sur l'homme, intérêt visuel, architecture, expérimentation, réalité virtuelle.

ACKNOWLEDGEMENTS

"Keep your face always toward the sunshine - and shadows will fall behind you." Walt Whitman

To my adviser Marilyne, thank you for leading me out of the dark infinite corridors of MIT and giving me the opportunity to pursue this work under the bright sun of les Alpes vaudoises. Collaborating with you these past 6 years has been one of the greatest adventures of my life. Your guidance, criticism, rigor, vision, example, and support have been profoundly influential and I look forward to ongoing collaboration in the years to come.

To EPFL and Velux Stiftung, for providing the financial support for this research.

To my colleagues at LIPID - Maria, Mandana, Parag, Emilie, Kynthia, Giuseppe, Minu, Giorgia, Luisa, Jan, Lorenzo, and Peter - who have helped to construct a truly special environment for interdisciplinary thinking and collaboration. May you stay open-minded, mindful, patient, and respectful of those who bring a different background and may you dare to ask the hard questions.

To my collaborators, Maria and Kynthia, you helped me bring this thesis to a new level of inquiry and I am forever grateful for your creativity, expertise, and enthusiasm. You've made these years particularly enriching and through your contributions and constructive criticism, my research has grown in leaps and bounds.

To my family, whose encouragement and support have fueled my ability and motivation to pursue this academic feat. To my father and mother, who raised me to fight for what I believe in and never back down from a challenge that peaks my interest. To my sister for setting an incredibly high bar and challengeing me to jump. I am forever grateful for your patience, creativity, and good example.

To my friends, near and far, who kept me smiling, laughing, loving, skiing, hiking, traveling, and partying...you have been the rock in my castle! I would like to put special emphasis on my main pillars of support outside of LIPID - Alex, Nettra, Allie, Cayley, Isa, Claudia, Galyna, Catherine, Dimi, and JuJu - without whom, I might have lost the better part of my mind along with most of my dignity.

And finally, to Romain, who emerged out of the wild in the final hours of my time in Switzerland - thank you for tolerating me through the completion of this dissertation and making me believe, once again, in the value of partnership.

CONTENTS

Abstract i
Résumé iii
Acknowledgements v
Contents vii

1 Introduction **1**
1.1 Daylight in Architecture 2
1.1.1 The Evolving Role of Daylight Performance 7
1.1.2 The Perceptual Quality of Daylight 10
1.2 Thesis Structure 15
1.2.1 Problem Statement 15
1.2.2 Thesis Objectives 16
1.2.3 Phases of Research 16

2 State of the Art **21**
2.1 Existing Task-Driven Daylight Performance Metrics 21
2.1.1 Illuminance-Based 21
2.1.2 Luminance-Based 23
2.2 Experimental Approaches to Study Perceptual Appraisal 24
2.2.1 Real Spaces 26
2.2.2 Physical Models 29
2.2.3 Digital Images 30
2.3 Strategies for Assessing Display Devices for Experimental Use 32
2.3.1 2D Displays 34
2.3.2 3D and/or Stereoscopic Displays 36
2.3.3 Tone-Mapped Images as a Surrogate for Real Space 38
2.4 Measuring Contrast in Digital Images 40
2.4.1 Global Measures 40
2.4.2 Local Measures 41
2.5 Summary of the Current State 43

3 Evaluating Dynamic Contrast in Daylit Renderings: A Proof-of-Concept **47**
3.1 Evaluating Contrast in Digital Images 47
3.2 Spatial Contrast 48
3.2.1 An Annual Approach to Spatial Contrast 52
3.3 Application of Spatial Contrast 54
3.3.1 Selection & Modelling of Case Studies 55
3.4 Results for Three Case Study Models 56
3.4.1 Case Study One: Direct Top-Lit Space 57
3.4.2 Case Study Four: Side-Lit Space 57
3.4.3 Case Study Nine: Indirect Top-Lit Space 57
3.5 Overview of the Results 58

3.6 Discussion....63
3.7 Chapter Summary....65

4 Comparison of Contrast Algorithms to Intuitive Rankings....**69**
4.1 Selected Algorithms....69
4.2 Comparison of Algorithms Using Case Study Renderings....74
4.3 Comparison to Intuitive Ranking....74
4.4 Chapter Summary....77

5 A Web-Based Experiment on Visual Interest Using 2D Renderings....**81**
5.1 Selection & Modelling of Architectural Case Studies....81
5.2 Design of Experiment....82
5.3 Rendering & Tone Mapping....84
5.4 Semantic Differential Scales....88
5.5 Pilot Study....88
5.6 Online Survey....89
5.6.1 Data Management....90
5.6.2 Data Analysis....91
5.7 Results of 2D User-Based Study....92
5.7.1 Analysis of Subjective Data....92
5.7.2 Effects of Experiment....94
5.7.3 Subject Ratings & 2D Contrast Measures....96
5.7.4 Excitement Prediction Using a Logit Model....98
5.8 Outlook....104
5.9 Chapter Summary....104

6 A Human-Centered Approach to Evaluating Visual Interest....**107**
6.1 A Visually-Immersive Approach....107
6.2 Application of mSC_5 to Visually Immersive Renderings....108
6.3 Simulation Workflow....110
6.3.1 Simulation Setup and Case Study Buildings....111
6.4 Application of mSC_5 to Hemisphere Views....114
6.5 Results....115
6.5.1 Case Study 1: Zollverein School of Management....117
6.5.2 Case Study 2: Test Room....119
6.6 Discussion....119
6.7 Chapter Summary....121

7 A Spatially-Immersive Simulation-Based Workflow....**123**
7.1 Development of an Integrated Platform....124
7.1.1 Simulation Workflow....124
7.1.2 Performance Modules....126
7.1.3 Visualization of Results....127
7.2 Selection of Architectural Case Study....127
7.2.1 Selection of Spaces for Comparative Study....129
7.2.2 Translation of Geometry File into RAD Format....130
7.2.3 Selection of Moments & Views....130
7.2.4 Interpretation of Performance Modules....130
7.3 Results....133
7.3.1 Visual Interest mSC_5....135
7.3.2 Health Potential nvR_D....135

7.4 The Impact of Architecture....137
7.5 Chapter Summary....138

8 A Visually-Immersive Experiment on Visual Interest Using Virtual Reality....**141**
8.1 An Immersive Display Approach....142
8.2 Simulation Workflow....142
8.2.1 Selection of Case Studies....143
8.2.2 360° HDR Renderings....143
8.2.3 Selection of Hemispherical View Directions....148
8.2.4 Projection of Final Scenes....148
8.3 Experimental Design & Procedure....150
8.3.1 Design of Experiment....150
8.3.2 Subjects & Experiment Procedure....151
8.4 Results for the 180° Group....155
8.4.1 180° Group: Distribution of Subject Ratings....155
8.4.2 180° Group: Effects of Space and Sky/View Direction....159
8.4.3 180° Group: Regression Analysis....162
8.5 Comparative Analysis of Results: 180° Group vs. 360° Group....163
8.6 Results for the 360° Group....166
8.6.1 360° Group: Distribution of Subject Ratings....166
8.6.2 360° Group: Effects of Space and Sky....167
8.6.3 360° Group: Regression Analysis....170
8.7 Head Tracking....172
8.8 Chapter Summary....173

9 Conclusion & Outlook....**179**
9.1 Findings & Contributions in Each Research Phase....179
9.1.1 Phase 0: Proof-of-Concept Approach....179
9.1.2 Phase 1: Comparative Study Using Contrast Algorithms....180
9.1.3 Phase 2: Web-Based Experiment Using 2D Renderings....180
9.1.4 Phase 3: Application of Visual Interest to a Simulation Workflow....182
9.1.5 Phase 4: Visually Immersive Experiment in Virtual Reality....184
9.2 Discussion....186
9.2.1 Comparing Online vs. Immersive Experimental Studies....186
9.2.2 Display of HDR Images & Tone-Mapping Considerations....186
9.2.3 Defining 'Visual Interest'....188
9.3 Next Steps....189

Appendix 1....**193**
Appendix 2....**195**
Appendix 3....**205**
Appendix 4....**207**
Appendix 5....**243**

Image Credits....**246**
References....**249**
Glossary of Acronyms & Terms....**259**
CV....**261**

.

1 INTRODUCTION

In the experience of architecture, our visual perception is greatly impacted by the ephemeral and dynamic conditions of the surrounding environment. Driven by changes in the sky, time of day and time of year, these variable conditions can alter our impressions and appraisal of the indoor environment. Daylight, including both direct sunlight and indirect skylight, drives a powerful range of perceptual phenomena, which transform structured geometry into a time lapse of slow and smooth or fast and harsh ethereal effects. Steven Holl (2011) writes about the impact of light on architectural design, stating that "it is not surprising that some architects have written that the entire intention of their work revolves around light…" (p. 63). He goes on to state that the "perceptual spirit and metaphysical strength of architecture are driven by the qualities of light and shadow shaped by solids and voids, by opacities, transparencies, and translucencies." Our sensual interpretation of architecture, according to Holl, is primarily formed according to the conditions of light and shadow.

Between the disciplines of architectural design and building performance engineering, there are many approaches to evaluating daylight performance, from purely qualitative (visual texture, color, spatial depth, and ambiance) to quantitative (compliance targets for illumination, energy, and comfort). Of course, some of these considerations, like comfort, are both quantitative in prediction method and qualitative in subjective assessment making validation and implementation into mainstream standards time-consuming. Unlike other environmental concerns such as temperature, air quality, embodied energy, and acoustics, daylight has direct impacts on our visual environment and cannot be divorced from the simultaneously aesthetic implications it has on architectural design. Over the past several decades, the divide between qualitative and quantitative approaches to daylight performance has begun to shrink, but if and when such a divide exists, it is deeply rooted in disciplinary culture. Through differences in value, research methodology, and means of communication between architecture and engineering, topics that bridge building performance with human well-being and design aesthetics are wracked with methodological challenges.

Research in daylight is necessarily interdisciplinary, as it must integrate the physics of light, with the psychology of emotion, the biology of visual perception, and the aesthetics of architectural design. This thesis began with a healthy skepticism towards existing quantitative criteria (commonly used in building performance engineering) and sought to develop evaluation methods that could integrate aesthetic and experiential factors. To integrate perceptual, aesthetic, and emotional considerations in lighting performance, this thesis explores the impacts of dynamic

environmental conditions on our visual perception of daylit space. Using both simulation-based and experimental methods to induce visual stimuli and measure perceived effects, this thesis introduces a method to predict attributes of visual interest such as excitement, stimulation, pleasantness, and contrast using digital images.

1.1 Daylight in Architecture

"Light for me is the fundamental basis of architecture. I compose with light." (Corbusier, 1991, p. 132)

Visual characteristics of light, such as brightness and contrast, are essential to our interpretation of space, which depends on the balance between light and dark, the eye's ability to perceive those differences, and the brain's ability to translate that information into an understanding of depth and complexity (Liljefors, 1997). Architects and interior designers have long understood the impacts of light on spatial perception and used daylight to sculpt, infuse, and define interior space. Millet (1996) wrote that "As light reveals the forms of architecture and the places made by it, it simultaneously reveals the meaning and the intentions that are released through the process of conceiving, designing, and building" (p. 2). In an essay on Steven Holl's work, theoretician Stanford Kwinter writes, "For Holl, architecture is the science of experience…Light is not itself the plenum of matter, but rather what reveals and conveys it (like water in the paper into which pigment is placed)" (Holl, 2011, p. 78). Light not only reveals spatial depth and material texture, but the ephemerality of these effects under natural lighting conditions can evade intuition and produce unanticipated and even surprising results, both positive and negative. In his own description of the Kurt Hamsun Center, Steven Holl describes the seasonal variations in sunlight as a primary driver in the design process. He describes the design concept as,

"…being activated by these extreme changes in seasonal light. With drawings, models, and calculations, the towering, shifting space of the interior was organized around light and time. Window openings were cut according to the movement of the sun through building section. The interior volume catches particular moments of light as an index of time" (Holl, 2011, p. 109).

In his seminal book titled, The Eyes of the Skin: Architecture and the Senses, Juhani Pallasmaa states that "In great architectural spaces, there is a constant, deep breathing of shadow and light; shadow inhales and illumination exhales light" (Pallasmaa, 2005, p. 47). While daylight interacts with and alters the visual appearance of every space it enters, what Pallasmaa refers to here is the dynamic interplay between light and shadow that great architecture is inclined to manipulate and employ. While not all daylit spaces can be considered great architecture, it would be difficult to find even one great architectural space where daylight is not manipulated to induce highly controlled visual effects. Pallasmaa (2005) later goes on to criticize our contemporary valuation of light by stating that "In our time, light has turned into a mere quantitative matter and the window has lost its significance as a mediator between two worlds, between enclosed and open, interiority and exteriority, private and public, shadow and light" (p. 2). What is perhaps most striking about this assertion is the idea that these 'worlds' as Pallasmaa describes, require

mediation – that they create a kind of living system through which architectural form plays an active role in choreographing a deeply dynamic perceptual experience.

In her book titled, The Architecture of Light, Mary Ann Steane discusses the key differences between daylight and artificial sources in lighting design. While artificial light sources can be carefully combined to match a desired composition, they produce a static effect and can never match the "nuance of mood created by the time of day and the wonder of the seasons" (Steane, 2011, p. 7). While Holl and Pallasmaa describe qualities that are primarily formal – sun, space, volume, time, shadow, interiority/exteriority -Steane's description of a 'nuance in mood' begins to address a dimension of daylight that has been little explored by scientific researchers. The idea that daylight can induce and alter our emotional state within the built environment challenges the idea that a two-dimensional photometric measure like illuminance can be used to evaluate daylight performance (Rea, 2000) without a complimentary set of measures that can evaluate perceptual impacts on the human, at the human scale.

Peter Zumthor describes the power of daylight in similar terms,

"Thinking about daylight and artificial light I have to admit that daylight, the light on things, is so moving to me that I feel it almost as a spiritual quality. When the sun comes up in the morning – which I always find so marvelous, absolutely fantastic the way it comes back every morning – and casts its light on things, it doesn't feel as if it quite belongs in this world…For an architect that light is a thousand times better than artificial light" (Zumthor, 2006, The Light of Things).

The visual effects of daylight; such as shadow, depth, contrast, and texture are highly valued by architects and often regarded in the narrative of design, but are not often described in quantitative terms. Aside from the obvious challenges associated with measuring qualities that are inherently subjective, our current discourse on lighting performance has focused on characteristics that are more objectively calculated, such as horizontal illuminance requirements or visual discomfort. While there is still some debate among experts over accepted criteria for illumination and even more so for visual comfort, these criteria share a focus on performance within working environments and are often threshold driven. This means that there is a specific quantity of something (i.e. illuminance, luminance, or a ratio of those) that has been determined to support visual acuity, energy savings, or on the other end of the spectrum, trigger discomfort. The 'almost spiritual' qualities that Holl and Zumthor have described, are revealed through our visual perception of space and as such, existing task-based and threshold-driven daylight metrics can not effectively evaluate them. Furthermore, we lack cohesive terminology and criteria for describing the visual effects of light between spaces that vary in composition and geographic location. This lack of structured terminology poses a complex problem when architects want to discuss the perceptual quality and design intent of daylight to engineers who are evaluating the visual acuity, comfort and energy autonomy.

Disciplinary challenges that emerge from the cultural divide between architects and engineers can deter the pursuit of topics that link subjective characteristics of light to quantifiable performance

models, but this rift can also creates a ripe environment for interdisciplinary inquiry. Reflection on the disconnect between design intent and performance evaluation in daylight raises a fundamental and perhaps abrasive question: Why do existing evaluation tools focus almost exclusively on horizontal illumination and visual discomfort, avoiding the positive impacts of visual perception? We design for composition, ephemerality, and spatial experience, yet due to the subjective nature of these factors, the building sciences too often consider them to be soft concerns. Many perceptual factors of daylight can in fact be measured and this thesis seeks to develop an evaluation method to better understand the dynamic impacts of light on human emotion in architecture. Section 1.1.1 will return to this topic again.

One could argue that this tendency towards surface-based performance measures like horizontal task illumination emerged from the very fundamentals of architectural representation. The most conventional methods for describing constructive and spatial design intent rely on 2D and 3D modes of orthographic projection (plan, section, and axonometric) such as can be seen in Figure 1.1. These modes allow us to describe the complex physical and organizational logic of architecture through a scalable and translatable language. Elements of architectural, structural, mechanical, and electrical systems can be overlaid and integrated, allowing for efficient coordination between disciplines. Architects and engineers are trained to read these conventions as a constructive syntax, but with any proper syntax comes a lack of flexibility in describing elements outside the language. Non-dimensional elements of space, such as light, heat, sound, and smell are not easily communicated using the existing conventions of orthographic representation. Steven Holl describes these very physical, but non-structural elements of spatial experience in the Chapel of St. Ignatius as

"Seven bottles of light in a stone box; the metaphor of light is shaped in different volumes emerging from the roof whose irregularities aim at different qualities of light: East facing, South facing, West and North facing, all gathered together for one united ceremony." (Holl, 2016, February 15)

To communicate these 'bottles of light,' Holl uses watercolor and perspective sketch to diagram his design intentions (Figures 1.2 - 1.3). When we look at photographs of this architectural project against plan and section documents, we understand the limitations of orthographic representation for communicating the immersive spatial experience he describes. The narrative of these light-driven qualities that Holl is known for producing cannot be described using orthographic projection, because they are qualities that must be experienced in 3-dimensions from the perspective of an observer who is immersed in space. He writes that the new generation of architects are

"interested in the plasticity of human sensory response and who manipulate experience and perception by deforming the ambient environment with sound, light, weather, and chemistry together represent but a shimmer on a new horizon toward which we are moving rapidly today." (Holl, 2011, p. 88).

Holl is not alone in his vision of space as a plastic sensory environment. In describing the Winter House, Phillipe Rahm describes a way to "broaden the field of architecture to the design of the invisible, of electromagnetic fields…" (Rahm, 2017, January 14). Architectural space, as

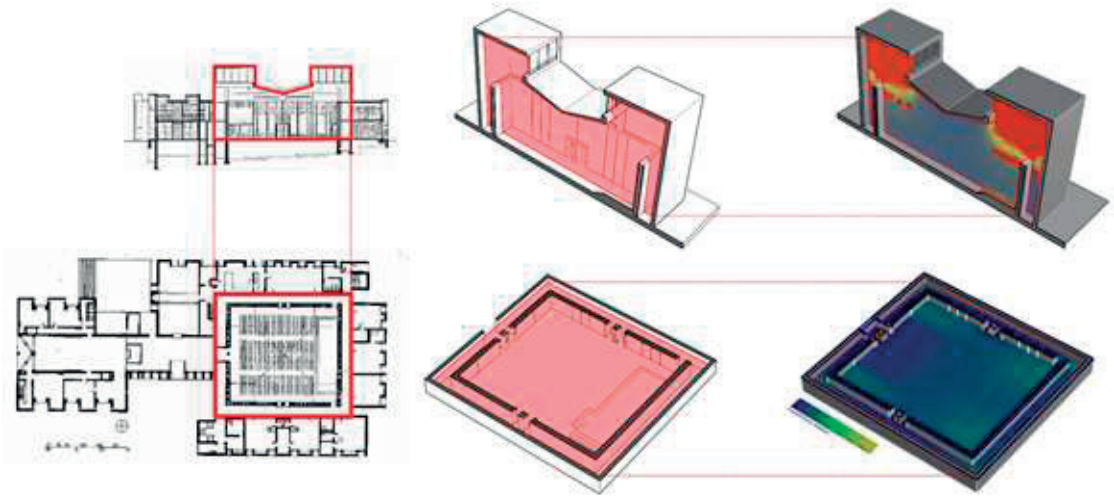


Figure 1.1 Orthograhic views of Louis Kahn's First Unitarian Church, with corresponding section and plan cuts showing how horizontal and vertical illuminance is often extracted from a 3D architectural space.

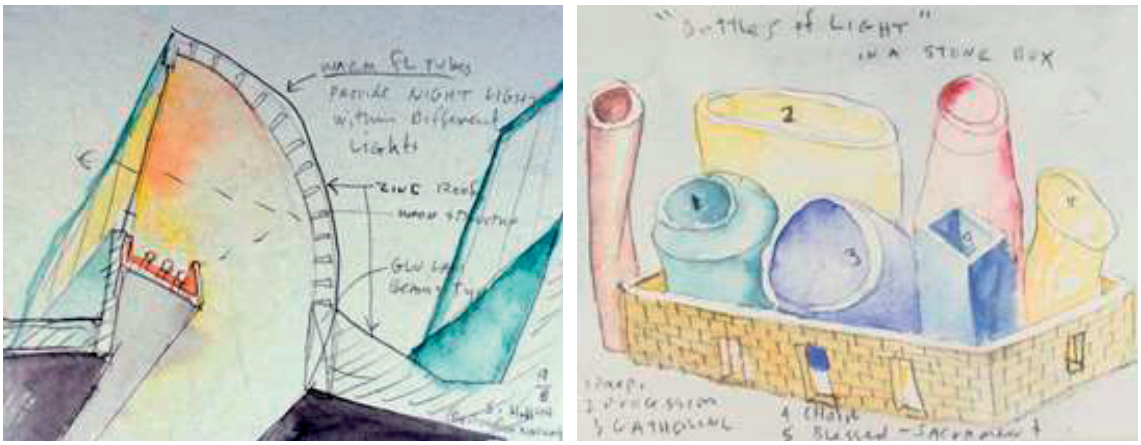


Figure 1.2 Watercolor drawings of the Chapel of St. Ignatius. See image credits page.

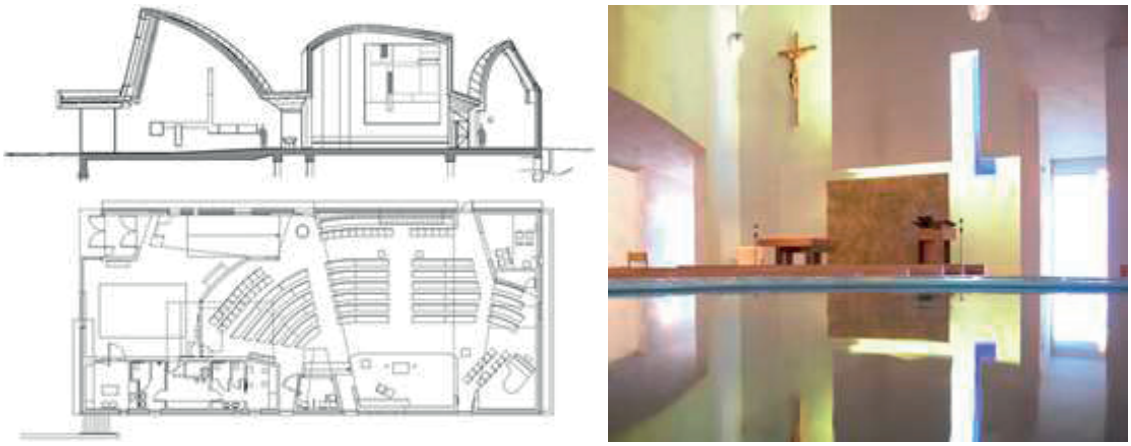


Figure 1.3 Plan and Section through the St. Ignatius Chapel by Steven Holl on the left with interior photograph on the right. See image credits page.

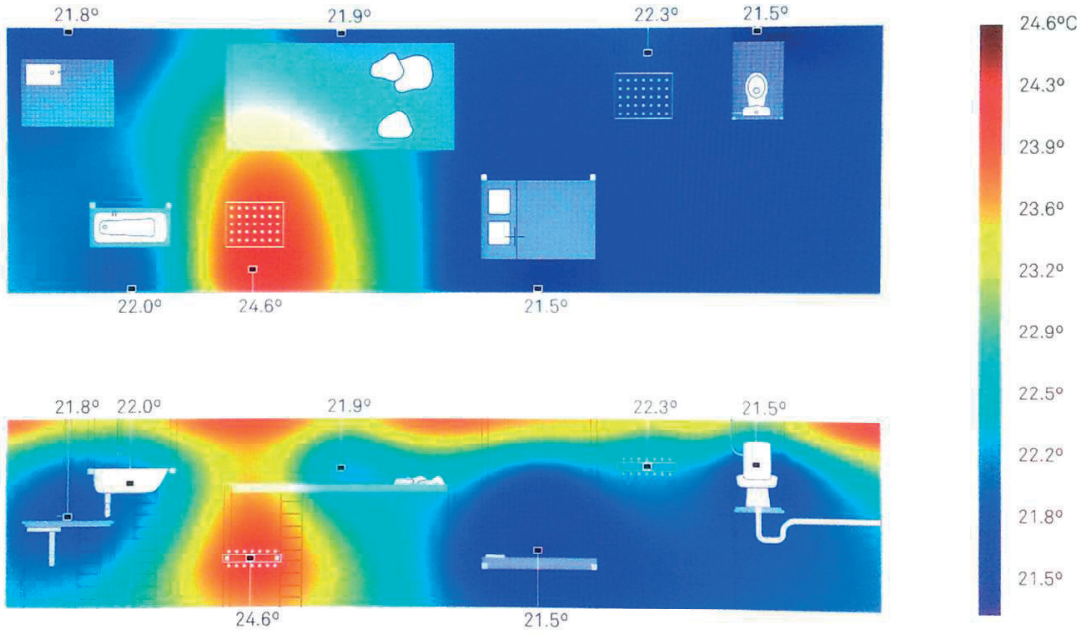


Figure 1.4 Domestic Astronomy by Phillippe Rahm. See image credits page.

traditionally defined by structured elements has, for Rahm been re-envisioned to "an architecture of air, invisible but physically modified." He no longer sees the experience of space encapsulated by orthographic convention,

"...no longer imagines...[it simply as]...a void, as an absence defined by walls, floor, and ceiling but as a less dense mass, disconnected, transparent and yet nevertheless filled with material; a void invisible to the eye, certainly, but in which the body is immersed." (Rahm, 2017, January 14)

For Holl, one might imagine the simple answer to his dilemma would be perspective rendering, which allows the designer to capture a snapshot of time and illustrate all the environmental, material, and human-related qualities that would fully define this experience. Perspective renderings, drawings, and walk-through animations are often used to communicate those aspects of design intent which rely on human perception or environmental influences, such as sunlight and programmatic function (i.e. other people, furniture, artifacts). While it is one thing to produce a static or even dynamic (in the sense of a walk-through) animation of architectural space, it is increasingly more difficult to describe how that space is transformed by the ephemeral conditions of the surrounding environment. For Rahm, whose plastic environment bridges the visual domain with thermal sensation, sunlight, weather, wind, heat, and sound, the spatial experience of architecture is a live phenomenon, in constant flux with nature and human interaction (Figure 1.3). We would have to create animations that are not only immersive, but sensually encapsulating and temporally dynamic.

If we rely on conventional representation to drive analysis methods in daylight performance, then it makes sense that most of our existing metrics use 2D orthographic projections of illuminance.

It is easy to coordinate and overlay this type of data as one would an electrical or mechanical plan, but much like the limitations of orthographic projection in describing the live phenomenon of experience in architecture, 2D projection does not help us measure or represent elements of daylight that we witness from an immersive human perspective. If we are conditioned to evaluate daylight through a 2D task-driven surface, then how can we describe, objectify, and evaluate experiential qualities of daylight from a human perspective?

1.1.1 The Evolving Role of Daylight Performance

To help situate this thesis within an appropriate historical context, this section will present a brief overview of the emergence of daylight as an element in building performance alongside other environmental factors so that we may understand why research expanded in some areas and narrowed in others. In the field of architecture, the last several decades have seen a dramatic shift in focus, from issues of post-war urbanization and housing to the energy crisis of the 1970s and a growing concern for carbon-neutral and environmentally-responsible building design. The term *environmental* has been used to define a broad range of social and economic initiatives (U.N., 1987), which – in the field of architecture – has translated most pervasively to the pursuit of energy autonomy and carbon-neutral building practices. The relationship between building and environment dates back to the very foundations of architectural practice. That being said, the advent of new technologies to understand, control, and reduce the impacts of human habitation on environmental degradation marks a dramatic transition in the history of building technology. Addington (2003) writes, "No other problem during the 20th century mobilized the public and private sectors across such a wide swath in such a short time." (p. 1).

Until the turn of the 18th century, there was very little change in our heating, cooling and ventilation practices, which were entirely passive and relied on low-grade sources of fuel (Banham, 1969). With the introduction of the Edison lamp in 1880 and the AC motor in 1888, the industry of artificial lighting was born. By 1910, it is estimated that 1 in every 10 houses in America had electricity and by 1930, 7 out of 10 homes were wired (Banham, 1969). With the invention of indoor Air Conditioning and its widespread distribution throughout the 1930s, architecture developed a more autonomous relationship from the external environment, relying on electric light deep within the building core and artificial 'coolth' where un-filtered sunlight was emitted through fully-glazed façade systems (Ackerman, 2010). Parallel advances in Modern Architecture of the 1930s and 40s show a gross misunderstanding of the ecological impacts of unrestricted energy consumption. In his book titled, *Precisions*, originally published in 1930, Le Corbusier (1991) introduces a mechanically-driven air system referred to as "respiration exacte":

> *"Every nation builds houses for its own climate. At this time of international interpretation of scientific techniques, I propose: one single building for all nations and climates, the house with respiration exacte." (Corbusier, 1991)*

The move toward an international style coincided with advances in electricity and artificial cooling which occurred throughout the 1930s and 40s and precipitated some rather fantastical

ideas about the potential for unlimited energy use and artificial climate within buildings. In the New York World Fair of 1939, the Carrier Corporation installed an exhibition they called the 'igloo of tomorrow,' complete with artificial snow to showcase the power of air-conditioning (Cooper, 1998). This fantasy of unlimited energy for artificial indoor environments continued its rapid expansion into the early 1970s, as evidenced by the high-tech architecture of the 1960s and may have continued were it not for the U.S. oil crisis of 1973 (Banham, 1969). With a sudden peak in energy costs, the age of *environmentally* conscious design was born.

As the building sciences continue their rapid transformation into a field of critical importance for contemporary architecture, it is important now, more than ever, that we re-examine our definitions of and methods for measuring performance. We are conditioned to think that zero net-energy is the ultimate target in environmental design and while no one denies the importance of building energy use, the centrality of energy has become pervasive in our current discourse. As daylight can help offset energy consumption and create high-quality illumination autonomous from electricity, energy-related performance considerations have taken precedence over seemingly less urgent, but important human-centered topics. Comfort, perception, and health have taken a back seat to more immediate environmental concerns. Due to the interdisciplinary nature of research in these topics, funding resources have also been more difficult to secure as researchers found themselves trying to fit into either humanities, life-sciences or engineering-based grants.

In the field of civil engineering, there is a rich history of collaboration between architect and structural engineer, dating from the master builders of the 16th century, who transcended the professional silos of disciplinary boundary, to contemporary practice, where there are numerous examples of positive collaboration (Calatrava, 1997; Allen & Zalewski, 2009; Ochsendorf, 2013). This is due, in part, to the invasive impact of structure on form, which requires early and continuous collaboration throughout the design process. In mechanical and environmental engineering, however, where the 'material' (i.e. air, heat, and water) is less visually integrated than structure, there often exists a tenuous relationship between architect and engineer. When it comes to daylight evaluation, much of the criteria and tools for passive building performance have been developed by engineers, computational scientists, and physicists, with little input from architects on the complexity and reiterative nature of the design process. Working on the periphery of the discipline through groups that specialize in modeling and simulation and less often on design integration, there often exists a pedagogical disconnect between the architect, who guides and coordinates the design process, and the engineer, who analyses performance to verify targets and recommend adjustments post-hoc.

To further complicate this relationship, many performance metrics in building science (especially in energy use and thermal comfort) are non-spatial, creating a block between the visual and structured language of architectural design and the quantitative language of evaluation. If the thermal comfort of a building is outside an acceptable range, how do the numbers that define that analysis translate to a new design recommendation?

In lighting photometry, the earliest measurements were developed by French engineers in the 1890s in response to the boom in gas lighting and were used to recommend the height and distribution of fixtures to create a uniform appearance (Cuttle, 2004). Candle power, a unit of luminous intensity developed to describe the burning rate of a spermaceti candle, was an early unit used to describe lighting performance (Cuttle, 2004). As electric light was introduced in the 1920s and beyond, illuminance levels and ratios across a horizontal plane became the primary measure behind most lighting standards. The notion of our legal 'right-to-light' didn't emerge until the following decades, when the footcandle came to be regarded as a sufficient unit of light for reading (Waldram, 1954).

In daylight analysis, early performance measures were developed with the goal of providing illumination for visually demanding tasks. These studies sought to recommend illuminance levels to compensate for task difficulty using measurements gathered on-site in schools and offices (Weston, 1961; Rea & Ouellette, 1984; IES, 1966). While these studies revealed that illuminance levels as low as 6-12 lux could be used to meet reading task requirements for healthy adults, recommendations of 300 - 1000 lux (for schools and offices respectively) or higher were found in that year's edition of the IES Lighting Handbook (IES, 1966). Why the substantial increase in recommended levels? Christopher Cuttle argues that

"The illuminance levels that must be provided to meet peoples' expectations for adequately illuminated surroundings exceed the levels that they need in order to cope with the simple visual tasks that they encounter in modern, well-equipped workplaces. While lighting standards may claim to be performance-based, this is false, as the levels specified cannot be justified on the basis of visual performance" (Cuttle, 2004, p. 76).

These simple photometric recommendations for various task-oriented activities were developed to ensure visual acuity during daylight hours, yet ensuring task acuity does not necessarily equate to the desired appearance of a space during those tasks. Cuttle goes on to state that

"More generally, when people in workplaces equipped with modern, efficient lighting complain about the lighting, their objections are likely to be directed towards the appearance of their surroundings. They may find the appearance of the workplace to be dull or gloomy, or the effect of the lighting to be harsh, producing dense and unattractive shadows" (Cuttle, 2004, p. 76).

We will return to Cuttle's argument, but we must first acknowledge parallel advances in horizontal illumination metrics. To better understand the annual performance and overall electricity consumption that resulted from a rapid increase in artificial light through the 1980s and 90s, daylight metrics began to shift away from single point-in-time lux measures. By computing the influence of sky conditions and occupancy, climate-driven illumination metrics could help designers reduce the demand for electric lighting. Due to advances in computational capacity, metrics like Daylight Autonomy became the new preferred target in contemporary practice (Reinhart, et al., 2006). Ever since their emergence, these climate-based illumination metrics have dominated the field of contemporary research, promoting concerns related to energy autonomy over other important design criteria such as visual comfort, human health, and perceptual experience.

Numerous studies have sought to quantify the effects of window geometry, glazing type, and surface reflectance on attributes of perception and preference (discussed in more depth in Section 2.2). That being said, most of these studies resulted in categorical recommendations (i.e. window size, glazing transmittance values, etc) and not generalizable methods to predict these effects. Those methods that have been developed to measure brightness, diversity, or distribution, do not integrate a time-series to account for dynamic effects or provide a computational approach that could be integrated into a simulation-based method. In other words, existing methods have measured the effect of various parameters on subjective attributes, but have not produced a metric to predict these effects.

Our perception and appreciation of daylit space is largely defined by the ephemeral human-centric conditions in our environment, and so like Cuttle, we began to ask ourselves: Why are existing metrics still focused on non-perspectival, threshold-driven, and surface-based measurements? Rather than attach ourselves to existing analysis methods, perhaps we should explore the characteristics of light that really define daylight performance. How can we predict these characteristics and what kinds of tools and methods can be developed to integrate perceptual daylight considerations within the design process? In Chapter 2, we will present a more systematic review of past and current literature regarding daylight performance measures, but first we will discuss the role of perceptual quality more broadly, building the motivation for this dissertation and work within.

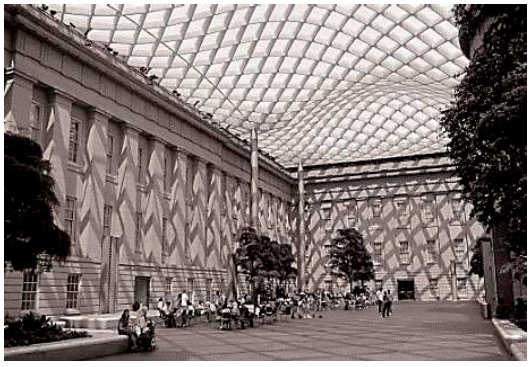


Figure 1.5 Kogod Courtyard, Norman Foster. See image credits page.

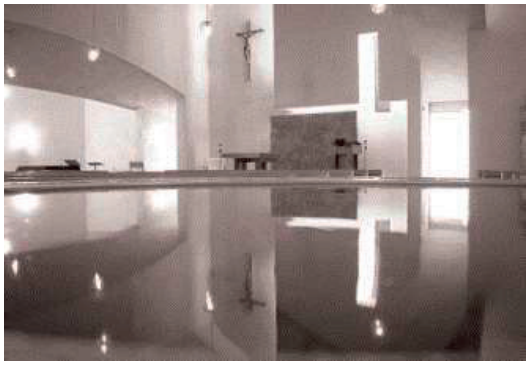


Figure 1.6 Chapel of St. Ignatius, Steven Holl. See image credits page.

1.1.2 The Perceptual Quality of Daylight

To define the perceptual characteristics of daylight performance in contemporary architecture, we will look at a range of strategies, starting with a comparison between three examples which differ in their strategy for integration, and end with a categorical and typological matrix. In our first three examples of contemporary architecture, we will discuss the differences inherent in their expression of spatial and temporal variability. The first example is Norman Foster's renovation of the Kogod courtyard of the Smithsonian American Art Museum Washington, D.C (Figure 1.5). Completed in 2007, the articulated glass roof, which was inserted into the existing building, emits direct sunlight through a 'fishnet' pattern of light and shadow across the walls and floor of the interior (Ouroussoff, 2007). Designed for

Figure 1.7 The Chicago Art Institute Modern Wing, Renzo Piano. See image credits page.

temporary occupation and public gathering, the space does not require a tightly controlled lighting strategy. On the contrary, the Kogod Courtyard uses transparency to create a diverse and visually engaging environment, embracing direct sunlight and dynamic visual effects.

The second example is the Church of St. Ignatius in Seattle, Washington, designed by Steven Holl (Figure 1.6). Unlike the Kogod Courtyard, this space transforms sunlight through a series of soft and indirect luminous forms (Holl, 1999). This architectural space creates a carefully calibrated glow (Ryan, 1995) and produces a dramatically different use of sunlight than the Kogod Courtyard. The interior maintains a dynamic relationship with exterior light levels as shifting sun angles and weather patterns create smooth, yet dramatic transformations to the chapel.

The third example is the Modern Wing at the Chicago Art Institute, designed by Renzo Piano Building Workshop and completed in 1997 (Figure 1.7). Art galleries require controlled lighting environments to protect artwork from UV damage and minimize veiling glare. The roof of the modern wing is composed of two layers: the first layer is made up of white, curved louvers that block direct sunlight, while the second layer contains translucent glass to further diffuse incoming light. As a result, the galleries receive diffuse and uniform daylight which is dynamic in overall brightness, but relatively static in contrast levels due to the lack of direct sunlight.

These three examples illustrate the diverse role of luminous composition in the visual performance of daylit architecture. While each of the strategies varies in its integration of spatial and temporal diversity, all three could be considered successful in achieving an intended set of visual effects and ambient luminosity which strengthen an occupant's spatial experience. Furthermore, the composition of luminance levels, rather than mean luminance or luminance range, appears to play a critical role in our perception of contrast. These three examples are exemplary of a larger global survey of contemporary architecture which was originally presented by the author in (Rockcastle & Andersen, 2011, 2013a) and then refined and discussed again in (Rockcastle & Andersen, 2013b) to catalogue a range of light-based visual effects and rank the resulting typologies in terms of spatial and temporal variability.

To develop the matrix of architectural examples presented in (Rockcastle & Andersen, 2013b), each architectural example was studied using the authors' trained intuition and then positioned within a linear gradient to represent the degree of perceived spatial and temporal variability within each photograph. When individual images appeared to contain similar contrast characteristics,

Figure 1.8 Matrix Showing Spatial and Temporal Variability for 50 Architectural Spaces (High Spatial Contrast & Temporal Varibaility on the Left to Low Spatial Contrast & Variability on the Right).

they were added to an existing category, but when images showed unique characteristics, they formed a new category. Although a total of 75 architectural spaces were initially placed into 15 categories (Rockcastle & Andersen, 2013a) the authors narrowed the final matrix down to 10 categories, each of which contain 5 exemplary spaces (Rockcastle & Andersen, 2013b) (Figure 1.8).

The horizontal axis of the matrix shows a linear gradient from high spatial and temporal variability on the left to low spatial and temporal variability on the right. The 50 examples presented here were taken from across the world and represent a diverse mix of architectural designers. Each photograph was selected for its representative perspective of the interior space and most photographs show choreographed views chosen by the architects for publication purposes.

Those typologies that fall on the left end of the spectrum are labeled as Direct and Exaggerated, Direct and Dramatic, and Direct and Screened. The Direct and Exaggerated category includes highly variable top-lit spaces such as the Kogod Courtyard by Norman Foster and the Milwaukee Art Museum by Santiago Calatrava. The Direct and Dramatic category includes side-lit spaces that emit large light patches, such as the Mikimoto Store by Toyo Ito and the Zollverein School by SANAA. The Direct and Screened category contains examples of facades or roofs that emit small, but frequent patches of direct sunlight, like the Benavidas Warehouse by Guillermo Hevia and the Dominus Winery by Herzog and deMeuron.

Those typologies that fall toward the middle of the spectrum are labeled as Partially Direct, Direct, Selectively Direct, and Direct/Indirect. The Partially Direct category contains side-lit spaces that emit sunlight through louvers or repetitive façade elements. Spaces in this category include the Magney House and the Fletcher Page House by Glenn Murcutt. The Direct category includes side-lit spaces with minimal obstructions (i.e. no louvers) such as the Bomballa Farmhouse by Collins & Turner or the Farnsworth House by Mies van der Rohe. Selectively Direct contains spaces that emit sunlight in discreet instances, such as the Tulach a Tsolais monument by Scotta Tallon Walker or the Imperial War Museum by Daniel Libeskind. The Direct/Indirect category is composed of spaces that emit sunlight through thickened openings in the building envelope, resulting in both direct sun patches and an indirect wall wash. Spaces in this category include the Poli House by Pezo Von Ellrichshausen and Notre Dame de Haut by Le Corbusier.

The categories that fall toward the right end of the spectrum are labeled Spatial Indirect, Indirect, and Indirect & Diffuse. The Spatial Indirect category is defined by spaces that emit indirect light across interior surfaces. This category includes spaces such as the Chapel of St. Ignatius by Steven Holl and the First Unitarian Church by Louis Kahn. The Indirect category contains spaces that emit indirect light through north facing monitors or openings in the roof. Spaces in this category include the Dia Beacon Museum by Open Office and the High Museum of Art by Renzo Piano. And finally, the Indirect and Diffuse category contains spaces that utilize diffusing surfaces to minimize the dynamic effects of light and shadow. The Chicago Art Institute by Renzo Piano and the Louis Vuitton Building by Jun Aoki are examples from this category. These ten categories, although not exhaustive, illustrate a broad range of daylight strategies in contemporary architecture.

Using the architectural matrix as reference, a simplified spatial model for each of the ten categories in Figure 1.9 was created during the final phase of the author's master thesis work (Rockcastle & Andersen, 2011). These simplified spatial models were used to generate annual renderings and compare the impacts of daylight over time through a preliminary algorithm developed to measure the compositional contrast within each image (Rockcastle & Andersen, 2012). A brief summary of this work will be presented in Chapter 3, as this doctoral thesis builds upon the simulation-based study begun during the author's master thesis (Rockcastle, 2011).

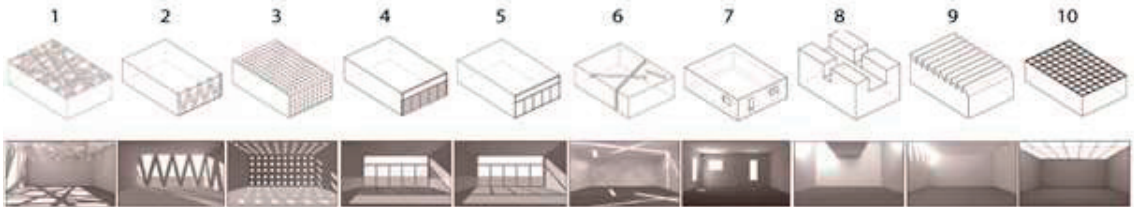


Figure 1.9 Matrix Showing 10 Typological Models (High Spatial Contrast & Temporal Variability on the Left to Low Spatial Contrast & Variability on the Right.

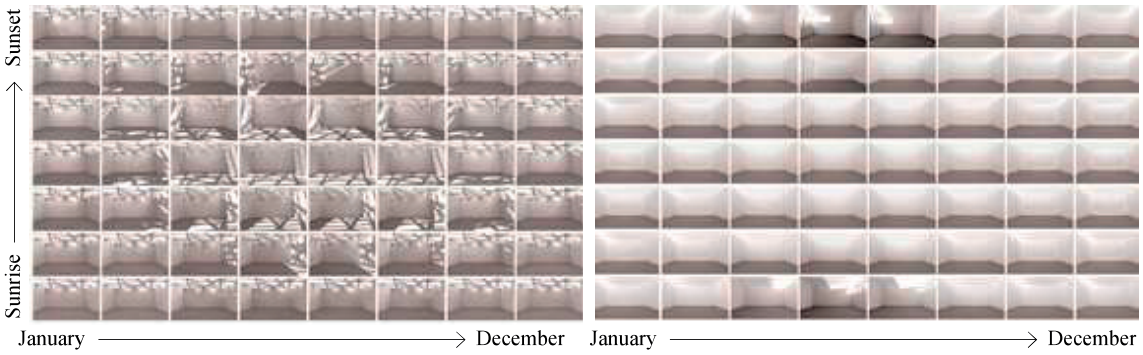


Figure 1.10 Annual Renderings for a Two Top-lit Spaces in Boston 56 Radiance renderings with even daily hourly and daily subdivisions to represent a full year, Latitude 42N, sunny skies.

An example of these annual renderings is shown in Figure 1.10, which reveals the degree of variability that occurs throughout a selected view from two abstract top-lit spaces. The date and time of these 56 renderings was established using a time-segmentation method explained in more depth in Chapter 3.

The Direct & Exaggerated top lit space in Figure 1.10a shows a highly contrasted interior with variable strength and composition due to the temporal dynamics of sunlight. The Indirect top-lit space in Figure 1.10b, however, shows a relatively static interior with low contrast - except for sunrise and sunset in the summer months when sun penetrates the North-facing roof monitors. These annual sets of images show the degree of luminous variability that occurs throughout each selected view and illustrates the need for prediction algorithms that can assess the spatial and temporal diversity of light from an occupants' perspective. While spatial diversity can be analyzed within a static image, temporal diversity (resulting from daylight) requires a time-series of images taken throughout the year to help designers evaluate the strength and diversity of contrast-based perceptual effects over time.

The images shown in Figure 1.9 capture a single point of time each and are meant to illustrate a similar gradient of effects as the full architectural matrix in Figure 1.8. While photographs of existing architectural spaces provide us with more complex information about the effects of sunlight throughout our visual field, HDR renderings of abstract spatial models allow us to more objectively compare the resulting perceptual effects as they vary over time from a fixed view position.

1.2 Thesis Structure

The authors' master thesis ended with a preliminary set of measures called Spatial Contrast and Luminance Variability. While promising in their ability to differentiate (in relative terms) between the abstract spatial models shown in Figure 1.9, comparisons to subjective contrast perception were limited to the author's own intuition. In Section 1.1, we introduced a broad historical and cultural overview of concerns surrounding the integration of daylight performance in architectural design. In chapter 2, a range of experimental studies that address the topic of brightness, distribution, and daylight diversity in driving various subjective assessments of occupant perception will be presented.

Using the simple geometry models introduced in Section 1.1.2 and presented in more detail in Chapter 3, Chapter 4 will compare an expanded series of contrast-based algorithms. Using a selection of those algorithms, Chapter 5 will then introduce an Online experiment using rendered images of more complex architectural models to collect subjective ratings related to visual interest. Data from this experiment will then be compared to the contrast-based algorithms and a prediction model for predicting exciting and stimulating attributes will be proposed. In Chapters 6, this model will be applied to a simulation-based workflow for a perspectival field-of-view and then across an array of view positions in Chapter 7. In Chapter 8, a final experiment using visually-immersive rendered scenes will collect responses from subjects wearing a virtual reality headset. To evaluate the predictive capability of our proposed model under immersive view conditions, algorithms proposed in Chapters 5-7 will be compared to subjective data from the virtual reality experiment.

The following section will outline the main objectives of this thesis and summarize each of the distinct phases of research. This overview of thesis structure assists the reader in understanding the structure and trajectory of the work.

1.2.1 Problem Statement

Daylight is a dynamic source of illumination which creates variable impacts on our perception of architecture due to hourly, seasonal, and weather-based fluctuations. While many prominent architectural theoreticians have stated the importance of daylight in our visual perception and emotional responses to the indoor environment, there is a lack of research seeking to measure, define, and/or predict these responses. Visual characteristics of daylight such as brightness, contrast, and diversity have shown some dependence to occupant preferences and appraisals in the indoor environment, but there are no robust simulation-based models for predicting these responses across space and over time. A simulation-based method for predicting the variable effects of daylight on impressions of visual interest would help designers to understand *where*, within a given scene, daylight can drive emotional responses, *when* those responses may occur, and *how* strongly they are affected by changing sky conditions. This thesis will propose a computational method to measure elements of contrast and predict impressions of visual interest in rendered scenes, using a model that was derived from subjective data in experimental studies.

1.2.2 Thesis Objectives

1) To evaluate the relationship between changing sky conditions and subjective evaluations of composition, contrast, and visual interest in rendered daylit scenes.
2) To develop an algorithm for predicting characteristics of visual interest in daylit architectural renderings using subjective responses from user experiments.
3) To develop a simulation-based framework that can assess the strength and temporal changes in of visual interest across a series of view positions and view directions within a digital architectural scene.
4) To provide architects with the ability to evaluate the perceptual performance of their own designs alongside existing performance measures for a more holistic assessment of daylight in architecture.

1.2.3 Phases of Research

The five phases of research (Phase 0 – Phase 4) in this thesis are outlined below:

Phase 0: Evaluating Dynamic Contrast: A Proof-of-Concept Approach *Chapter 3*

This phase will introduce a pai of algorithms called spatial contrast and annual spatial contrast. These measures, which were originally proposed by the author through a proof-of-concept simulation-based study, is then expanded upon during the initial phases of this dissertation (Rockcastle & Andersen, 2011, 2013b, 2014). This phase of research uses a set of 10 abstract case study spaces, modelled in Rhinoceros and rendered in Radiance to compare the temporal variability of contrast between spaces and across an annual time series. As the author did not have access to subjective data on contrast perception in this phase, she used her own intuition to rank images from high to low and then compares those relative rankings to predictions from the proposed algorithms. This comparative approach serves as a preliminary proof-of-concept upon which further validation and development will be addressed in the following phases of research.

Phase 1: Comparison of Existing Contrast Algorithms *Chapter 4*

Phase 1 of this thesis uses the renderings developed for Phase 0 to compare a broad range of image-based algorithms developed to predict contrast perception. Whereas Phase 0 relied on a single-level neighbourhood algorithm modified from (Rizzi, et al., 2004) and then applied to a time-series, Phase 1 looks at a broader set of methods, both global and local in approach. The following image-based algorithms have been selected for this comparative study: Michelson (Michelson, 1927), RMS (Pavel, et al., 1987), Spatial Contrast (SC) (Rockcastle & Andersen, 2014), RAMM (multi-level average: RAMMG and single level: RAMM1) (Rizzi, et al., 2004), Difference of Gaussian (DOG) (Tadmor & Tolhurst, 2000), and RSC (Simone, et al., 2009). After implementation into MATLAB (https://www.mathworks.com) and application to each of the case study renderings, results for each algorithm are normalized and compared to the author's own

intuitive image ranking using linear regression. A short list of algorithms will then be identified for use in the following phase.

Phase 2: Web-Based Experiment on Visual Interest *Chapter 5*

In Phase 2, an online survey has been designed to collect subject ratings of contrast and visual interest in 2D renderings of existing architectural spaces which vary in the complexity and strength of contrasting luminance values and dynamic daylight composition. After distributing the survey and collecting responses, subject ratings were then compared to the selection of global and local contrast-based algorithms from Phase 1 to search for relationships between the data and predictive measures. While Phase 0 used simplified renderings of shoe-box models, Phase 2 uses higher quality renderings of detailed existing architectural spaces, rendered across a range of instances to test the effect of both spatial and temporal parameters. From the regression analysis between subjective data and contrast algorithms, a new measure called modified spatial contrast 'mSC' will be proposed due to its goodness of fit with ratings of calming - exciting, subdued – stimulating, diffuse – direct, and low contrast - high contrast. After fitting mSC values for each rendering against the distribution of subjective ratings for that image, a proportional odds model was extracted and will be used in the following phase to predict impressions of calm and excitement. This work was published in part in (Rockcastle, et al., 2015; Rockcastle, et al., 2016).

Phase 3: Application of Visual Interest Model to a Simulation Workflow *Chapter 6 & 7*

The third phase of this thesis proposes a simulation-based strategy for implementing the mSC algorithm within an immersive human-centric view format. In Chapter 6, the mSC algorithm was first adapted to a hemispherical image format and then applied to a series of view directions from a fixed central view position. Using the proportional odds model from Phase 2, this hemispherical implementation allows for the evaluation of multiple view directions from a fixed eye-level position to predict the proportion of subjects that would rate that view direction as calming, neutral, or exciting. A simulation-based framework will be introduced in Chapter 7, to propose a spatially-immersive approach across *an array* of view positions in an architectural case study, allowing the user to compute the performance of mSC alongside a non-visual health-based metric called non-visual direct-response nvRD (Amundadottir, 2016). This work was published and/or submitted in part in (Amundadottir, et al., 2016; Rockcastle, et al., 2017b).

Phase 4: Visually Immersive Experiment on Visual Interest Using VR *Chapter 8*

The final phase of this thesis introduces an immersive 3D study using projected scenes in virtual reality. Designed for the Oculus Rift CV1 headset, chapter 8 describes an experiment using semi (180°) and fully (360°) visually-immersive rendered scenes for a series of eight architectural spaces. For each space, mSC was used to select a high instance of spatial contrast from a semi-

annual time-series using an adapted 3D algorithm. High-quality 360° spherical renderings were then generated under clear and overcast sky conditions for this selected instance to provide a range of relatively high and low contrast conditions for each scene. Subjects were randomly assigned into one of two groups: one that saw 180° scenes and one that saw 360° scenes and were then asked to rate the scenes using a series of verbal response scales. In the 180° group, subjects saw a randomization of two view directions (one high and one low view direction relative to the 18 possible view directions present in the scene). By varying view direction and sky condition within this virtual environment, this experiment compares subject ratings within a visually immersive virtual environment to predictions using the mSC measure, alongside other variants of this algorithm. This experiment seeks to validate the effectiveness of mSC at predicting recorded responses under visually-immersive scenes, as proposed in Chapters 6 & 7. This work was published in part in (Rockcastle, et al., 2017a).

Conclusion & Outlook *Chapter 9*

To conclude, Chapter 9 will summarize the findings of each phase of research. Following a general discussion about limitations, we will provide an overview of future steps for experimental research, including a discussion on the potential applications of this research within the architecture and building performance professions.

2 STATE OF THE ART

This chapter will present a state of the art in several parallel areas of scientific study: task-based daylight performance, perceptual daylight performance, the digital display of images for experimental use, and algorithms to predict contrast perception. Starting with existing daylight performance metrics, both illuminance and luminance-based, this literature review will cover the evolution from static to dynamic illumination and visual comfort indicators. Transitioning from task-based visual acuity to the perceptual field-of-view, we will move into a review of relevant experimental approaches for assessing visual aspects of daylight in buildings and end with a review of display methods used to test various subjective qualities in digital scenes. In our review of display methods, we will also discuss the impacts of tone-mapping and visual immersion on experimental studies related to lighting perception. Shifting from daylight applications into the fields of image processing and computer vision, we will then explore a range of algorithms developed to predict contrast in digital images, discussing their limitations and potential application in qualitative daylight assessments. This interdisciplinary review will situate the thesis between complimentary, but often disparate fields of study - daylight, physics, vision, psychology, and image processing - to underline the need for new approaches to measure the positive perceptual impacts of daylight within our field-of-view.

2.1 Existing Task-Driven Daylight Performance Metrics

The most common daylight measures used today typically focus on task performance, whether in regards to illumination (task-specific horizontal *illumuinance* measurements) or to visual comfort (eye-level glare risk using *luminance* measurements).

2.1.1 Illuminance-Based

Over the past several decades, there have been significant improvements in our understanding of daylight as a dynamic source of illumination. We have transitioned from static illuminance-based measures such as Daylight Factor DF (Moon & Spencer, 1942), to annual climate-based metrics like Daylight Autonomy DA (Reinhart et al., 2006), Continuous Daylight Autonomy (cDA) (Rogers, 2006), Daylight Saturation Percentage (DSP) (CHPS, 2006), Useful Daylight Illuminance UDI (Nabil & Mardaljevic, 2006), Acceptable Illuminance Extent AIE (Kleindienst & Andersen, 2012), and Spatial Daylight Autonomy (sDA) (Heschong, et al., 2014). These climate-based metrics address the performance of horizontal illuminance across the day and year, accounting for a more statistically accurate method of quantifying internal illumination levels (Mardaljevic,

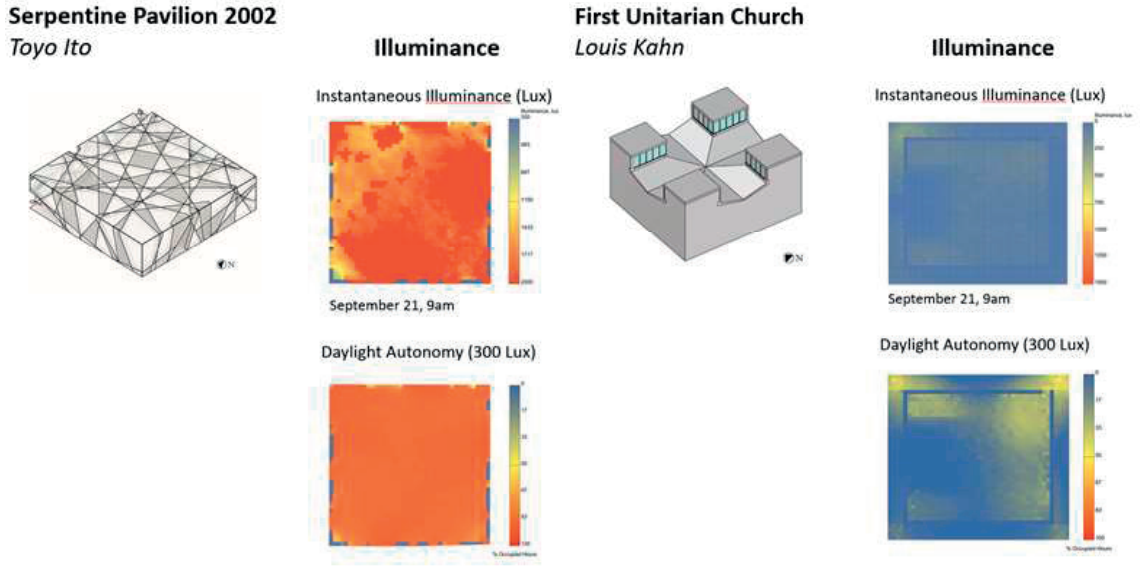


Figure 2.1 Illuminance-based measures (instantaneous illuminance on the top and daylight autonomy on the bottom) for two spaces which vary in architectural form: The Serpentine Pavillion by Toyo Ito on the left and the First Unitarian Church by Louis Kahn on the right.

2000). Despite improvements in computational capability that have enabled the application of climate-based analysis methods, DF remains one of the most widespread task-based illumination metrics currently used in practice. This measurement, which calculates the ratio between an internal illuminance level under overcast skies and a simultaneous exterior measurement, limits our understanding of daylight, assuming a more-is-better attitude regardless of the sky type or intended programmatic use of the space under consideration (Reinhart, et al., 2006). Dynamic illumination metrics such as DA, cDA, DSP, UDI, AIE or sDA can evaluate annual illumination thresholds, taking into account building orientation and climate-driven sky types to provide a more accurate assessment of task-plane illuminance. While there are numerous studies that have shown the advantages of using climate-based illumination metrics (Mardaljevic, et al., 2009) over static point-in-time measurements like DF, the implementation of these metrics has been slow to reach the level of standardization. LEED v.4 now requires an integrated approach using sDA and Annual Sunlight Exposure (ASE) (Heschong, et al., 2014; USGBC, 2014) after more than a decade of debate on the subject. Comfort, health, and perceptual metrics will arguably face even more scrutiny during an attempted implementation into mainstream criteria and standards, but have found a home in widespread software tools designed for architects, like DIVA for Rhino (http://diva4rhino.com/), Sefaira (http://sefaira.com/), and Lightsolve (http://lightsolve.epfl.ch/). Figure 2.1 shows an example of static and climate-based illuminance metrics applied to two architectural geometry models using the DIVA for Rhino toolbar. A path towards the implementation of non-illumination-based measures and holistic daylight evaluation criteria will be discussed later in this thesis, with an emphasis on simulation-based approaches in Chapters 6 & 7.

2.1.2 Luminance-Based

Unlike task-based illumination metrics that rely on illuminance, visual comfort metrics (typically pertaining to glare) rely on luminance (CIE, 1926). Of the four photometric quantities (luminous flux, intensity, illuminance, and luminance), luminance is most closely related to how the eye perceives light, and as such, appears as the only quantity capable of expressing visual discomfort. As luminance, brightness, and contrast are subjectively evaluated, methods to analyze glare discomfort are fragmented across no less than seven established metrics (Wienold & Christofferson, 2006; Rea, 2000; Osterhaus, 2005). While these indices do not always agree, partly due to the fact that some were developed for electric lighting sources (Hopkinson, 1970) and others for daylight, most are derived from the same four quantities: luminance, size of glare sources, position of glare source, and the surrounding field of luminance that the eye must adapt to (Wienold, 2009). Glare-based visual comfort metrics, such as Daylight Glare Probability DGP (Wienold & Christofferson, 2006), considered the most reliable index for side-lit office spaces under daylight conditions and Daylight Glare Index DGI (Hopkinson, 1970), have also evolved into dynamic annual metrics such as DGPs (Wienold, 2009) which provides a comprehensive yearly analysis of glare, with limited computational intensity (Jakubiec & Reinhart, 2012) and DGIw, which provides an annual summary of visual comfort performance across 5-minute time-lapse images under clear sky conditions (Konis et al., 2011). Figure 2.2 shows an example of DGP applied to two architectural geometry models over a single instance and also a set of annual instances. Note that this example does not use the annual approach defined by Jakubiec & Reinhart (2012) or Konis et al. (2011), but is merely a point-in-time application over 56 annual moments using the time segmentation method proposed by Kleindienst & Andersen (2008).

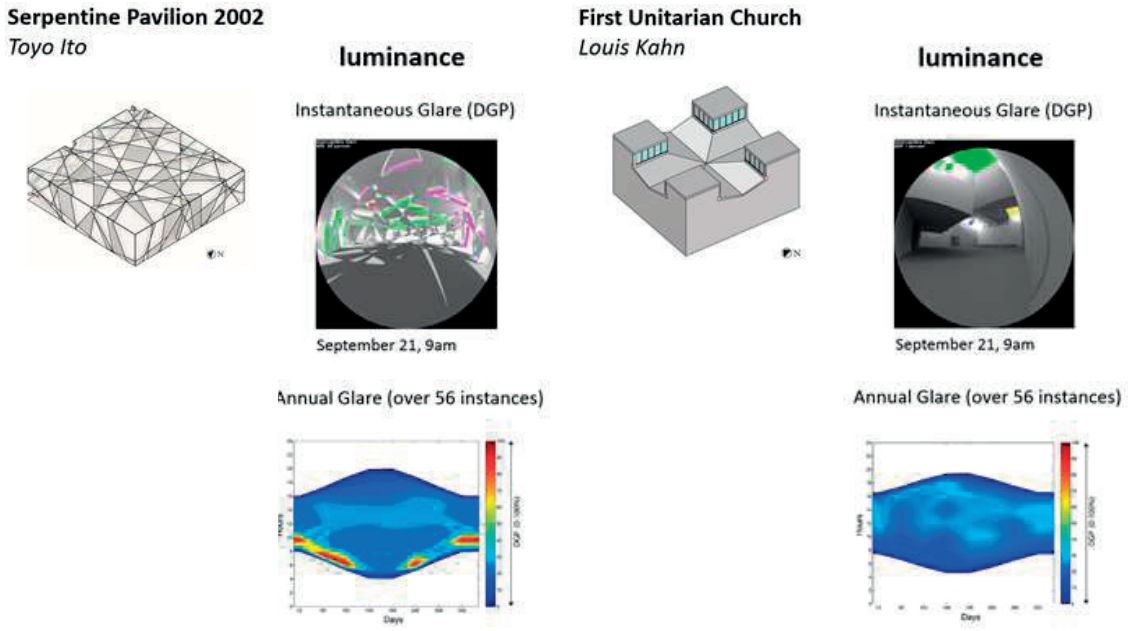


Figure 2.2 Luminance-based measures (daylight glare probability DGP) for two spaces which vary in architectural form: The Serpentine Pavillion by Toyo Ito on the left and the First Unitarian Church by Louis Kahn on the right.

While discomfort metrics have evolved to accommodate hourly and daily time-steps as well as climate-based inputs, perhaps the most important development in these measures has come from an expanded understanding and integration of view position in comfort prediction measurements. View-dependent metrics, such as those pertaining to daylight glare (DGP and DGI), have traditionally depended on a single view position, framed through a 2-dimensional hemispherical, or fisheye lens. Glare based on the assumption of a fixed view position can be problematic for two obvious reasons: 1) subjects are free to shift their gaze away from the potential glare source and 2) gaze behaviour is likely impacted by dynamic conditions such as those which are present in a daylit space. The 'adaptive zone' approach introduced by (Jakubiec & Reinhart, 2012) address the first of these concerns, while the 'gaze-driven' approach introduced by (Sarey Khanie et al., 2015; Sarey Khanie, 2015) focuses on the second. Whereas illumination metrics are generally independent of occupant behaviour such as pupil or gaze-direction, glare models are highly dependent on such inputs and have undergone rigorous study to predict how the visual system is affected by variable conditions both compositionally and in terms of climate and time.

2.2 Experimental Approaches to Study Perceptual Appraisal

While climate-driven illumination and task-driven comfort have seen substantial improvements in model development and integration within performance guidelines and software packages over the last decade, the development of quantifiable measurements for predicting the perceptual performance of daylight has been limited. There have been numerous studies relating subject preference to factors such as window size and view, some of which produce generalizable recommendations for dimension or ratio of opening (Ne'eman & Hopkinson, 1970). To evaluate the visual impacts of luminosity within daylit architecture, existing research has compared subjective assessments of rendered images, photographs and/or real scenes to luminance-based measures derived from those scenes. These studies have most often relied on global measures such as average luminance or '*brightness*', threshold luminance, luminance ratios or '*contrast*', and luminance variation or '*standard pixel deviation*' in line with occupant surveys to seek which (if any) values can be used to predict preferential ratings. The computation of these measure is most commonly extracted from high-dynamic-range HDR images, digital photographs or renderings produced using raytracing software such as Radiance, which provide an expanded range of photometric information, allowing for the evaluation of characteristics such as brightness and contrast (Ward, 1994; Newsham, et al., 2002; Inanici, 2003).

Two factors that are widely accepted to impact the field-of-view in daylit architecture are average luminance and luminance variation (Veitch & Newsham, 2000). The former has been directly associated with perceived lightness and the latter with visual interest (Loe, et al., 1994). Some studies have found that both mean luminance and luminance variation within an office environment contribute to occupant impressions of preference (Cetegen, et al., 2008; Newsham, et al., 2005, 2010), whereas others have discovered that luminance distribution across an occupant's field-of-view (Boubekri, et al., 1991; Tiller & Veitch, 1995) as well as the strength of variation are factors of preference (Wymelenberg & Inanici, 2009; Parpairi, et al., 2002).

To discuss the state of the art behind qualitative lighting research, this section will begin by introducing experimental studies that address perceptual lighting quality. Table 2.1 shows a range of subjective experiments that have been categorized by: 1) real spaces, 2) scaled architectural models, and 3) digital images – both printed and projected on digital display medium. For each of the relevant studies, we will introduce the experimental setup, discuss the scene or image type (real, physical model or photograph/rendering), and subjective rating methods used to conduct the qualitative research.

As the compression of luminance values from a real space to a digital image is a subject of particular importance when evaluating subjective impressions of lighting in an experimental set-up, section 2.3 will discuss this issue in more depth. Section 2.3 will cover a review of experimental strategies aimed at validating the effect of specific display devices (LDR, HDR, 2D, panoramic, stereoscopic, and 3D virtual reality) on subject evaluations of emotional and preferential characteristics of daylight in digital images (both photographic and rendered). As both the capture technology, display device, and tone-mapped luminance range of any digital image are critical for an accurate evaluation of lighting levels within a given scene, this section will end with a look at specific tone-mapping operators and their ability to mimic real world scenes for use in qualitative lighting research.

Table 2.1 – Experiments assessing subjective qualities of daylight

Real Spaces		*Year:*	*Image Capture:*	*Scene Display Method:*	*Evaluation Method:*
1.	Boubekri et al.	1991	luminance & illuminance meters	Real space (private office) under varied sky conditions	Subjective ratings of emotional response and preference
2.	Tiller & Veitch	1995	Digital camera	Real space (IRC subjective reactions lab)	Brightness matching (using lighting controls)
3.	Veitch & Newsham	2000	Digital camera	Real space (open-plan office)	Manipulation of lighting & surveyed preference rating
4.	Parpairi et al.	2002	Luminance meter	Real space (libraries at Cambridge University)	Semantic differential ratings (9 scales)
5.	Dubois	2007	Digital camera	Real space (café)	Observation of spatial distr. And task of occupants
6.	Wymelenberg & Inanici	2009	Digital camera + HDR technique	Real space (private office)	Manipulation of blinds and surveyed preference rating
7.	Ozdemir	2010	Illuminance meter	Real spaces (private offices in campus building with range of orientations & view)	Semantic diff. scales (view, satisfaction, brightness, spaciousness)
8.	Moscoso et al.	2016	Digital camera & illuminance meter	Two test rooms with varied window area & surface reflectances	Semantic differential ratings (9 scales)

Physical Models		*Year:*	*Image Capture:*	*Scene Display Method:*	*Evaluation Method:*
9.	Ne'eman & Hopkinson	1970	N/A	1:12 arch. model with varied layouts/view/sky cond.	Subjective ratings of preference toward window geom. & dim.
10.	Keighley	1973	N/A	1:12 office model with varied window geometry and projected cityscape views	Subjective ratings of visual satisfaction and acceptability.
11.	Dubois	2007	N/A	1:7.5 arch. Model with varied glazing materials	Semantic differential ratings (shadows, light level, beauty, sharpness, etc.)

Digital Images		*Year:*	*Image Capture:*	*Scene Display Method:*	*Evaluation Method:*
12.	Newsham et al.	2005	Digital renderings (Lightscape)	2D LDR Image projection ≤ 140 cd/m2	Semantic differential ratings (attractiveness)
13.	Cetegen et al.	2008	Digital camera + HDR technique	HDR LCD display ≤ 2,700 cd/m2	Semantic differential ratings using 8 scales
14.	Lo & Steemers	2009	Digital renderings (Radiance)	2D Image projection ≤ ? cd/m2	Semantic differential ratings (profane/ sacred, relaxing/ stimulating, dull/ dynamic, functional/ poetic)
15.	Franz et al.	2012	Digital Renderings (Radiosity 360° panoramic)	Spherical wide-angle projection (Elumens VisionStation)	Semantic differential ratings (pleasure, interest, beauty, normality, calm, spaciousness, brightness, openness)

2.2.1 Real Spaces

Those studies that have been conducted in real space are generally limited to typical office environments (open and closed plan) or test rooms and allow subjects to manipulate what is often a single set of experimental factors, such as window blinds, lighting controls, surface distribution, or sky conditions (Boubekri, et al., 1991; Tiller & Veitch, 1995; Veitch & Newsham, 2000; Wymelenberg & Inanici, 2010). These spaces are typically viewed by a participant, manipulated by either the subject or the controller to produce a variety of luminous effects, and then evaluated using subjective ratings or in some cases, character matching (like brightness). To compare these subjective ratings to quantitative measures, photometric quantities are then captured using HDR photography, luminance or lux meters. From these HDR images, luminance data can be extracted and photometric quantities are compared to rating responses to explore relationships between quantitative and qualitative factors.

In a study by Boubekri, et al. (1991), subjective responses for a set of emotional factors were collected in a typical office environment, whereby 40 subjects were divided into two groups: one that saw the window from a frontal position, and one that saw the window from the side. Within each group, subjects were shown 4 possible window-to-wall-area ratios which resulted in 4 different sunlight penetration-to-floor-area ratios. Subjects were asked to evaluate each scene using a seven-point semantic differential scale for words related to excitement and relaxation, afterwhich they were asked about their satisfaction. The research team found a significant effect of window size on ratings of relaxation in the group with side-exposure and a moderate effect on ratings of excitement, but only weak significance with the other scales. Due to the small population size (5 subjects per factor), it is difficult to draw further conclusions about these findings, although the researchers do suggest a sunlight penetration of 15 – 25% is ideal (with higher percentages acceptable) with small sun patches preferred over a large flood of light.

In a study by Tiller and Veitch (1995), subjects were asked to match the brightness between mock office scenes, where the distribution of luminances were varied. The authors concluded that a non-uniform luminance distribution across an occupant's field-of-view increased their perception of brightness over a uniformly lit office space. This study also concluded that 5-10% less task-plane illuminance was required in the non-uniformly lit office space, proving that luminance distribution can impact more than just qualitative lighting preference. Qualitative issues like brightness and daylight distribution can have an impact on general illumination requirements and ultimately building energy use.

A study by Veitch & Newsham (2000) asked subjects to manipulate electric lighting controls to achieve a preferred set of luminous conditions in a mock office space. In pairs of two, one subject selected his/her desired luminous settings and the two subjects then worked for the remainder of the day under those fixed settings. Lighting satisfaction measures were collected at the end of the session with questions regarding the lighting quality and overall environmental satisfaction. The authors reported high satisfaction ratings for the selected scenes, but were not able to draw broader conclusions about what combination of indirect and direct lighting conditions were the most desirable as there was too much variation between selected conditions and subject responses. When subject ratings were compared to luminance ratios, the authors suggest that desired variations may be higher than suggested by the IESNA and CIBSE, which suggest a ratio of 10:1 (from task to surround brightness). In this study, ratios in preferred scenes ranged from 11:1 to 68:1. Furthermore, the most preferred luminous conditions for office work may not be the most interesting scene as determined by (Loe, et al., 1994). The authors call for consensus on a method and procedure for measuring luminance variation across the field-of-view as they believe that the *distribution and frequency of these variations may be desirable* depending on their size and location.

To address the importance of light distribution, Parpairi et al. (2002) developed the Luminance Difference Index (LD index), which proposes a spatially dependent method for measuring luminance diversity across a selected view position. This method relies on eye-level luminance

measurements to calculate the difference in luminance levels across a range of acceptable angles corresponding to eye and head movement. An experiment which calculates the LD index across three selected view positions in real world settings, found that luminance variability was highly appreciated by the participants and that variability rather than intensity were found to contribute to occupant satisfaction. This method by Parpairi, et al. (2002) is quite compelling in that it links an objective measure of daylight diversity to subjective responses about perceptual preference. The draw-back to this method is, however, the reliance on physical measurements taken during a single point in time, and which require a space that is already constructed or a model that is already built. Daylight is a dynamic source of illumination and understanding its impact of the appearance of architecture requires many instances over time and between sky conditions.

A study by Wymelenberg & Inanici (2009), asked participants in a side-lit office space to adjust a set of horizontal blinds until the light distribution was most preferable and then again into a position that they felt was just disturbing. These scenes were photographed and then compared to a number of existing luminance measures, applied to the HDR photographs. While DGP showed no significant trends between the most preferable and just disturbing spaces, threshold luminance values and standard deviation did produce some dependence to preference. The authors of the study concluded that *adequate variations* in luminance tended to create a *stimulating visual environment*, while excessive variability tended to create uncomfortable spaces. In line with Veitch & Newham's suggestion, there seems to be some agreement that luminance variation can be a positive factor in lighting preference, but the degree of variation and its location in the field-of-view has yet to be determined (Veitch & Newsham, 2000).

To study the role of luminous diversity on the preferred spatial distribution of occupants in a daylit space, Dubois, et al. (2007) conducted a study in a student café, whereby behavioural mapping and HDR photography was used to determine where occupants chose to sat while conducting various activities under different natural lighting conditions. Findings revealed that the majority of subjects chose to occupy the brightest zones of the space near the windows, but the preferred locations for tasks which require even the highest visual acuity (like reading and writing) were diverse and often took place in parts of the space which would not meet IES recommendations. In a self-reported survey, 'lighting' was chosen as the most important factor in the choice of location for their spatial distribution, but the spread in this distribution suggests that diverse lighting conditions, while not always meeting prescribed criteria or standards, may provide more opportunity for a population to find spaces that fit their individual wants and needs. This supports the thesis by Steane & Steemers (2004) which advocates for a diversity of environmental conditions to meet a diverse set of human wants and needs.

A study by Ozdemir (2010) studied how view openness and naturalness are related to a users' satisfaction with room, perceived spaciousness, and view. This experiment was conducted in 18 seperate rooms (offices) in a campus building in Turkey. The researchers first asked expert landscape architects to assess pictures of the windows, captured from an eye-level interior view position using two 5-point bi-polar scales (open – closed and natural – built). They then asked

occupants in each office to rate their view satisfaction, room satisfaction, perceived spaciousness, and brightness to see if expert assessments correlated with user surveys. The authors found that openness was correlated to user ratings of perceived spaciousness, brightness, and room satisfaction while naturalness was correlated to view satisfaction, room satisfaction, and perceived spaciousness. The decision to relate descriptive expert ratings to evaluative user responses makes this an interesting experimental approach.

To study the impact of window size and room reflectance on perceived quality attributes in two small test rooms, Moscoso, et al. (2016) used paired attributes and semantic differential ratings related to order, complexity, pleasantness, and spaciousness. Experiments were held at midday and subjects were divided into two groups - one that saw a black test room (reflectance = 5%) and one that saw a white test room (reflectance = 87%). In each group, a random presentation of stimuli was presented (with three levels of window size). Results showed that 8 out of 9 rating scales were significantly impacted by window size and all 9 scales were impacted by wall reflectance (white vs black interior). The authors found strong linear correlations between many of the rating scales, like pleasantness and excitement, spaciousness and openness, and spaciousness and spatial definition. The authors also plotted mean luminance against semantic ratings to see a positive linear relationship between higher luminance and higher ratings of pleasantness (amongst others). The caveat behind this finding is that the only variables were surface reflectance and window area for one spatial layout, so we cannot generalize higher mean luminance as a predictor for higher pleasantness without comparing more spacial configurations.

2.2.2 *Physical Models*

Those studies that have used scaled architectural models are often similar in procedural methodology to those studies conducted in real space, but they have the advantage of being able to display a broader range of spatial or material configurations in an efficient manner (Ne'eman & Hopkinson, 1970). While architectural models are more easily manipulated than full-scale spaces, if not purely for the economics of material and installation, they still have the same temporal limitations in terms of sky condition. To recreate realistic photometric qualities in the test apparatus, both real spaces and scaled architectural models are limited to the current sky condition at the test location. This makes it possible to conduct experiments over a series of days or months, but difficult to compare subject impressions over time. While sun angles can be more easily manipulated in a physical model than a real space, the accuracy of resulting sky conditions are questionable.

In a study by Ne'eman & Hopkinson (1970), a 1:12 scaled architectural model was shown to subjects who were asked to adjust the window width to define a minimum opening below which he would lose satisfactory contact with the outside. The following variables were altered for each subject: window heights, number of windows, size of room, outside view, and weather/ daylight levels. Two observations were made; one without electric light, and one with overhead fluorescent light. Illuminance measures were collected inside the model alongside sky luminance.

The authors found an average window width of 7.5 ft (for a 7 ft window), but this was dependent on the subject's view angle to the outside and the dimensional proportions of the room. Other findings suggest that windows serve a function beyond that of task illumination and that while artificial light may satisfy specific task needs, quality and connection to the outside may be secured by windows which do not necessarily provide much daylight. The authors commented on the difficulty associated with testing a broad combination of variables and the complicated relationship between window size and room size. They also confirmed that accordance between results from the physical model and a parallel appraisal in the field suggest that a model apparatus can be a valid analogue for investigating psychophysical problems of vision and light in buildings.

In a study by Keighley (1973), subjective ratings of visual satisfaction and acceptability were collected using a 1:12 office model with variable window geometries. The model represented a full scale room with a mirrored side wall. Subjects were asked to rate scenes within the model while aperture, window height, window area, mullion width, and type of view were altered by the researcher. The author found that subject's satisfaction was proportionally affected by widow area and inversely proportional to the number and width of mullions. In other words, people preferred larger windows with smaller and fewer mullions. Subjects also preferred a wide lateral view of the horizon and skyline.

To evaluate the effects of coated glazing materials on subjective impressions of view and interior in scaled physical models, a study by Dubois (2007) used two identical scale models (1:7.5) of an office room. During the experiments, samples of various glazing materials were applied to the window opening while subjects viewed the interior of the model from a hole on the other side. Using one model as a base case (with clear double-glazed window), the other model was used to display 6 glazing material samples. Subjects were asked to first view the base case model and answer a questionnaire, after which they were asked to do the same for the second model. The questionnaire used 7-point bi-polar scales related to shadows, light level, light colour, comfort (glare), naturalness, beauty and pleasantness, and sharpness. Scales were sometimes negative-positive and other times positive-negative. The effect of various glazings (when compared to the base case) was significant on all scales ($P<0.01$) except for those related to shadows. The author also found that glazing types of with higher transmittance resulted in more positive ratings for factors of naturalness, beauty, pleasantness, and sharpness.

2.2.3 Digital Images

Studies based on digital images have the advantage of using a much broader range of variables that are not dependent on real-time sky conditions or sun positions. Where real spaces provide a realistic distribution of luminance levels and physical models can reproduce daylight distribution instantaneously, digital images are more easily manipulated and reproduced to provide a range of experimental variables. The limitations associated with display luminance and factors related to scene reproduction for perceptual accuracy will be addressed in Section 2.3.

In a study by Newsham, et al. (2005), subjects were shown a projected set of rendered images with varied outputs for 4 independent sets of luminaires. Images were projected to 33% of real size using a standard projector with compressed luminance range ($\leq$ 140 cd/m2). Ratings were collected using a 10-point scale, with bi-polar ratings for attraction (ugly – beautiful, pleasant – unpleasant, comfortable – uncomfortable), non-uniformity (varied – unvaried, simple – complex, and uniform – nonuniform), and brightness (bright – dim, dark – light, radiant – murky). To compare qualitative ratings to quantitative measure, various image-based values were computed for each projected scene, namely; WAV (average luminance of the image) and RMS (non-uniformity). The authors found that ratings of brightness were related by WAV, while ratings of non-uniformity were related to RMS. Attractiveness was related to a combination of WAV and RMS, or in other words – brightness and non-uniformity. The authors concluded that rendered scenes could be used as a powerful surrogate to conduct qualitative research for lighting, although this study was conducted with artificial and not natural sources.

To study the impacts of various design parameters (panel height, panel material, and window blind condition) on occupant satisfaction in an office setting, Cetegen et al. (2008) asked subjects to rate a series of HDR photographs, presented on a HDR monitor (3 - 2608 cd/m2). Asked to imagine the scene as their own workplace, subjects viewed each image for 30 seconds and then rated it using a series of eight semantic scales. The authors discovered a positive trend between increased average luminance levels and satisfaction for the view in an office setting, but they also saw a trend between increased *luminance diversity* and the participant's *impression of excitement.*

In (2009), Lo & Steemers published a paper which summarizes two experimental approaches: one using projected images of abstract rendered scenes and photographs and another field-based study in two Le Corbusier buildings in France. The authors collected subjective responses for two bi-polar pairs using a seven-point scale, profane – sacred, and functional – poetic, after which they asked some open-ended questions. In the image-based study, subjects were asked to rate the abstract renderings (which varied in window shape and number of windows) using the discrete 7-point scales. They were then shown photographs of real daylit spaces and asked to answer questions about the space use and lighting quality. The researchers observed a positive correlation between brightness and both perceived functionality and profanity as well as between contrast and both sacredness and poeticism. In the field-based study, subjects were asked to rate their perception of the bi-polar scales mentioned above, alongside discrete contrast perception in two famous buildings. Luminance and illuminance measure collected on site during the field study allowed the authors to compare subjective ratings to physical measure. Similar to the image-based study, the authors saw positive correlations between brightness and functionality and profanity as well as contrast and ratings of sacredness. What is perhaps most interesting in this approach is the decision to use abstract renderings - which offer controlled testing parameters - alongside real spaces - which offer more realistic scenes, but less control. The trade-offs of these two approached will be discussed later in the thesis to justify approaches used by the author in Chapters 5 & 8.

In a study by Franz, et al. (2012) radiosity-rendered interior scenes with varied window area, wall openness ratio, balustrade height, size of single window, room area, room length/width, and door height were presented to a group of participants in a spherical, wide-angle projection system (Elumens VisionStation™). Subjects were allowed 30 seconds to freely explore these scenes (presented in random order) and then asked to rate the scenes using seven-point bi-polar likert-scales for the following semantic pairs: unpleasant-pleasant, boring-interesting, ugly-beautiful, strange-normal, arousing-calm, narrow-spacious, dark-bright, and enclosed-open. The authors found a negative correlation between ratings of beauty and increased proportions of length and width. They also found a positive correlation between ratings of beauty and increased proportions of width/height. The authors also saw high correlations between ratings of beauty, pleasure, and interestingness and these ratings were most highly correlated to physical openness. Surprisingly for the authors, window proportion showed no significant effect on these same ratings. In a post-experiment evaluation, subjects rated their sense of presentness as moderate and were highly content with the appropriateness of display medium and realism of the simulations. While the focus in this experiment was on the relationship between room and window proportion and subjective ratings under a consistent set of illumination and surface properties where daylight was not a factor, the use of an immersive display device is relevant to this thesis.

In the following section, our literature review will focus on those experiments where *display medium* is, itself, the parameter being studied. So far, we have seen that factors (and or measures) related to contrast, non-uniformity, brightness, variation, proportion, and diversity have shown some significant relationships to subjective ratings of pleasantness, attractiveness, beauty, excitement, and interestingness. In the next section, we will look at those studies that seek to validate the use of 2D and 3D display devices and modes for use in experimental studies relating to daylight perception.

2.3 Strategies for Assessing Display Devices for Experimental Use

Where the previous section looked at studies that used real spaces and digital images to collect and compare subjective assessments of daylight qualities, this section will look at studies where the *type of scene display* (real space, 2D, 3D projection and/or stereoscopic device) is the parameter being tested. In other words, this section looks at experimental studies where one or more displays is compared to real world conditions (or other display devices) in an attempt to measure the impact of their specific projection type on subjective factors. These studies seek to understand which projection methods are most perceptually accurate and useful as surrogates for real world scenes. Table 2.2 shows experimental studies organized by 2D and 3D/stereoscopic display device.

Experimental studies that rely on 2D image projection have compared real spaces to digital photographs of real space and/or renderings (Salters, et al., 2012; Engelke, et al., 2013) as well as digital photographs presented to subjects in either HDR ($\leq$ 4,000 cd/m2) or LDR mode ($\leq$ 200 cd/m2) (Newsham, et al., 2010). While we grouped the study by Murdoch & Stokkermans (2014)

in with others that investigate the role of 3D projection, this experiment also looks at impact of HDR luminance display modes on 2D devices. A study by Newsham et al. (2010) used a HDR LCD display developed by Seetzen, et al. (2003) and Whitehead, et al. (2005) which can display luminance values up to 4,000 cd/m2, while Murdoch & Stokkermans used a screen with display luminance values up to 1,800 cd/m2 . The importance of tone-mapping operators (TMOs) in the projection of HDR images on either LDR or HDR displays will be discussed in the following section, but in this section we will focus on the display rather than the TMO.

Experiments that compare real world scenes to images (both photographic and digital) projected in 2D and 3D display devices have relied on a series of techniques; pair-wise comparison, semantic differential ratings, and single question comparisons. A study conducted by Cauwerts & Bodart (2011) compared semantic differential ratings between real spaces and projected photographs of those spaces using both 2D and 3D projection methods. Studies relying on digital renderings, instead of photographs, include work by Villa & Labayrade, (2010; 2015), and Murdoch & Stokkermans (2014). A final study by Chamilothori, et al., (2016) uses the Oculus Rift DK2 headset to compare semantic ratings in a real world scene to renderings of that scene projected in virtual reality.

Table 2.2 - *Digital display devices in experimental use*

2D displays		*Year:*	*Image Capture:*	*TMO:*	*Scene Display:*	*Evaluation Method:*
16.	Newsham et al.	2002	Images provided by (Vietch & Newsham, 1998)	N/A	2D LDR Image projection ≤ 110 cd/m2 and HDR LCD display ≤ 1,800 cd/m2	Semantic differential ratings (attractiveness)
17.	Newsham et al.	2010	Digital camera + HDR technique	N/A	LCD display in HDR mode ≤ 4,000 cd/m2 & LDR mode ≤ 200 cd/m2	Semantic differential ratings (attractiveness)
18.	Salters et al.	2012	Digital Renderings (Vray+Indigo and LightTools)	Reinhard iCAM	Real Space and LDR screen ≤ 200 cd/m2	Semantic differential ratings (brightness, diffuseness, contrast, uniformity, shadow visibility)
20.	Engelke et al.	2014	Digital camera + digital renderings	Reinhard	Real Space, LDR screen ≤ 400 cd/m2	Semantic differential ratings (overall impressions, light distribution, and perceived atmosphere)

3D/Stereoscopic displays		*Year:*	*Image Capture:*	*TMO:*	*Scene Display:*	*Evaluation Method:*
21.	Villa & Labayarde	2010	Digital Renderings (Mental Ray, Maxwell, V-ray, Inspirer)	Reinhard	Real space vs. renderings (printed) and then pair-wise stereoscopic simulator ≤ 60 cd/m2	Ranking on a linear scale and pair-wise comparison
22.	Cauwerts & Bodart	2011	Digital 3D camera + HDR technique	Reinhard	Comparison: real space & projected photographs (2D & 3D) ≤ 300 cd/m2	Semantic differential ratings (17 scales)
23.	Cauwerts & Bodart	2013	Digital 3D camera + HDR technique	Reinhard	Comparison: real space, (2D & 3D) ≤ 300 cd/m2	Semantic differential ratings (17 scales)
24.	Villa & Labayarde	2013	Digital Renderings (V-ray)	Reinhard	Online uncontrolled vs. stereoscopic 3d simulator ≤ 170 cd/m2	Question regarding suitability of environment for office work
25.	Heydarian et al.	2015			Comparison: real space vs virtual space in Oculus Rift DK2	Object identification, reading speed and comprehension
26.	Chamilothori et al.	2016 - 2017	Digital Renderings (Radiance)	PCOND	Real Space vs. rendering of real space using Oculus Rift DK2 ≤ 100 cd/m2	Semantic differential ratings (pleasant, interesting, exciting, complex, satisfied with view)

2.3.1 2D Displays

In the first of several studies by Newsham et al., photographs of 3 office test rooms originally captured and used in a study by Veitch & Newsham (1998), were presented to subjects using 2 display modes with scenes luminance compressed to HDR (<1,800cd/m^2) and conventional LDR (<100cd/m^2). Both display modes were presented in random order on a 17" backlit screen in a black box. Subjects were asked to rate each of the presented images using 15 bi-polar adjective scales (bright - dim, uniform - non-uniform, interesting - monotonous, pleasant - unpleasant, comfortable - uncomfortable, stimulating - subdued, radiant - gloomy, tense - relaxing, dramatic - diffuse, spacious - cramped, glaring - not-glaring, friendly - hostile, simple - complex, formal - casual, realistic - unrealistic) registering their rating using a pencil stroke on a 100mm line printed on paper. Subjects were also asked to rate their agreement with a number of satisfaction-based scales (from strongly disagree to strongly agree) ranging from satisfaction with the workplace to satisfaction with the lighting. The authors hypothesized that rating of brightness would be higher for the HDR than the LDR images and that higher luminance contrast images (HDR) would lead to lower satisfaction. Their findings supported that hypothesis, but also showed that *both* HDR

and LDR display modes produced similar ratings to the real scenes (data used from a previous experiment). Univariate analyses between display modes for each rating scale show small, but significant effects on ratings of dim-bright, unform-non-uniform, radiant-gloomy, and realistic-unrealistic. Larger effects were found on ratings of pleasant-unpleasant, comfortable-uncomfortable, tense-relaxed, dramatic-diffuse, glaring-not glaring, friendly-hostile, simple-complex. HDR displays produced mean ratings (on the bi-polar scale from 0 to 100) which placed them closer to the following adjectives: bright, non-uniform, unpleasant, uncomfortable, radiant, tense, dramatic, glaring, hostile, complex, and unrealistic. LDR displays produced mean ratings which placed them closer to these adjectives: dim, uniform, pleasant, comfortable, gloomy, relaxed, diffuse, not-glaring, friendly, simple, and realistic. In terms of environmental satisfaction and features, subjects rated the LDR display mode as slightly higher than the HDR display mode. They could not, based on their data, support or refute the hypothesis that HDR images were perceived as more similar to the real scenes, but suggest that future work should compare real spaces to images of those spaces from similar view positions.

In another study by Newsham, et al. (2010), subjects were shown a selection of 6 daylight spaces in real and then again in either HDR and LDR mode on a HDR display. After exploring each scene or image, the subjects were asked to rate it using semantic differential scales (dim – bright, nonuniform – uniform, unpleasant – pleasant, glaring – not-glaring). For the digital images, subjects were shown photographs of the real scenes in both HDR and LDR display mode (with the order presented in random) and then asked to select which mode is more realistic for each scene. The authors found a significant effect of screen display (HDR or LDR mode) on ratings of dim – right, nonuniform – uniform, unpleasant – pleasant. There was also a significant effect between HDR display mode and the real space for ratings of dim – bright, and glaring – not glaring. Conventional display mode vs. real space was not presented. The results across the 6 selected spaces show that 2 of the 6 show significant differences in ratings between HDR and LDR mode and that these spaces had the largest areas of high luminance. From their analysis, there was a significant difference in ratings between high luminance scenes (the lobby and staircase) shown in HDR display mode and conventional LDR display mode, but the authors also noted a significant effect of display sequence (real or digital image first).

A few key questions arise regarding the design of experiment in this study. Was there a significant effect of HDR vs. LDR display sequence? While it was not mentioned in this study, we wonder if the authors would have found a significant difference between HDR and LDR mode if digital images had not been presented to the same subjects in the same session. When subjects are shown a random presentation of images that jump from cd $\leq$ 200 cd/m2 up to cd $\leq$ 4,000 cd/m2 in back-to-back sequence, the LDR images may appear *relatively* lower in brightness, glare, uniformity, and pleasantness due simply to the obvious shift in brightness capacity of the screen between display modes. The question is, would there be an effect of HDR or LDR mode on subject ratings (as compared to the real space) if they saw all digital images in only LDR *or* HDR mode?

It can be difficult to compare the effects of display device and mode on subjective assessments due to many technical factors (display luminance range, tone mapping operator, detail in screen capture or rendering technique), but there is also disagreement based on statistical methods for determining the effects of that comparison. In a study by Salters, et al. (2012), ratings from 10 semantic differential scales in real world scenes were compared to 4 HDR renderings displayed on a conventional LDR display. Each of the 4 HDR renderings was made using 3D studio Max or Light Tools and was tone-mapped using icam or Reinhard. The analysis of data from this study covers a comprehensive spread of statistical techniques, resulting in an important question: How do we evaluate which rendering/tone-mapping combination best matches the real world scene when statistical indicators (Euclidean Distance, Means Test, KL Divergence, Correlation, ANOVA) give us different assessments of goodness? Real world ratings varied from those made using virtually displayed renderings, but ratings from each of the 4 rendering/tone-mapping combinations also varied from one another. The authors concluded from the range of available tests for comparing the 'badness' of fit, that the combination of 3D studio Max and Reinhard was the best combination in that it was the 'least bad' or equal to the other options in all 5 statistical tests.

A study by Engelke, et al. (2013) compared subjective attribute ratings from a full scale test room, where a number of lighting stimuli were presented, to a collection of HDR photographs and renderings of that same test room, this time presented on a 2D screen ($\leq$ 400 cd/m2). Subjects used a 7-point scale from low to high to rate attributes for pleasantness, brightness, diffuseness, contrast, uniformity, shadow visibility, coziness, liveliness, tenseness, and detachment. Initial results showed no significant difference in responses between the HDR photographs and the real scene, but did show a significant difference between the renderings and the real scene. A second experiment was then conducted using renderings with improved materials and geometry, and this time the authors found that the more 'realistic' renderings achieved a similar result as the HDR photographs. This finding, in line with Salter, et al. (2012) supports the use of renderings, when modelled and tone-mapped accurately, as a surrogate for real scenes in the study of luminous characteristics.

2.3.2 3D and/or Stereoscopic Displays

In a study by Villa & Labayarde (2010a), a comparison was made between real world scenes and printed digital renderings of those scenes. Geometry models of three spaces (a classroom, bathroom, and corridor) were created in 3dsmax2009 and rendered using 4 software packages (Mental-Ray, Maxwell, V-Ray, and Inspirer). Subjects were taken in front of the three scenes in random order and then shown the printed images and asked to place the image on a scale from "not at all" to "extremely" for two questions regarding the reality and lighting atmosphere. V-Ray was selected as the best for producing photorealism and correct lighting atmosphere.

In (2013), a second paper by Villa & Labrayarde presented two experimental studies to evaluate the reliability of an online protocol for assessing the luminous environment using digital renderings.

In the first experiment, subjects saw a series of rendered scenes with manipulated lighting stimuli on two controlled displays: a sterescopic 3d virtual reality simulator (cd≤ 170 cd/m2) and a web-based application. They were asked to rate the suitability of the lighting for work using a sliding scale from 0 - 10 (not at all to extremely). The authors found no statistically significant difference between responses from the 3d virtual reality simulator and the web-based application. A second experiment used the same online protocol and design, but collected 786 subject responses from a broad range of subjects under uncontrolled settings. While some significant effects were observed in operating systems, browser, and brightness of surrounding area, the authors found that 40 subjects provided a large enough sample size to reduce the error caused by these uncontrolled settings.

In a pair of experiments on the influence of presentation modes, Cauwerts & Bodart (2011) and Cauwerts, et al. (2013) compared subjective ratings in real daylit environments to projected photographs of those same scenes in 2D and 3D projection, 2D panoramic QTVR, and 2D images on LDR and HDR displays. The 3D projection was stereoscopic (one image per eye with 3D glasses), while the 2D panorama offered greater view direction range. This comparison demonstrated that only the 2D panorama projection mode, where the user could explore the environment, was able to replicate the evaluation of perceived pleasantness and light distribution of a real space. From their findings, the authors recommend the use of panoramic displays for research on daylight distribution, conventioanl 2D displays for the perception of light level, coloration, and contrast, and HDR screens for the evalution of glare risk. This supports the need for immersive displays in the evaluation of perceptual impressions of daylight composition.

The importance of immersion and interactivity of presentation modes has been highlighted by various researchers, such Bishop and Rohrmann (2003). Following this principle, the use of virtual reality headsets, is suggested by Kuliga et al. (2015) as an empirical research tool that allows a more immersive experience due to the lack of conflicting stimuli in the observer's peripheral vision. To the authors' knowledge, very few studies have used an immersive virtual reality display in the investigation of impressions of lighting. Heydarian et al. (2015a) demonstrated that user performance in object identification, reading speed and comprehension in an office space was similar in a real office environment and its virtual counterpart projected in an immersive virtual reality headset. In a later experiment (2015b), Heydarian et al. explored the lighting preferences of users in a virtual scene through the users' control of the blinds and artificial lights in the virtual environment, using the virtual reality headset Oculus Rift DK2.

A study by Chamilothori et al. (2016), used the same device to investigate the influence of façade patterns on the perceptual impressions of a simulated daylit space. In this study, the authors used a the Oculus Rift DK2 to display a series of immersive renderings of a side-lit demonstration room which showed a range of exterior facade conditions. Subjects were allowed to freely explore the renderings from a fixed view position and then asked a series of questions about their appraisal of the daylight using a 5-point likert scale (interest, excitement, contrast, and satisfaction with view). While preliminary in their findings due to the limited number of subjects, this study revealed a

significant increase in ratings of interest and excitement between exterior screen conditions, with the most *compositionally diverse* composition achieving the most *positive appraisals*.

Recent work by Chamilothori et al. (2017) compares the subjective evaluations of a real space and HDR photographs of that space under different lighting conditions projected in the Oculus Rift CV1 headset. The results of this study are promising for the adequacy of the studied methodology and device as a substitute for real environments in the investigation of perceptual qualities of daylit spaces.

2.3.3 Tone-Mapped Images as a Surrogate for Real Space

When using images to gather subjective impressions of daylight relating to perceptual factors such as brightness and contrast, it is important that light levels are accurately captured, through HDR imaging techniques, and displayed through screen or projection equipment that can reproduce a broad range of luminance values. The impacts of tone-mapping on images (where a compressed luminance range due to display specifications is necessary) are linked to more accurate subjective impressions of light. To create accurate and detailed luminance maps of interior space, both photographic and simulation-based methods can be utilized. With a conventional camera, multiple photographs can be captured across a range of exposures, combined, and calibrated using software which extracts per-pixel luminance values to create a high-dynamic-range HDR image of the space (Inanici, 2003). Using ray-tracing software like Radiance (Ward, 1994), HDR images can also be produced through simulation, a method which offers many advantages when conducting research on temporal characteristics of light. Due to the abundance of tone-mapping algorithms that have been developed over the last two decades, a number of studies have sought to evaluate which algorithms provide the most accurate compression of luminance values for perceptual accuracy. The results of these studies depend on the display medium, from LDR screens (Yoshida, et al., 2005; Cadik, et al., 2008;), to printed photographs (Villa & Labrayade, 2010b), to a virtual reality headset (Chamilothori, et al., 2017). Table 2.3 summarizes the image capture technique, tone-mapping operator (TMO), screen display, and evaluation method for each of these studies.

The study by Yoshida, et al. (2005) asked subjects to rate real-world scenes and HDR photographs of those scenes on a LDR monitor. Several existing tone-mapping algorithms were used to compress luminance values to the range of that LDR screen, including local operators such as Reinhard, et al. (2002), Ashikhmin (2002), and Durand & Dorsey (2002), and global operators such as *pcond* by Ward, et al. (1997) and Drago, et al. (2003). The authors found that global operators resulted in higher brightness and contrast than local operators, but that more detail in bright regions of the images were perceived using local operators like Reinhard. That being said, *pcond*, Reinhard and Drago were all perceived as the most natural looking tone-mappers.

Cadik, et al. (2008) used to range of local and global operators to make direct comparisons between tone-mapped HDR photographs and real-world scenes as well as ratings of HDR photographs

without a real-world reference. In these studies, the authors' found that the best overall quality was generally observed by global operators such as *pcond* by Ward (1997). In the top six best performing TMOs in this study, Reinhard, et al. (2002) was the only local operator.

In a (2010b) study by Villa & Labrayade, printed photographs were used to compare tone-mapping algorithms to 5 real-world scenes. In this study, the authors considered a range of local and global operators, including Reinhard, et al. (2002), Drago, et al. (2003), Durand & Dorsey (2002), and Ashikhim (2002). They found that Reinhard, et al. (2002) and Drago, et al. (2003) received the best average scores from a series of ratings related to luminosity, contrast, detail, and artefacts, with Ashikhim (2002) performing the worst.

A study by Chamilothori, et al. (2017), used pcond by Ward (1997), Durand & Dorsey (2002), and Reinhard, et al. (2002) to tone-map HDR photographs to the luminance rang of a virtual-reality headset. Subjects were presented a real world-scene and tone-mapped photograph in random order and then asked to rate the scenes using for series of attributes. This preliminary study found no significant difference in ratings between the available tone-mapping algorithms.

In conclusion, there are a number of tone-mapping operators available for compressing HDR images down to the display luminances of a given printed or projected medium. While some operators seem to consistently out-perform others (Reinhard, et al, 2002 and Ward, 2997), most of these studies have been conducted using 2D image content and more work is needed to validate the findings in Chamilothori, et al. (2017).

Table 2.3 - *Experiments on Tone-Mapping Algorithms*

Tone-Mapping		*Year:*	*Image Capture:*	*TMO:*	*Scene Display:*	*Evaluation Method:*
27.	Yoshida et al.	2005	HDR Photographs	pcond, Pattanaik, Drago, Reinhard, Ashikhim, Durand	Real space vs. LDR screen	Subjects compared real world to HDR using ratings for appearance & realism
28.	Cadik et al.	2008	HDR Photographs	pcond, Drago, Durand, Reinhard, & more...	Real space vs. LDR screen $\leq$ 280 cd/m2	Subjects rated image attributes both with and without real space comparison
29.	Villa & Labrayade	2010b	HDR Photographs	Ashikhim, Drago, Durand, Fattal, Reinhard (02 & 04)	Real space vs. Printed Photographs	Subjects rated images in comparison to the real space using a drawn scale
30.	Chamilothori et al.	2017	HDR Photographs	pcond, Durand, Reinhard	Read Space vs. virtual image using Oculur Rift CV1	Semantic differential ratings & multiple choice

2.4 Measuring Contrast in Digital Images

The following section will introduce two categories of algorithms that were developed to measure contrast perception in digital images: those that rely on global measures and those that rely on local measures. The first category describes 'global measures,' which rely on non-spatially dependent methods of evaluating individual pixels, such as a simple ratio between high and low pixel values. The second category describes a range of 'local measures,' whereby contrast is computed as the difference between localized or neighboring pixels in a compositionally-dependent framework.

2.4.1 Global Measures

The problem with those experimental studies presented in the previous sections, which rely on simple photometric measures such as average luminance and luminance variation to draw connections with various subjective attributes, is that they generally do not address the spatial diversity of luminance values within an occupants' field-of-view. In these studies, luminance variation or contrast is most commonly defined by a global measurement, such as Michelson or Root Mean Square (RMS) contrast. The Michelson contrast (Michelson, 1927) is defined as,

$$Michelson = \frac{P_{max} - P_{min}}{P_{max} + P_{min}}, \tag{2.1}$$

where pixel intensities p_{max} and p_{min} are two single points of extreme brightness and darkness in an image. The RMS contrast measures the root mean squared difference of individual pixel intensities from the mean (Pavel, et al., 1987)

$$RMS = \sqrt{\frac{1}{WH}\sum_{i=1}^{W}\sum_{j=1}^{H}\left(p_{i,j} - \bar{p}\right)^2}, \tag{2.2}$$

where $p_{i,j}$ are the pixels intensities at position *(i,j)* in an image of size *W* by *H* and $\bar{p}$ is the average pixel intensity.

These global contrast metrics provide a single comprehensible value, which existing studies in daylight perception have used for straight forward comparisons with subjective rankings (Wymelenberg & Inanici, 2010). Yet, they cannot effectively predict perceived contrast between two images that vary in composition (Simone, et al., 2012). To overcome this limitation, more sophisticated contrast metrics have been developed in the fields of image analysis and vision research.

2.4.2 Local Measures

Local contrast metrics were developed to overcome the limitations associated with global metrics by quantifying the effect of composition, i.e. contrasting areas of brightness and darkness. Included within this group of metrics are methods that evaluate spatial frequencies in the Fourier domain (Hess, et al., 1983), those that measure a weighted color contrast based on the distance between chroma regions (Tremeau, 2000), and those that calculate the difference between a single pixel and a surrounding region or neighborhood (Tadmor & Tolhurst, 2000) (Rizzi, et al., 2004) (Matekovic, et al., 2005) (Rockcastle & Andersen, 2014). The authors have focused on the latter group of neighborhood metrics for their ability to quantify the local contrast values between pixels within a neighborhood or sub-region and assign a singular measure which represents the strength of local variation across all pixels.

RAMMG, a multi-level algorithm developed by (Rizzi, et al., 2004) measures the local variations in brightness across an image by computing the mean local pixel variation across a pyramidally sub-sampled structure, taking into account perceived differences in brightness across multiple image resolutions. The overall measure for N number of levels is described as

$$RAMMG = \frac{1}{N}\sum_{l=1}^{N} \overline{c_l} \quad , \tag{2.3}$$

where $\overline{c_l}$ is the mean contrast in the level l:

$$\overline{c_l} = \frac{1}{WH}\sum_{i=1}^{W_l}\sum_{J=1}^{H_l} c_{i,j} \quad . \tag{2.4}$$

The image resolution is halved in each subsequent level, where $W_l = W_{l-1}/2$ and $H_l = H_{l-1}/2$ are the width and height of the image at level l and $c_{i,j}$ is the contrast of each pixel, calculated as:

$$c_{i,j} = \sum_{k \in K_8} \alpha \left| p_{i,j} - p_k \right| , \tag{2.5}$$

where pixels p_k are the 8 neighbouring pixels of $p_{i,j}$ and the weight α applied to each of the 8 surrounding pixels k is

$$\alpha = \frac{1}{4+2\sqrt{2}} \begin{vmatrix} \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \\ 1 & & 1 \\ \frac{\sqrt{2}}{2} & 1 & \frac{\sqrt{2}}{2} \end{vmatrix} . \tag{2.6}$$

Multi-level metrics like RAMMG were developed under the assumption that a multi-resolution analysis can better explain different mechanisms of perception in the human visual system

(Adelson, et al., 1984) (Simone, et al., 2009). Whereas large image resolutions (>100,000 pixels) provide the detail to compute small, localized contrast evaluated between pixel neighbors, small image resolutions (<25,000) provide the opportunity to measure the difference between larger areas of brightness (i.e. larger neighbourhoods). While a pixel does not directly translate to the solid view angle between a subject and 3-dimensioanl scene due to the flattening of pixels in a 2D image, larger pixels may be seen as representing a larger view angle to the scene. Large image resolutions may therefore correspond to higher-level vision function such as object and pattern recognition, while small image resolutions may correspond to lower-level functions such as the extraction of light stimulation to construct surfaces, edges, and features. We believe that some resolution levels may relate to contrast-driven subjective responses more than others, which is also suggested by Rizzi et al. based on their findings in (Rizzi, et al., 2008) and discussed in depth in Chapter 5.

The Difference of Gaussian (DOG) metric, developed by (Tadmor & Tolhurst, 2000), differs substantially from the previous three contrast metrics in that it computes local differences between two bi-dimensional Gaussian filters with a center component $R_c(x,y)$ and a surround component $R_s(x,y)$

$$DOG(x,y) = \frac{R_c(x,y) - R_s(x,y)}{R_c(x,y) + R_s(x,y)}. \tag{2.7}$$

The center component $R_c(x,y)$ can be described as:

$$R_c(x,y) = \sum_{i=x-3r_c}^{x+3r_c} \sum_{j=y-3r_c}^{y+3r_c} G_{center}(i-x, j-y) p_{i,j} \quad , \tag{2.8}$$

with a bi-dimensional Gaussian filter G_{center} with center (x,y) defined as:

$$G_{center}(x,y) = e^{-(x^2+y^2)/{r_c}^2} \quad , \tag{2.9}$$

where $p_{i,j}$ are the pixel intensities at position (i,j) and r_c is the radius of the Gaussian filter $G_{center}(x,y)$.

The surround component $R_s(x,y)$ in Equation 2.9 can be described as a bi-dimensional Gaussian filter:

$$R_s(x,y) = \sum_{i=x-3r_s}^{x+3r_s} \sum_{j=y-3r_s}^{y+3r_s} G_{surround}(i-x, j-y) p_{i,j} \ , \tag{2.10}$$

with a Gaussian filter with surround (x,y) and a radius r_s defined as:

$$G_{surround}(x,y) = 0.85 \left(\frac{r_c}{r_s}\right)^2 e^{-(x^2+y^2)/{r_s}^2} \ , \tag{2.11}$$

where the $0.85\left(\frac{r_c}{r_s}\right)^2$ sets the integrated sensitivity (or volume) of the surround component to be 85% of that of the center.

In the original metric, the final calculation was computed as an average of *DOG(x,y)* taken across 1,000 randomly selected pixels within an image to save computational time. In this thesis, implementations of DOG are computed using the average *DOG(x,y)* across all pixels in a given image with width W and height H:

$$\overline{DOG} = \sum_{i=1}^{W}\sum_{j=1}^{H} DOG(x_i, y_j). \tag{2.12}$$

In 2009, Simone et al. combined the multi-level approach developed for RAMMG and the DOG metric to create a multi-level metric called RSC (Simone, et al., 2009)

$$RSC = \frac{1}{N}\sum_{l=1}^{N} \overline{DOG_l}\ , \tag{2.13}$$

where *N* is the number of levels and $\overline{DOG_l}$ is the mean contrast in each level *l*

$$\overline{DOG_l} = \sum_{i=1}^{W_l}\sum_{j=1}^{H_l} DOG(x_i, y_j). \tag{2.14}$$

Each of these metrics will be applied to a range of 2D and 3D images to compare their computed value to subjective ratings of daylight composition, gathered through a series of studies presented in Chapters 3-5 & 8.

2.5 Summary of the Current State

Task-driven illumination metrics such as DF and DA can be used to determine whether an interior space is sufficiently illuminated for the performance of visual tasks whereas comfort-based luminance metrics such as DGP allow us to evaluate the visual field for sources of glare-based discomfort. While the shift toward climate-based metrics such as DA, cDA, DSP, UDI, AIE or sDA represents a significant improvement in daylight analysis, this data is limited to a two-dimensional task-surface and does not correspond to the three-dimensional view of space that is perceived by an occupant. Although dynamic glare metrics such as DGPs evaluate a three-dimensional view position across the day and year and novel approaches to view and gaze-response combine dynamic occupant behaviour, they were designed to predict the negative impacts of contrasting luminance levels on visual comfort. Of the many established glare indices, not one addresses the notion of contrast as a positive visual effect, although there is some mention of luminance-based attraction in the work of (Sarey Khanie, et al., 2015).

Furthermore, task-driven illumination and visual comfort metrics are only applicable in spaces where visual tasks are frequently encountered. For spaces where visual tasks are less indicative of lighting performance, we have few, if any, broadly accepted metrics to help guide designers toward improved performance. In the absence of quantitative criteria, architects are tasked with creating acceptably bright or visually engaging environments, based on subjective (Cuttle, 2010). For many architects, this task is made difficult by the dynamic nature of sunlight and the challenges associated with predicting the range of visual effects that will occur across the day and year.

There are a number of methods for creating and displaying interior scenes to assess subjective qualities of daylight, each of which has its own set of advantages and limitations. While real spaces obviously produce the most accurate impression of light for subjective assessment, experimental conditions are limited by the physical sky conditions available on site and it is difficult to compare a range of spatial configurations or temporal conditions in an efficient manner. Architectural models are more affective at comparing various spatial configurations, but are limited by many of the same conditions as real spaces. Image projection, especially 3D set-ups, can display simulated renderings from a human point-of-view, allowing for a broad range of spatial, temporal, and artificial lighting conditions, but these experiments are limited by the luminance output of the projector. Backlit LCD displays produce the most realistic range of luminance values for subjective assessment, but are arguably limited by the size and viewing position of the screen. All in all, it is necessary to select a display device and presentation mode that is most suited for the specific experimental parameters.

As was discussed in section 2.2, existing studies that address daylight perception within architecture have focused on either physical attributes (like wall/window ratio) or global measures such as standard deviation, luminance range, and/or average luminance due to the ease of comparing single global measures against subjective rankings (Wymelenberg & Inanici, 2010). Given a lack of consensus in the use of global measures that accurately assess contrast perception in lighting design, contrast is often regarded as a qualitative element of daylight performance. As such, attributes related to lighting composition are generally considered to be subjective and are evaluated by designers using illustrative renderings or photographs that capture the luminous character of a space.

Although we can view an architectural rendering and describe the location and distribution of contrasting values and compositional dynamics within it, there are few methods that successfully quantify the spatial diversity of this distribution and link it to subjective appraisals or evaluations. Such a method could help designers to contextualize the relative strength of perceived contrast composition within an architectural space, use it to compare daylight-driven visual effects, and visualize how and when these effects vary over time. While the LD index proposes a method for analysing the spatial diversity of luminance values across an occupant's point-of-view, it does not address the temporal impacts of these visual effects (Papairi, 2002). Furthermore, the method relies on physical measurements in live space, which can pose a number of practical problems, such as the movement of people and the disruption of equipment over extended studies.

A simulation-based measure which relies on renderings would allow designers to assess design proposals before they are constructed and make iterative improvements to meet a desired set of dynamic ambiances and emotional responses.

Existing research has produced promising indications that both luminance distribution and diversity play an important role in an occupant's perception of and preference toward the luminous environment. We have also seen that multiple display modes, including tone-mapped 2D images on LDR screens($\leq$400 cd/m2) , HDR screens, wide-format images displayed on panoramic displays, and 3D images displayed using 3D projection, stereoscopic, and virtual reality headsets have all been evaluated as potential surrogates to real spaces in the study of perceptually accurate daylight attributes. Engelke, et al. (2013), Cauwerts & Bodart (2011, 2013) and Villa & Labrayarde (2013) found that LDR screens could serve as reasonable surrogates for real world settings when using 2D renderings or tone-mapped photographs. While, Newsham, et al. (2010) found some significant differences between attribute ratings using LDR and HDR display modes, Newsham, et al. (2002) found that ratings from neither the HDR *or* LDR display modes provided a significant difference to ratings from the real world scene. Furthermore, subjects rated the environmental satisfaction and features as slightly higher in the LDR display mode. To evaluate characteristics related to lighting distribution, Cauwerts & Bodart (2011, 2013) recommended the use of a visually-immersive display mode such as a panoramic monitor. Chamilothori, et al (2016) further corroborated this recommendation through the use of a virtual reality headset.

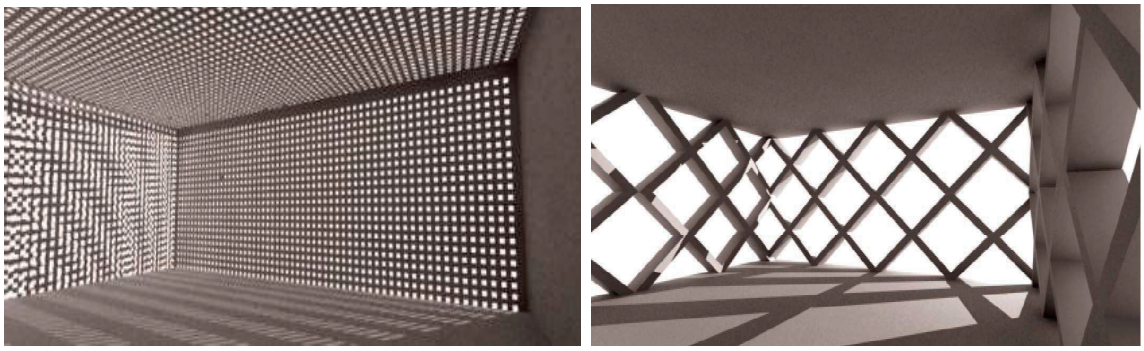
Using both 2D and 3D display modes, the author of this thesis will seek to evaluate the impacts of architectural composition, sky condition, and daylight distribution on subjective ratings of attributes related to visual interest. Using the data from these experiments, a range of image-based algorithms will then be compared to subjective data in search of a new, immersive method of predicting those attributes.

3 EVALUATING DYNAMIC CONTRAST IN DAYLIT RENDERINGS: A PROOF-OF-CONCEPT

This chapter introduces a simulation-based approach for measuring compositional contrast and its temporal variability in renderings of simple architectural spaces. Building upon the work that was completed by Rockcastle (2011), this chapter synthesizes the development of an image-based algorithm and its application to a time series of renderings. Starting with an introduction on algorithms developed to measure contrast in digital images, this chapter will introduce the application of an adapted algorithm, first referred to as *Spatial Contrast* (SC) by the authors in Rockcastle & Andersen (2011) and adapted in Rockcastle & Andersen (2014), to interior renderings across an annual time series to see how the composition of contrast varies locally across a scene with shifting sun positions. As both contrast and its temporal variability are believed to impact perceptual impressions in daylit architecture, this application of a local contrast algorithm to digital renderings illustrates the *potential* for a simulation-based approach in predicting the strength and variation of visual effects *over time*. Once we establish the proof-of-concept for a local contrast-based algorithm to differentiate between interior renderings and moments in a time series, Chapter 4 will then compare a range of contrast-based algorithms and further develop the mathematical approach proposed by the SC algorithm.

3.1 Evaluating Contrast in Digital Images

Local contrast measures were developed to overcome the limitations associated with global measures by quantifying the effect of composition on contrasting areas of brightness and darkness. Chapter 2 introduced a range of algorithms for measuring contrast, including methods that measure spatial frequencies in the Fourier domain (Hess, et al., 1983), those that measure a weighted color contrast based on the distance between chroma regions (Tremeau, 2000), and those that calculate the difference between a single pixel and a surrounding region or neighborhood (Tadmor & Tolhurst, 2000; Rizzi, et al., 2004). Through this range of methods, there is disagreement on how to produce a single number that represents the contrast perception of an entire image when contrast can be localized to a neighbourhood of values (Simone et al., 2012). On the one hand, a single number cannot easily distinguish between two images that vary in composition. This is due to the loss of information that occurs between a matrix of values or a power spectrum (resulting from a Fourier transformation) and the final value, which is often an average of a threshold. On the other hand, a single number is a compact measure that can be compared to subjective experiments, which often produce a single value from occupant surveys.



a) avg. brightness: 120, standard deviation: 18 **b)** avg. brightness: 132, standard deviation: 22

Figure 3.1 Two abstract renderings that differ in composition, but share similar values for mean pixel brightness and standard deviation.

For example, if we compare two renderings of daylit spaces side-by-side (figure 3.1) and relied on existing methods of global contrast analysis, such as average brightness or standard deviation, we would not be able to differentiate between the two images despite their varied compositions. The image on the left (figure 3.1a) shows a dense pattern of light and shadow - with an average brightness of 120 (out of 256 distinct RGB channels) and a standard deviation of 18. The image on the right (figure 3.1b) contains larger patches of sunlight, but has a similar average brightness of 132 and standard deviation of 22. The obvious drawback to measurements like average brightness and standard deviation, is the loss of information that occurs when individual values are removed from the spatial surround of their composition.

As was discussed in chapter 2 existing studies that address contrast perception within *daylit architecture* have focused on global measures such as standard deviation, luminance range, and average luminance (or brightness), due to the ease of comparing single global measures against subjective rankings (Wymelenberg & Inanici, 2009; Demers, 2007). Given the lack of consensus over the use of global measures that accurately assess contrast perception in lighting design, contrast is often regarded as a qualitative element of daylight performance and is subjectively evaluated by designers through renderings or photographs that capture the luminous character of space. Although we can view an architectural rendering and *describe* the location and distribution of contrasting values within it, there are few methods that successfully quantify the spatial diversity of this lighting distribution. Such a method could help designers to contextualize the relative strength of perceived contrast within an architectural space, use it to compare daylight-driven visual effects, and visualize how and when these effects vary over time.

3.2 Spatial Contrast

In this chapter, Spatial Contrast (SC) will be used to describe the sum of local variations in brightness across a digital image, or - more specifically - the sum of variation between neighboring

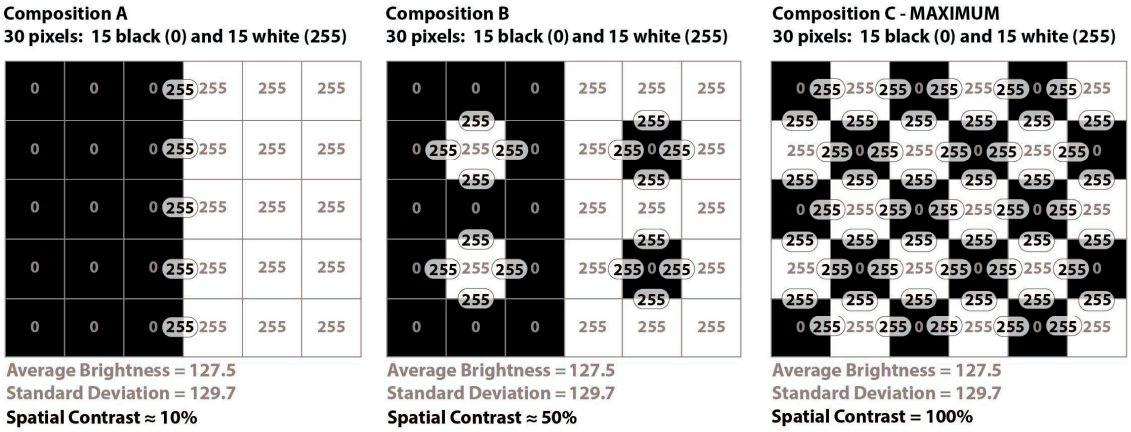


Figure 3.2 A demonstration of spatial contrast in three compositions with the same number of black and white pixels, the same RGB average and standard deviation.

pixels within a rendering or photograph of daylit space. Spatial Contrast resembles an existing method of local contrast measure proposed by Rizzi, et al. (2004) which computes the difference between every pixel and its neighboring 8 pixels across various frequency levels (Equations 2.3-2.6). Due to computational feasibility in this proof-of-concept, SC is simplified to one grey-scale color channel, single pixel resolution, and only the four immediate pixel neighbors. While luminance-based analysis methods often rely on HDR images due to their expanded range of pixel values, it was important, in this phase of work, that any image-based contrast-analysis algorithm could accommodate 8-bit images. This allows for flexibility in utilizing a broad range of image-generation techniques such as point-and-shoot photography and renderings from a range of geometric modeling platforms that may not accommodate HDR imaging. As such, renderings in this chapter used a standard grey-scale pixel value range of RGB 0 (black) to RGB 255 (white), tone mapped down from the HDR luminance range produced in Radiance and tone-mapped using the *pcond* tone-mapping operator (Ward, et al., 1997). As the thesis moves from a purely simulation-based approach to an experimental validation phase in Chapters 5 and 8, the role of tone-mapping and display medium will be discussed in more depth as it is essential to map luminance values relative to the printed or projected range that subjects will see. For now, all computed values are relative, as no subjective data is collected in this chapter.

Figure 3.2 illustrates the problem associated with global measures like standard deviation. When composition A is split down the middle, with half the pixels representing RGB 0 (black) and the other half representing RGB 255 (white), the average pixel brightness is 127.5, with a standard deviation of 129.7. Both average brightness and standard deviation are compact global measures used to describe luminance distribution in existing daylight research (Wymelenberg & Inanici, 2009; Demers, 2007; Cetegen, et al., 2008; Tiller & Veitch, 1995). As we rearrange the composition to create more perimeter between white and black pixels, such as can be seen in compositions B and C, the average pixel value and standard deviation remain unchanged, despite the obvious shift in our perception of localized variations between neighbouring pixels.

To account for these localized differences in brightness, which were thought to contribute to human perceptions of contrast, we introduce an initial algorithm called Spatial Contrast (SC) that calculates the difference in brightness between each pixel and its four neighbouring pixels. This algorithm is a simplification of that developed by Rizzi et al. (2004), which takes an average difference in brightness between each pixel and its eight neighbouring pixels for a pyramid of sub-sampled resolutions (Section 1.4.2). Instead of an average difference, as used in the approach by Rizzi et al. (2004), Spatial Contrast computes the *sum* of local contrast values, as it was initially thought that the sum of local contrast values more accurately accounts for the compositional complexity of contrast across an image. In Chapters 4 & 5, we will look more closely at this assumption and compare the SC algorithm to variations of this approach which use the average neighbourhood difference across a single native resolution, a single resolution within a sub-sampled series of levels, or an average of sub-sampled levels, as Rizzi proposed. As the cumulative sum can vary greatly between images that range in size, the overall measure for spatial contrast is expressed as a ratio between the cumulative sum of local contrast values for a given image size (n x m) and a maximum hypothetical value based on the image size (n x m).

If we have an image matrix *P* of size (n x m), each element *p* contains a brightness value for a relevant pixel in the image composition. To calculate spatial contrast, the difference in brightness is taken between each pixel *p* and its four neighboring pixels (column and row - figure 3.3) and then averaged to produce a local contrast matrix of size (n x m):

$$\overline{\Delta p_{i,j}} = \frac{1}{4}\left(\left|p_{i,j} - p_{i+1,j}\right| + \left|p_{i,j} - p_{i-1,j}\right| + \left|p_{i,j} - p_{i,j+1}\right| + \left|p_{i,j} - p_{i,j-1}\right|\right) \tag{3.1}$$

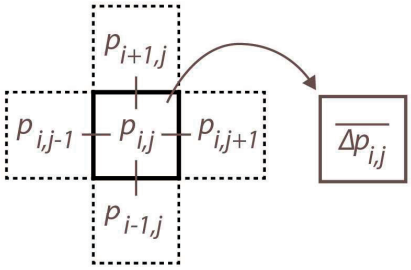


Figure 3.3 The average difference between each pixel $p_{i,j}$ and its four neighboring pixels.

for all i=1→n and j=1→m.
From these average local contrast values, spatial contrast is defined as:

$$Spatial\ Contrast(\%) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\overline{\Delta p_{i,j}}}{\overline{\Delta pmax_{i,j}}} * 100 \tag{3.2}$$

Where $\overline{\Delta pmax_{i,j}}$ is a hypothetical maximum value which is computed as a black and white checkerboard of size (n x m) where every pixel has an average local contrast of 255:

$$\overline{\Delta pmax_{i,j}} = 255 * (n)(m) \tag{3.3}$$

As the hypothetical maximum value used to calculate spatial contrast is far from being possible for any architectural space to achieve (as a symmetrical checkerboard of black and white pixels is two dimensional without spatial depth), the resulting values must be adjusted to a range that is meaningful. Figure 3.4 shows the distribution of normalized values and data quantiles for spatial contrast achieved by the ten case study spaces in this chapter (which will be introduced in section 3.4.1), where the middle quantile is used to scale values in the medium range and the upper and lower quantiles are used to scale high and low values, respectively. While the upper and lower thresholds of this scale would need adjustment based on a broader set of architectural spaces and renderings, the current range allows for an adequate comparison of the selected case studies in relative terms.

When applying SC to the pair of (584 x 564) images introduced in figure 3.1a and b, we can produce the corresponding spatial contrast (583 x 563) matrices illustrated through figure 3.5a and b. As evidenced when compared to one another, global and local contrast measures for each image differ greatly in their ability to differentiate between variations in the spatial distribution and density of contrasting pixel values. Figure 3.5a shows an interior space with many small openings that emit highly articulated patterns of light and shadow, leading to a spatial contrast equal to roughly 90% (after normalization to the range described above). Figure 3.5b, on the other hand, shows a rendering of a dia-grid façade with larger patches of light and shadow and a spatial contrast equal to roughly 60%. The distinguishably stronger spatial contrast in figure 3.5a, when compared to figure 3.5b, corresponds to an intuitive difference in composition between these two examples better than the global measures provided in section 3.1 (average brightness

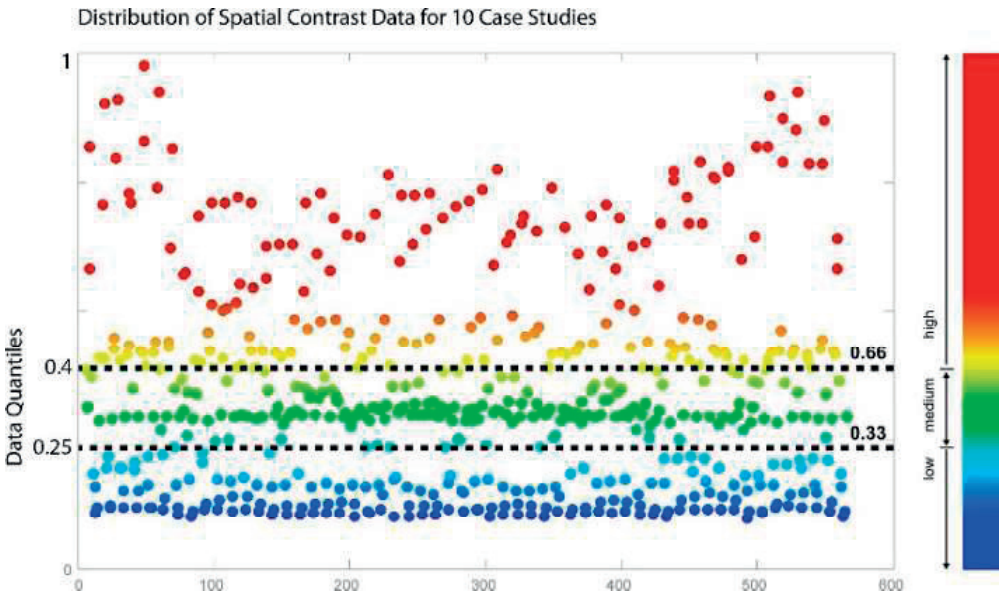
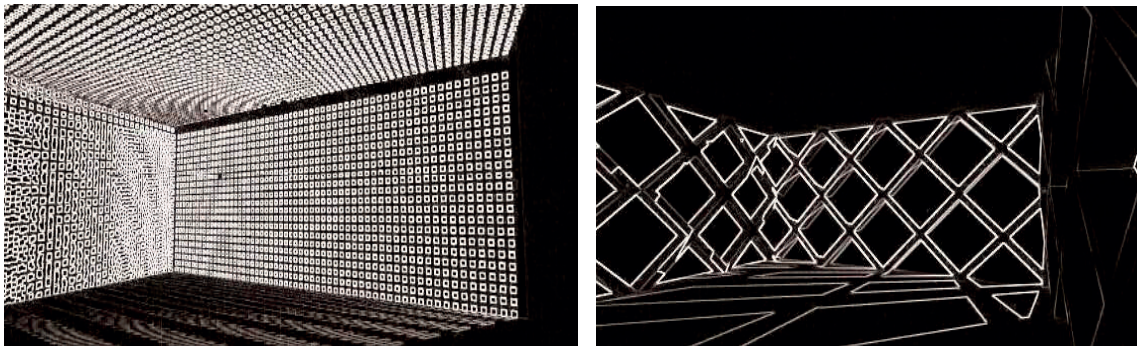


Figure 3.4 All values for spatial contrast (56 time steps x 10 case study spaces = 560 instances) have been normalized against the maximum value achieved through the case studies presented in this chapter.



a) **b)**

Figure 3.5 Spatial Contrast across two case study spaces with a) Spatial Contrast equal to 90% and b) Spatial Contrast equal to 60%.

and standard deviation). This example illustrates the need for a robust local measure, which is sensitive to compositional differences in brightness, to evaluate contrast in daylit scenes which are often spatially diverse. While global metrics like Michelson (Section 2.4.1) were developed to measure contrast between a single object and its background, interior architecture is composed of many objects across a diverse background, making those specific measures unreliable for contrast prediction within a three-dimensional scene. The following section builds upon these concepts by integrating the component of time to see how spatial contrast varies, temporally throughout a selected scene.

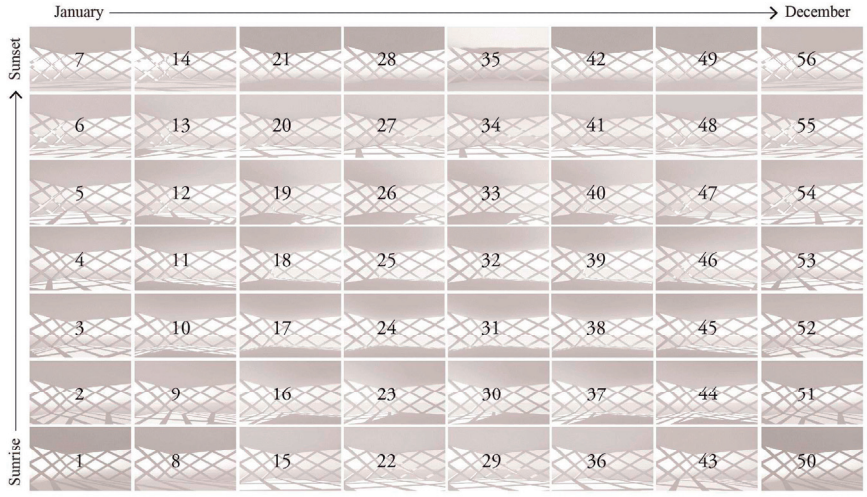
3.2.1 An Annual Approach to Spatial Contrast

In order to understand the impacts of daylight dynamics throughout the year, we applied the spatial contrast measure introduced in section 3.2 to a time series of 56 renderings in order to measure the cumulative effects of spatial contrast over time and visualize when and where these local variations occur throughout the selected scenes. Using time steps to divide the year into a series of symmetrical moments, we can provide a representative cross-section of daylight-driven visual effects. We can then calculate the cumulative value of spatial contrast across these symmetrical annual instances and plot the results across a temporal map to visualize daily and seasonal variations.

As the chosen time-step will influence the outcomes of any performance visualization, it is important to choose a resolution capable of describing climate and variations in sun course for a given location with sufficient detail. Towards this end, a method of time segmentation originally developed for Lightsolve by Kleindienst et al. (2008) will be used. Developed by Andersen and her research group, originally at MIT and now at EPFL, Lightsolve is a simulation platform that combines annual spatio-temporal maps with user-defined goals and associated annual daylight renderings for a climate-based analysis of daylight performance that splits the year into 56 representative time periods. It allows the designer to establish goal-based performance

criteria and navigate the resulting maps of annual data alongside daylight renderings to provide a quantitative and simultaneously visual representation of performance in daylit space (Andersen, et al., 2013b; Kleindienst & Andersen, 2012; Gagne, et al., 2011; Andersen, et al., 2008). Its automated production of climate-driven annual renderings and temporal mapping capabilities make the method an appropriate starting point for new image-based metrics that address perceptual factors of daylight performance on an annual scale.

The present study adopts the 56 time periods proposed by the Lightsolve method using the average sun position for the considered period, but uses a CIE clear sky model for all 56 renderings, instead of a weather-based selection from multiple sky models. This simplification removes the variability of weather to produce a consistent set of time-segmented rendering that can be analyzed for relative changes in spatial contrast under a single sky condition. The resulting set of clear sky renderings can also be considered an upper boundary for spatial contrast as overcast sky conditions produce less direct sun penetration and lower values, a notion that will be explored further in chapter 6-7. It should be noted that these 56 moments have not been validated for use with luminance-based performance predictions and only in climate-based analyses relying on illuminance calculations. Further work is needed to explore the appropriate number of time steps for an annual analysis of this nature. Figure 3.6 shows the 56 dates and legal times (figure 3.6b) for Boston, MA that are used to produce the annual set of renderings (figure 3.6a). The method for producing these renderings, including model geometry, building orientation, materials, and view-port setting will be described in more detail in Section 3.3 with the application of metrics to a series of simple case studies in Section 3.4. The 56 instances listed in Figure 3.6b will be used to produce all annual sets of renderings used for visual analysis in this chapter.



a)

Jan-13	1	07:52	Jul-15	29	06:25
	2	09:12		30	08:33
	3	10:33		31	10:41
	4	11:53		32	12:50
	5	13:13		33	14:58
	6	14:34		34	17:06
	7	15:54		35	19:14
Feb-28	8	07:09	Aug-29	36	07:03
	9	08:45		37	08:57
	10	10:21		38	10:51
	11	11:57		39	12:45
	12	13:33		40	14:39
	13	15:09		41	16:33
	14	16:45		42	18:27
Apr-14	15	07:03	Oct-14	43	07:43
	16	08:57		44	09:18
	17	10:51		45	10:54
	18	12:45		46	12:29
	19	14:39		47	14:04
	20	16:33		48	15:40
	21	18:27		49	17:15
May-30	22	06:15	Nov-28	50	07:30
	23	08:24		51	08:50
	24	10:33		52	10:11
	25	12:42		53	11:32
	26	14:50		54	12:52
	27	16:59		55	14:23
	28	19:08		56	15:33

b)

Figure 3.6 Annual approach: a) renderings for 56 annual moments listed in b) showing dates and times of the instances selected to represent a symmetrical annual cross-section (all times are legal and not solar).

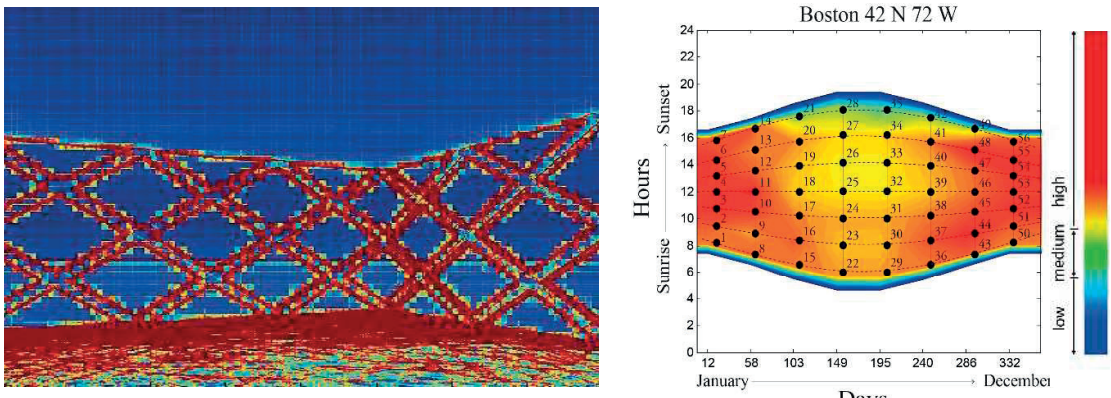


Figure 3.7 Annual spatial contrast results for a top-lit space showing a) cumulative annual map of luminance variability and b) Temporal map of luminance variability.

In order to calculate annual spatial contrast, instantaneous spatial contrast was computed across all 56 moments shown in Figure 3.6a. The cumulative sum of these instances, shown in Figure 3.7a, highlights the location and strength of spatial contrast throughout the year, revealing areas that receive consistently high (red) and low (blue) contrast. In order to visualize when these dynamic effects of spatial contrast vary throughout the year, the sum of spatial contrast for each of the 56 renderings is plotted temporally to express the frequency and magnitude of daily and seasonal variations (Figure 3.7b). The vertical axis of the temporal map shows daylight hours, from sunrise to sunset, while the horizontal axis shows days of the year, from January 1st to December 31st. Values are given a color scale to show relative strength from low to high, which has been adjusted to accommodate appropriate minimum and maximum values as mentioned in section 3.2 and determined by the case studies that will be introduced in section 3.3.1.

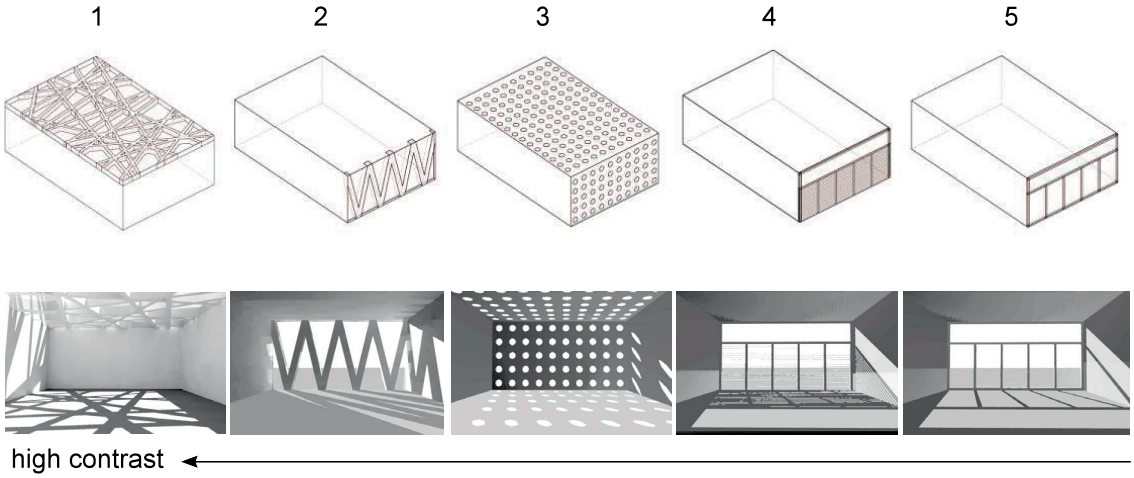


Figure 3.8 Ten case study models from high contrast and variability on the left to low contrast and variability on the right

3.3 Application of Spatial Contrast to Case Study Models

In order to evaluate these metrics across the year and determine whether they can adequately differentiate between a range of daylight conditions, 10 case study spaces were generated to represent an intuitive gradient of spatial contrast and its temporal variability from high to low (Figure 3.8). Each space was modelled with the same generic floor plan and ceiling height dimensions as well as a fixed camera position to produce a comparable set of annual renderings. Each case study was modelled to emulate abstract conditions of existing interior architecture and represent a gradient of luminous effects. Although a more rigorous study of existing architectural spaces is required to validate the range of results for each metric, these 10 initial case studies allow us to establish relative thresholds for 'high' and 'low' spatial contrast as shown in Figure 3.4.

3.3.1 Selection & Modeling of Case Studies

In order to select typological spaces that showcase a sufficiently broad range of daylight-driven visual effects, the authors referred to the survey of contemporary architecture presented in Section 1.1.2 This survey contained 60 architectural spaces from around the world and was organized into 10 categories that describe the strength of spatial contrast and termporal variability intuited by the authors within each interior view. This matrix from Figure 1.8 inspired the slightly-more-compact gradient of 10 case studies presented in Figure 3.8. Development of the typological ordering system, first created to intuitively describe spatial contrast and termpoal variability, is explained in more depth in the introduction, but each case study used in this chapter was designed to resemble abstract characteristics found from within the larger matrix of existing architectural spaces. Figure 3.8 shows the selected case studies in a line from 1 to 10, representing the author's intuitive gradient of visual effects (before application of the metrics) from high spatial contrast and variability on the left to low contrast and variability on the right.

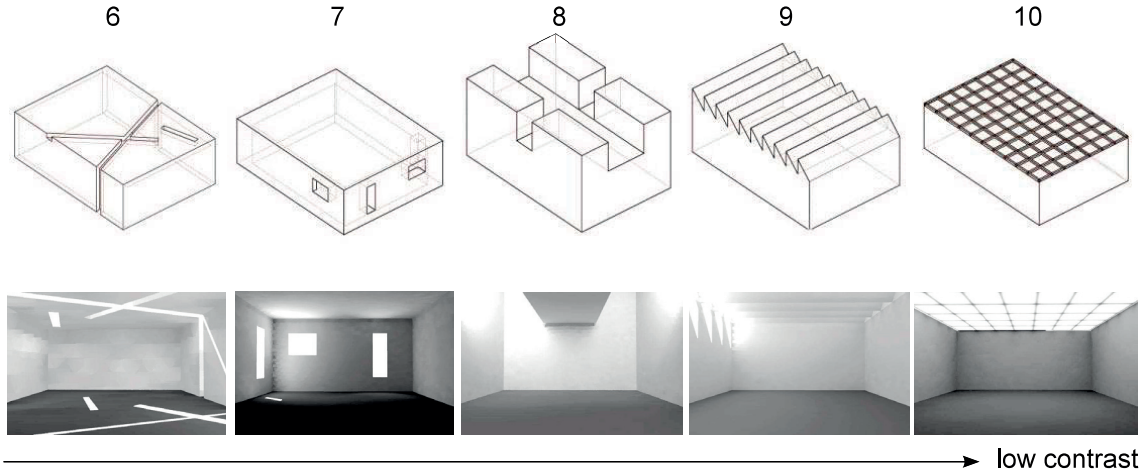


Figure 3.8 continued...

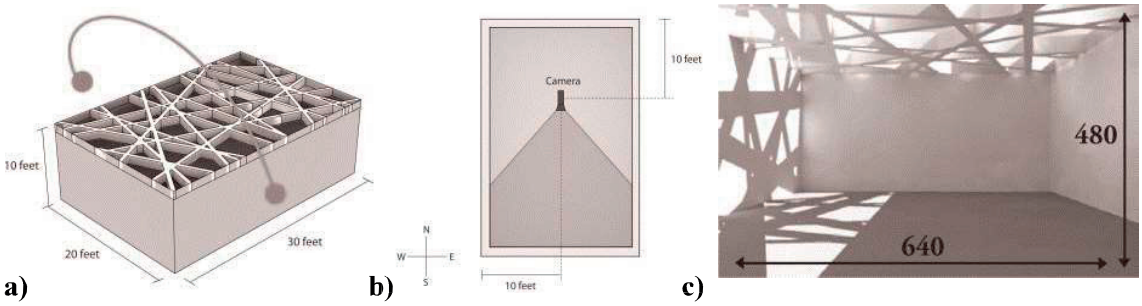


Figure 3.9 Modelling parameters showing a) exterior dimensions for each case study model, b) camera location in plan, and c) sample image dimensions for HDR renderings

The annual spatial contrast metric described in Section 3.2.1 was then applied to this set of ten case-study spaces as a preliminary feasibility study to verify whether the proposed metrics are capable of reproducing a similar gradient to the authors' own intuition. In order to compare these case studies, we produced a set of annual renderings for each space.

Using Rhinoceros (http://www.rhino3d.com, 2007), each case study was modelled with the same floor area, ceiling height, and camera location so that results could be compared in relative terms (figure 3.9a). Cameras were positioned to face South, centered in the East-West direction, and offset ten feet from the back wall to ensure an even distribution of wall, floor, and ceiling surfaces within each view (figure 3.9b). The Diva-for-Rhino toolbar (http://www.diva-for-rhino.com, 2009) was then used to export the camera view to Radiance (http://www.radiance-online.org/, 2009) with a vertical and horizontal viewport ratio set to –vv 40 and –vh 60. The specified materials were set to default reflectance values for floor, wall, and ceiling surfaces (0.3, 0.7, 0.9 respectively). The resolution of each image was rendered at high quality (a DIVA preset) to accommodate adequate detail with a 640 x 480 pixel aspect ratio (figure 3.9c). Although an individual rendering may be produced in Radiance using the DIVA toolbar, these proposed metrics require an automated set of 56 renderings for annual analysis. To automate this set, a Radiance batch script was used to generate a rendering for each date and time presented in figure 3.6b and Boston, Massachusetts was set as the location for all case-study renderings (Latitude 42 N, Longitude 72W). Although these metrics could eventually account for dominant sky conditions and evaluate the effects of climate on contrast and temporal variability, which will be presented and discussed further in Chapters 5-8, this chapter focused exclusively on a clear sky comparison, i.e. CIE clear sky with sun. In order to analyze the potential impacts of contrast over time, it was helpful to use a sky condition that allowed for maximized visual effects.

3.4 Results for Three Case Study Models

Annual spatial contrast is calculated for each set of radiance renderings produced using the method described in section 3.2. The results of these metrics are discussed in the present section for three

typological models that are representative of the whole range: case study one, four, and nine (figure 3.8). The relative numerical scale for spatial contrast has been determined by the results from all ten case studies using a statistical subdivision to generate a color scale of low, medium, and high values. Based on the distribution of resulting values, two thresholds divide the data for each metric into three parts, each representing a third of the population. A more comprehensive set of typologies would be needed to validate and further refine this initial categorization. As a result of the statistical subdivision in this chapter, spatial contrast values between 0 and 0.25 are considered low, values between 0.25 and 0.4 are considered medium, and values exceeding 0.4 are considered high (figure 3.4). In the remained of this section, we will use these thresholds to discuss the results in terms of relative high, medium, and low, although the generalizability of this range would, as previously mentioned, need to be studied further with a broader set of architectural examples.

3.4.1 Case Study One: Direct Top-lit Space

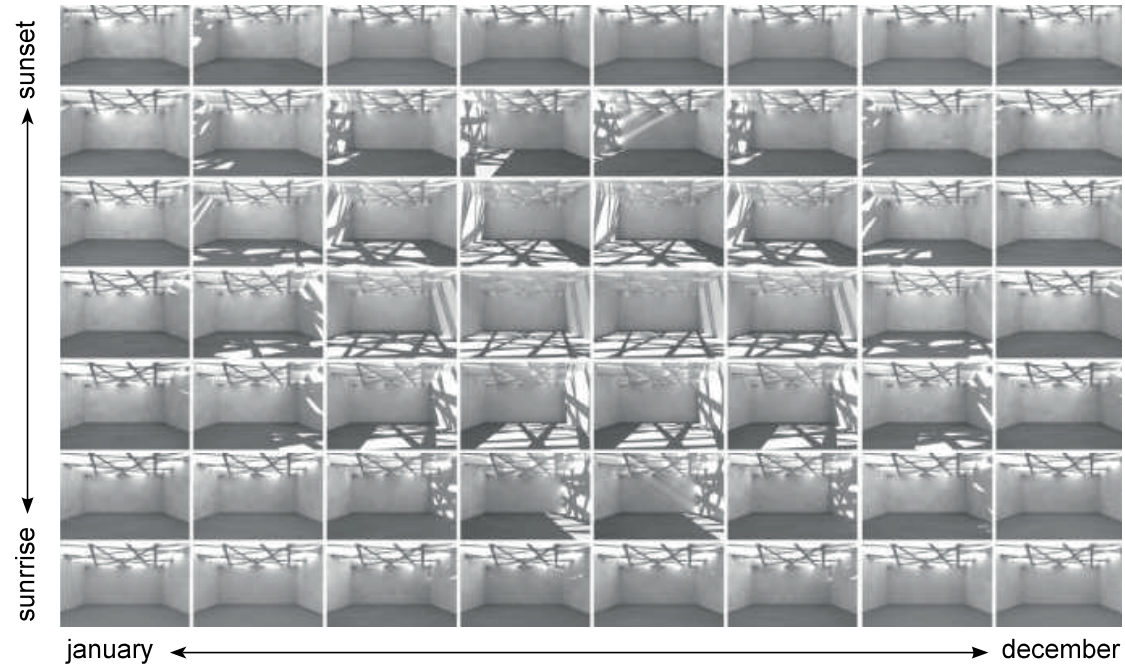
Case study one was modelled to represent a highly contrasted and variable interior daylight environment with an open roof structure that casts dynamic and articulated patterns of light and shadow down onto the walls and floor. From the dynamic renderings presented in Figure 3.10a, the results from case study one demonstrate a high degree of spatial contrast throughout the year. The temporal map in Figure 3.10c shows a peak in spatial contrast between 10 a.m. and 3 p.m. in the summer months when the sun is directly overhead. In Figure 3.10b, thick red lines signify locations where spatial contrast was most consistent, highlighting the roof structure as the most redundant source with secondary accumulations on the floor and walls.

3.4.2 Case Study Four: Side-lit Space

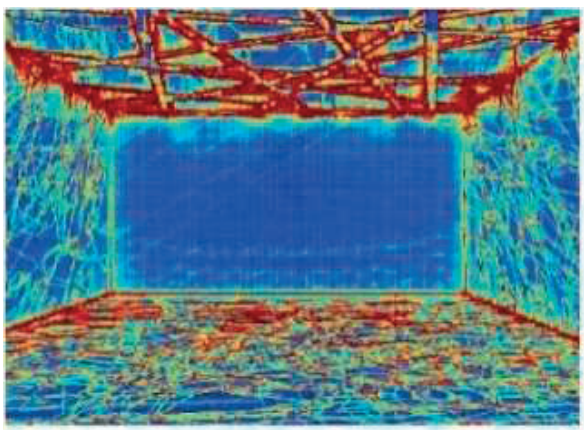
Case study four represents a more traditional side-lit daylight strategy with a clerestory window above and louvered screen below, creating varied effects across the year depending on solar altitude as shown in figure 3.11a. Here, the results for annual spatial contrast depict more seasonal variation, with a dramatic shift between the winter and summer months. The temporal map in figure 3.11c shows high spatial contrast between October and February, with moderate spatial contrast throughout the rest of the year. The location of these effects can be seen in the false color image in figure 3.11b, which shows the accumulation of contrast on the walls and floor closest to the wall of louvers.

3.4.3 Case Study Nine: Indirect Top-lit Space

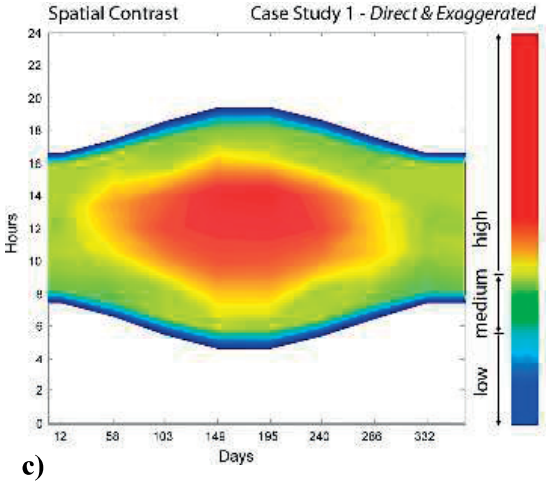
Case study nine contains a series of north-facing roof monitors that emit diffuse daylight down into the interior space. Across most of the day and year, case study nine achieves low spatial contrast (Figure 3.12c), however, there are slight variations that occur as sunlight penetrates the



a)



b)



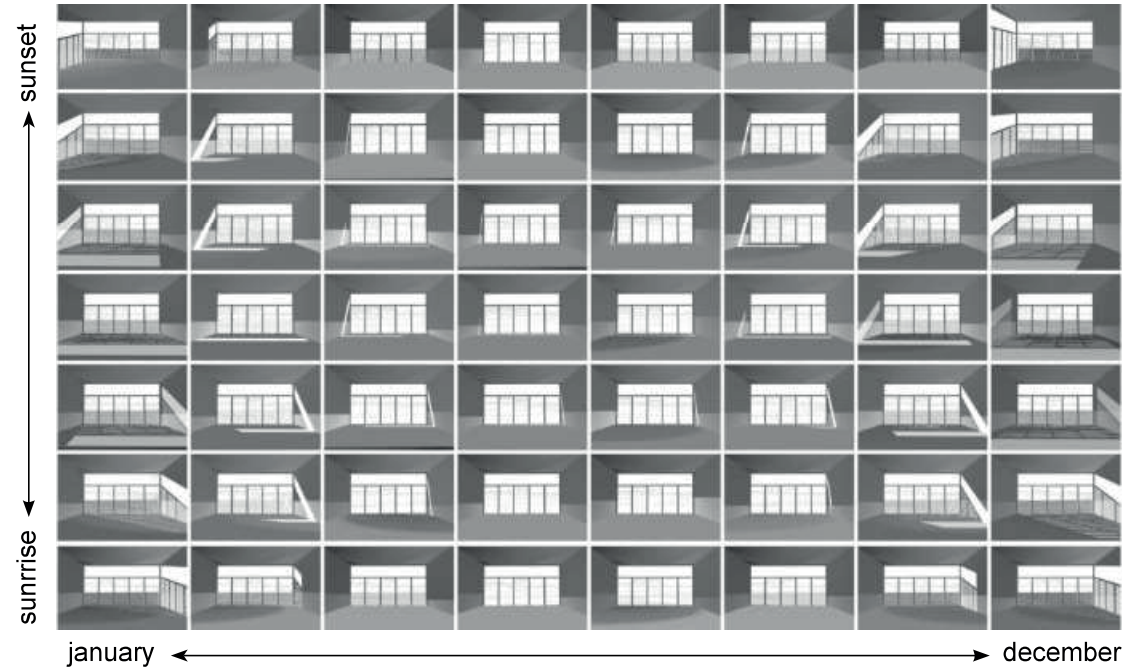
c)

Figure 3.10 Case Study 1 results for a) 56 annual renderings, b) spatial map of cumulative annual spatial contrast, c) temporal maps of dynamic results for spatial contrast.

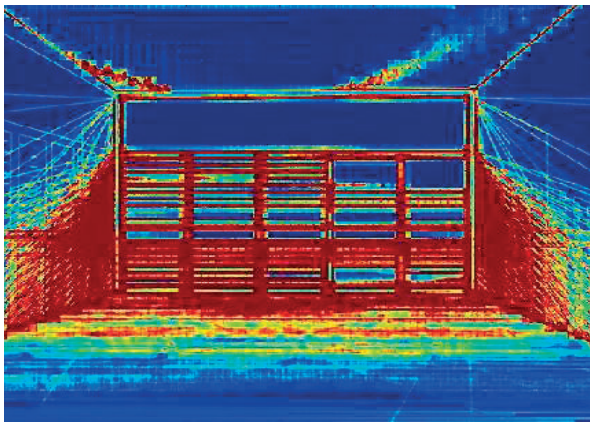
roof monitors in the early morning and late afternoon, as seen through the renderings in figure 3.12a. Figure 3.12b shows the location of cumulative spatial contrast the in roof monitors, with minimal contrast occuring on the floor and walls, due to the lack of direct sunlight penetration in this space.

3.5 Overview of the Results

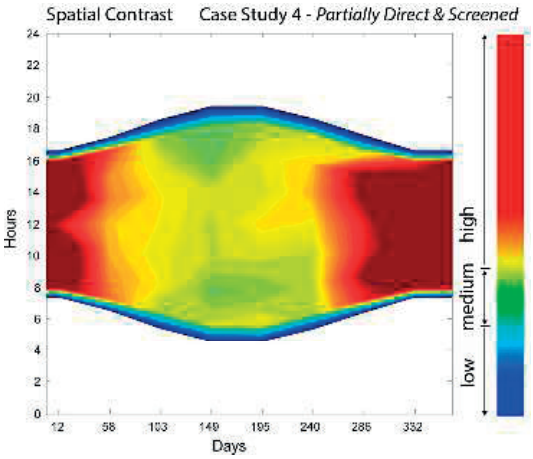
From the full set of ten case study results, a clear gradient can be identified between the high and low ends of the spatial contrast spectrum. The three case studies presented in section 3.4 show dynamic variations in spatial contrast as instantaneous values change over the course of



a)



b)

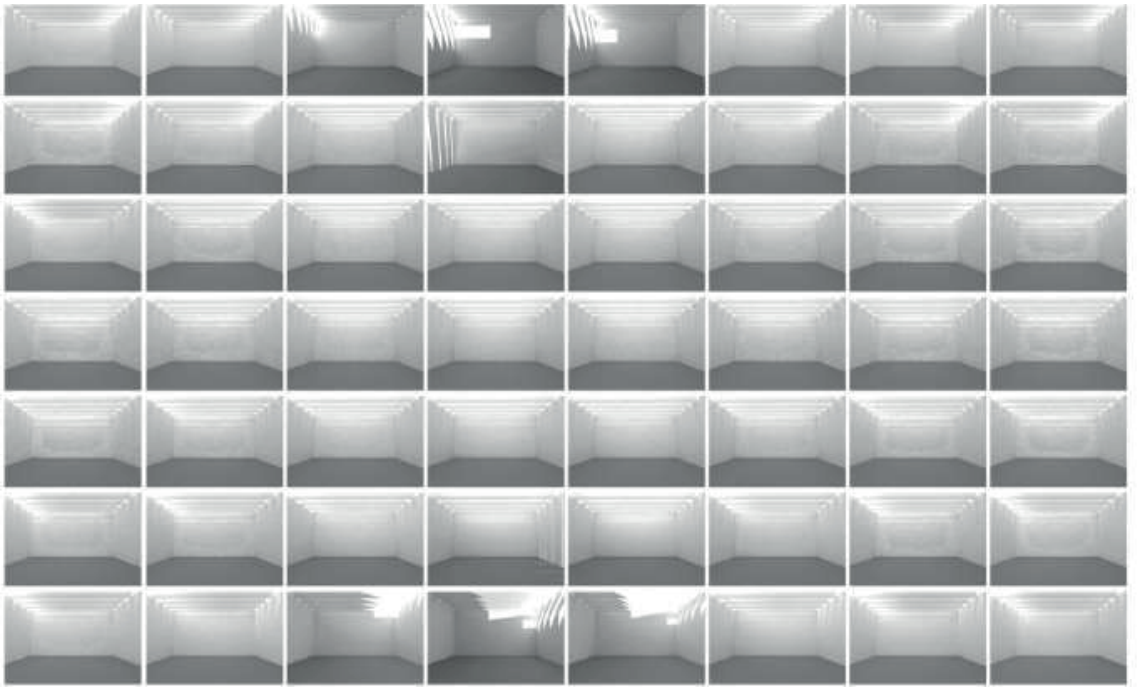


c)

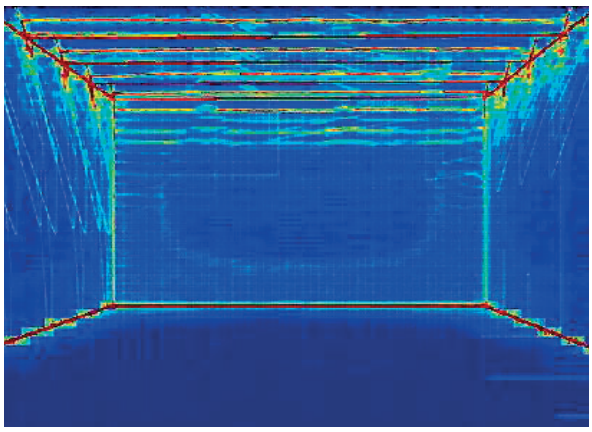
Figure 3.11 Case Study 4 results for a) 56 annual renderings, b) spatial map of cumulative annual spatial contrast, c) temporal maps of dynamic results for spatial contrast.

the day and year. It is clear from this proof-of-concept study that the SC algorithm proposed in this chapter must be considered on a temporal scale due to the dynamic nature of sunlight and its impact on the visual conditions within architecture over time. To compare the results for spatial contrast across the entire year, data for each of the 56 instances was converted into a single cumulative annual value. For annual spatial contrast, this number represents the sum of spatial contrast across all 56 instantaneous images.

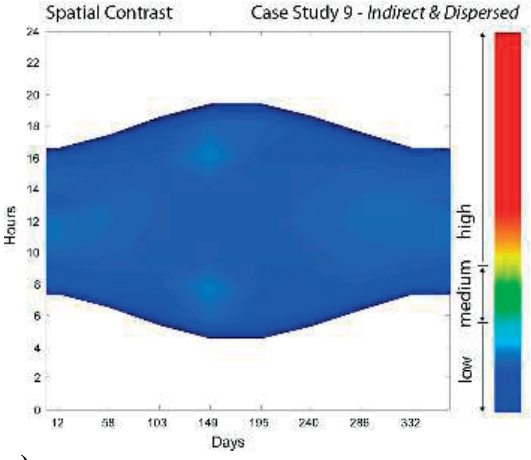
Figure 3.13 shows normalized values (between 0 and the maximum cumulative value achieved across all ten case studies) for annual spatial contrast. The resulting value is described by the sum of instantaneous spatial contrast across all 56 instances. The results are plotted against the author's



a)



b)



c)

Figure 3.12 Case Study 9 results for a) 56 annual renderings, b) spatial map of cumulative annual spatial contrast, c) temporal maps of dynamic results for spatial contrast.

intuitive ranking with a dotted line to the show the linear trend. This simple scatter plot between the normalized annual SC values for each space and the author's intuitive ranking is helpful in determining whether the proposed measures agree. As a proof-of-concept study, the proposed algorithm SC differentiates between spaces that were intuited, subjectively by the authors, to vary in terms of contrast composition and variability over time. A more rigorous validation is needed to prove that the authors' intuitive ranking is broadly consistent with a larger population of subjects. As such, chapter 5 will use subjective data from an online survey to compare ratings of contrast and characteristics related to visual interest to a larger, more detailed, and varied sample of architectural renderings.

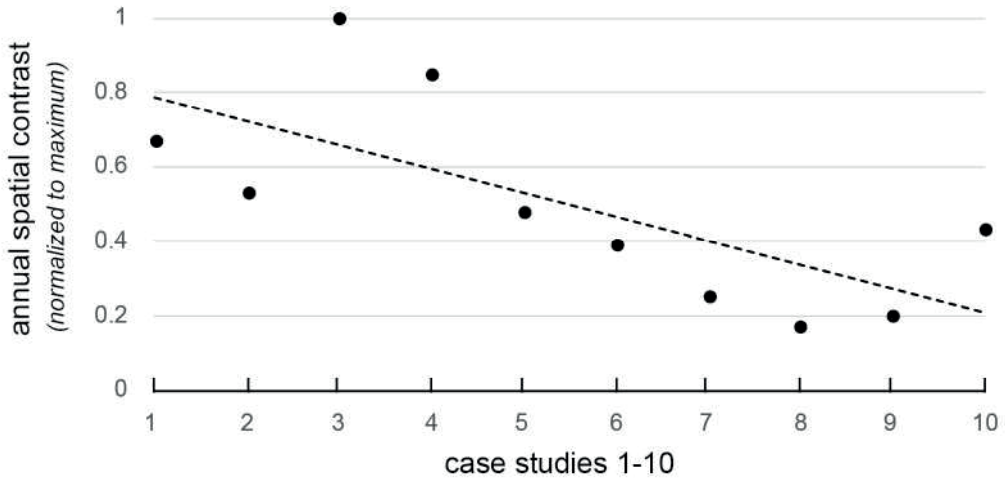


Figure 3.13 Results for annual spatial contrast for all 10 case studies. Values are normalized from 0 to the maximum across all spaces.

Figure 3.14 shows a side-by-side comparison of annual spatial contrast results for each case study. From figure 3.13 and figure 3.14, we can see that case studies one, two, and ten break the linear trend for annual spatial contrast, despite achieving relatively high instantaneous values in the time series simulation. In retrospect, the authors acknowledges that case study two was located in the wrong relative position within the intuitive gradient and that it should have been located closer to the middle of the matrix, between case studies five and six. The authors predicted that the dia-grid geometry would produce more dominant shadow patterns within the interior, while in actuality the results were quite similar to the side-lit office space in case study five. The only difference between case studies two and five is, of course, the specific geometry of openings, but they produce similar amounts of contrast and variability within the interior.

If we consider the temporal maps of instantaneous results in Figures 3.10c and 3.11c, the results for spatial contrast reveal surprisingly strong temporal changes between the two case studies 1 and 4. These changes were difficult to anticipate due to the change in incident solar angles, and we believe that the discrepancies between intuition for a single instance of time and the revelations revealed by the annual time series exemplifies the importance of dynamic visual analysis methods. In this case, the metrics helped to reveal perceptual changes within the visual field that were difficult to anticipate by the experts due to the dynamic nature of sunlight over time. If our intuition is unable to 'predict' the variation in contrast at different times of the day or year, then these annual metrics can help designers to make objective decisions in the selection of various design scenarios.

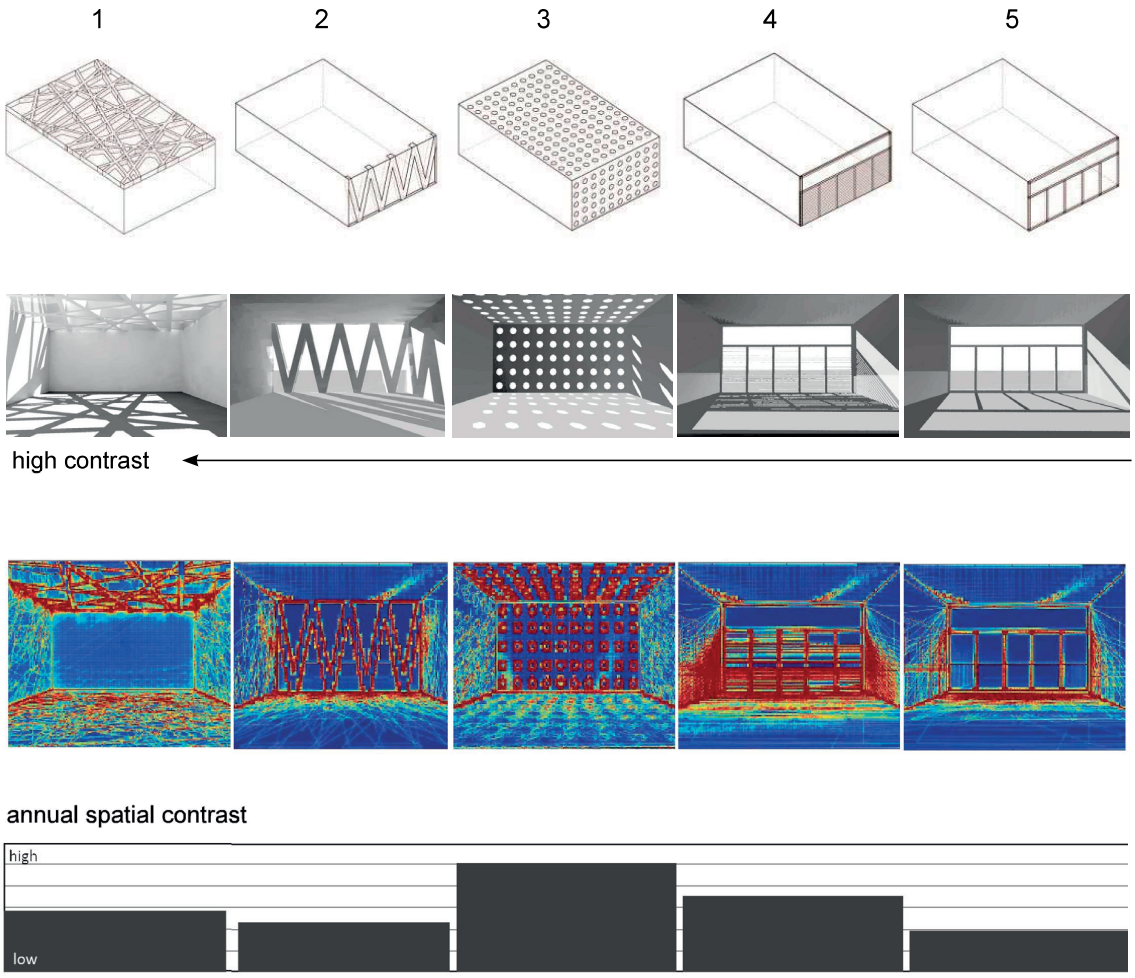


Figure 3.14 Case study result for annual spatial contrast for all 10 case studies. All values are normalized from 0 to the maximum cumulative value across all spaces.

Figure 3.15 show the annual results for each case study space 1-10, normalized from 0 to the maximum cumulative value across all spaces and plotted linearly. Each space is identified by a color, in a gradient from dark blue (case study 10) to light blue/light orange (case studied 6 and 5) to dark orange (case study 1). What is interesting to see in these results is the diurnal swing in the strength of spatial contrast, especially for those spaces in the orange (high) spectrum. This is due in large part to the variability of interior lighting conditions in these spaces due to large spans of glass and mullions or structural elements which create spatial diversity in neighboring pixel values and variations in these values over time. The spaces in the blue spectrum are much more stable across the day. The space with the largest diurnal swing is case study 1 due to the highly dynamic movement of sunlight across the floor and walls, which peaks in the summer around noon with higher horizontal sun angles. A look at the variability in these instantaneous values provides a useful insight when compared alongside the cumulative annual values presented in

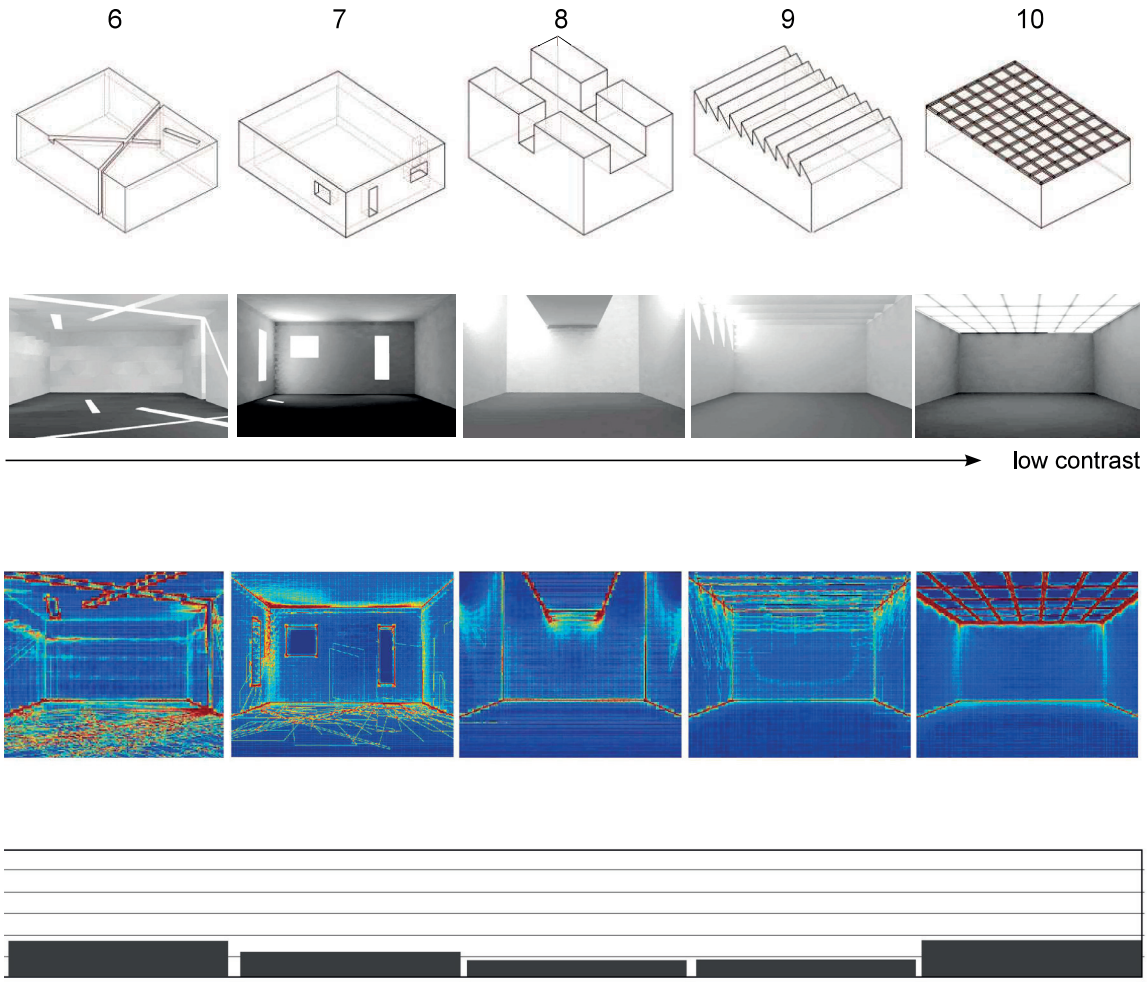


Figure 3.14 continued...

Figure 3.14. It is hard to compare the strength of contrast and its variability through a collapsed cumulative or average annual value as high instantaneous values can be reduced when averaged across a time series. As space is also not perceived as a cumulative set of instances, it is important to consider the variable temporal nature of contrast and not aim for a reduced annual value when making comparisons about the visual impacts of daylight on the indoor environment.

This proof-of-concept study allows for the relative comparison between each of the ten case study spaces using the proposed algorithm for spatial contrast. Although the proposed algorithm requires further development and will be compared alongside an expanded set of contrast algorithms in the following chapter, this proof-of-concept demonstrates the capacity of spatial contrast to measure a relative set of characteristics within daylit architectural renderings. Furthermore, it allows us to see when and to what degree those characteristics change over time as a result of dynamic daylight conditions.

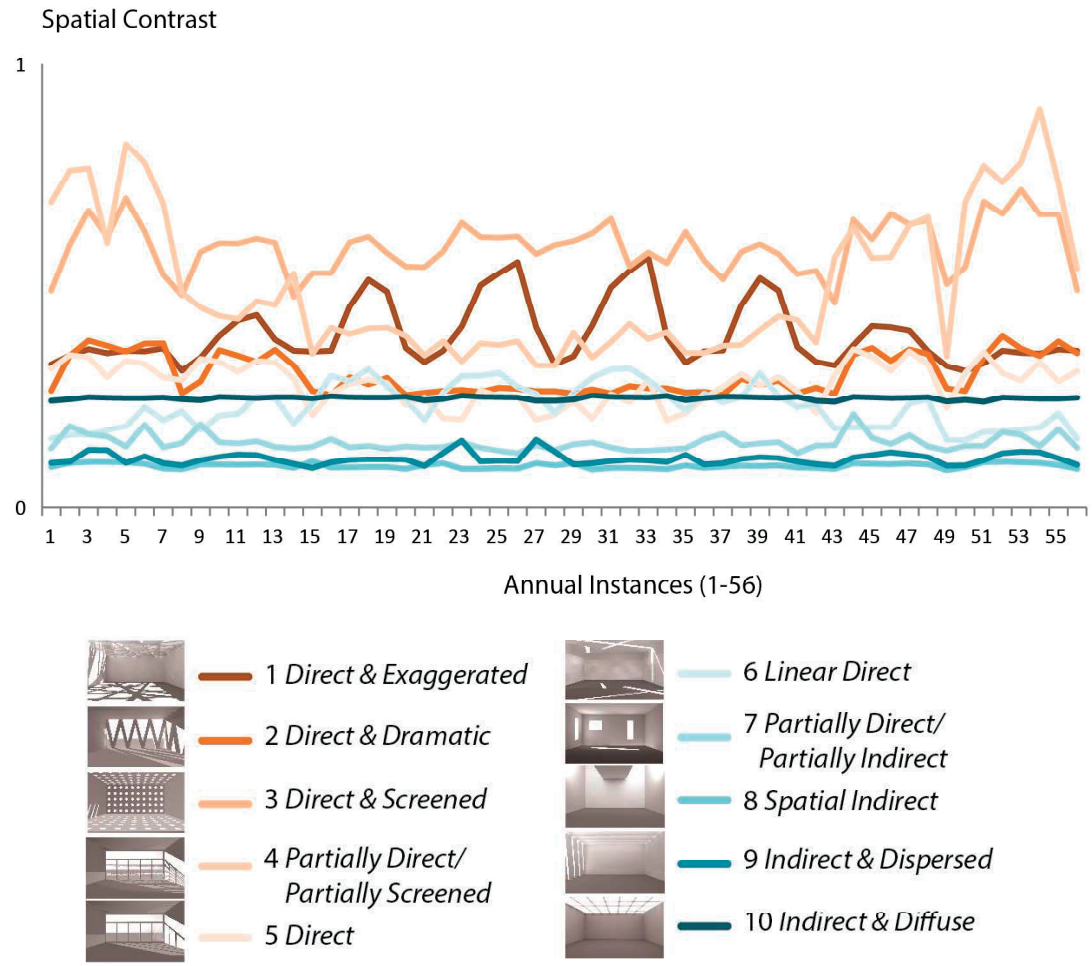


Figure 3.15 Linear temporal plot of spatial contrast for all 10 case studies. All values are normalized from 0 to the maximum value.

3.6 Discussion

The application of the spatial contrast to the ten case studies introduced in section 3.4 was an attempt to link an intuitive assessment of perceptual visual effects to an algorithm that may be used to predict them. To further develop and verify the utility of these new image-based measures, this thesis will progress in a few parallel pathways.

The first path will compare an expanded set of image-based algorithms to determine whether there are other, more effective methods of computing compositional contrast. Chapter 4 will used the same case study models presented here and compare a range of image-based algorithms to the authors own intuition. To compare those algorithms to subjective data from a broader population, more detailed architectural renderings will then be generated for use in an online user/based experiment. We will conduct a web-based survey to evaluate the relationship between subjective attributes of contrast and visual interest in 2D renderings and the algorithms proposed by the literature. A new model for predicting perceptual attributes of visual interest will then be developed

based on the resulting subjective data, which has only been discussed in this chapter based on the authors' trained, but limited, intuition. A refinement of the quantitative measures presented in this chapter is necessary as the current method takes an accumulative difference between neighboring pixels to produce a set of boundary conditions within a given image. Although this accounts for a fine level of detail in local luminance variation, the measure is dependent on pixel density and does not account for larger regions of contrast that may be perceived by the human eye.

The second path will consider the role of view direction, field-of-view and visual immersion in subjective impressions of daylight in architecture as well as algorithms used to predict these impressions. Where the current study is constrained by a 2D rectangular field-of-view, Chapters 6 - 7 will jump from a rectangular to a hemispherical view, applying the 2D algorithms developed over the next couple of chapters to a range of hemispherical images, representing a 360° view-range from a fixed position in space. Chapter 8 will then seek to validate this approach through an experiment which uses semi and fully immersive rendered scenes to collect subjective ratings in a virtual reality headset.

Finally, it is important that these novel metrics be integrated into a work flow so that perceptual performance may be measured alongside other performance measures to provide a more holistic evaluation of daylit space. The Lightsolve project originally,proposed the integration of these metrics alongside non-visual and dynamic comfort metrics as part of an integrated tool to assess human needs in daylit architecture (Andersen et al., 2013a & b). Through an integration of perceptual field-of-view metrics like those proposed in this chapter, Chapters 6 & 7 will propose a new integrated workflow and platform for simulating human-centric performance in an architectural case study alongside a non-visual health-based metric.

3.7 Chapter Summary

This chapter began with a review of existing contrast measures to establish the need for new indicators that can account for a range of perceptual and temporal qualities found within daylit architectural space. These qualities, once measurable, can then be positioned alongside existing task-based illuminance and visual comfort metrics to provide a more holistic analysis of daylight performance in architecture. To contextualize these perceptual and temporal qualities within the discipline of architecture, the authors used a typological matrix of daylit spaces using intuitive judgement to rank examples (Section 1.1.3) (Rockcastle & Andersen, 2013b). These typologies reveal a diverse range of daylight-driven visual effects. Using this intuitive matrix as context, this chapter then introduced a new local contrast algorithm: spatial contrast, which was developed to measure the effects of spatial and temporal diversity of daylight in architecture using a time series of renderings

In order to measure the intuitive effects described above, HDR renderings were used to record luminance levels within a selected view and then compressed them down into 8-bit RGB images, providing a compact range of values that could be analyzed. Although spatial contrast looks at the variation between neighboring pixel values within a selected image, annual spatial contrast accounts for the dynamics of contrast throughout the year. Through an analysis of an annual time series (56 renderings that represent an even subdivision of daylit hours across the year; 7 daily and 8 monthly) the designer can identify the magnitude of spatial contrast over time and visualize these dynamic effects through a combination of accumulative spatial images and annual temporal maps. When applied to the ten case study spaces selected for this study, spatial contrast produced a linear trend that relates to the author's intuitive ranking and serves as a proof-of-concept.

This new metric has shown the potential to communicate information about the spatial and temporal variability of daylight, providing architects with a potential new tool for comparing the magnitude of visual effects within architecture. The implications of this approach are widespread, from a simple analytical tool for describing dynamic daylight conditions, to an objective approach that compliments the use of task-based illumination and visual comfort metrics in a variety of programmatic conditions. By establishing an intuitive gradient of visual effects and producing a method for quantifying those effects over time, we are able to re-focus the discussion on daylight performance to regard those perceptual qualities of light that are so essential to contemporary practice. The following chapter will build upon this proof-of-concept with the comparison of an expanded set of contrast-based algorithms to develop a computational approach that best fits the authors own predictive intuition before jumping into an experimental phase in Chapter 5.

4 COMPARISON OF CONTRAST ALGORITHMS TO INTUITIVE RANKINGs

This chapter will introduce a comparative analysis of existing contrast algorithms using the ten abstract case study renderings presented in the previous chapter. As the results in chapter 3 only took into consideration a single algorithm (spatial contrast), this chapter will compare an expanded set of local and global algorithms to the author's intuitive contrast ranking. The objectives of this analysis are two-fold: 1) to compare contrast predictions between existing algorithms and 2) to compare algorithm predictions to the author's intuitive ranking of contrast to develop the approach presented in the previous chapter. In the end of chapter 3, the authors discussed the need to evaluate additional measures, especially those multi-level algorithms introduced in Chapter 2 that can assess contrast across micro and macro scales (Section 2.4.2). This chapter will compare existing and modified contrast algorithms to one another as well as to the authors intuitive ranking to select an algorithm for use in the following chapter, where an online experiment will collect subjective ratings of contrast and contrast-based characteristics. The finding from this chapter build upon the proof-of-concept in Chapter 3 and provide a building blocks or the experiments presented in Chapter 5.

4.1 *Selected Algorithms*

The algorithms selected for this comparative study were introduced in Section 2.4.2 and represent a range of computational approaches to predicting contrast perception. The following contrast algorithms will be considered: Michelson (Michelson, 1927), RMS (Pavel, et al., 1987), Spatial Contrast (SC) (Rockcastle & Andersen, 2014), RAMM (multi-level average: RAMMG and single level: RAMM1) (Rizzi, et al., 2004), Difference of Gaussian (DOG) (Tadmor & Tolhurst, 2000), and RSC (Simone, et al., 2009). From global measures like Michelson and RMS which evaluate pixel values independently from their compositional surrounds to local metrics like SC and RAMM (single and multi-level) and Gaussian metrics like DOG (single level) and RSC (multi-level), there are a range of possible approaches. In chapter 3, we introduced spatial contrast, a simplified version of the RAMMG algorithm (Rizzi, et al., 2004), using 4 neighbouring pixels rather than 8 and looking at the first resolution level (in this case 404 x 640 pixels). This simplified measure was computed across a time series of renderings for a number of abstract case studies to evaluate the potential for a contrast-based algorithm to differentiate between architectural typologies and changes in daylight composition.

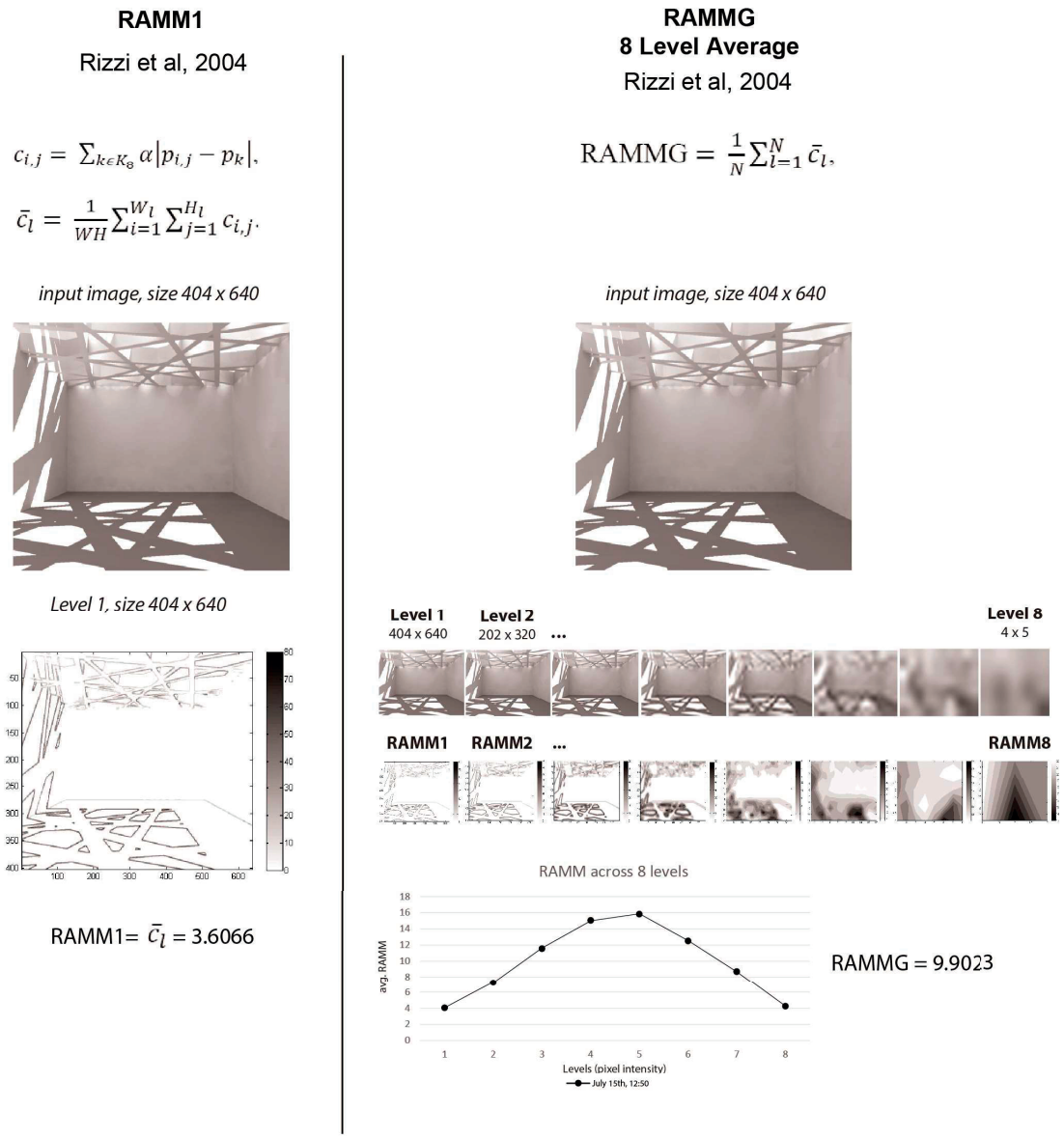
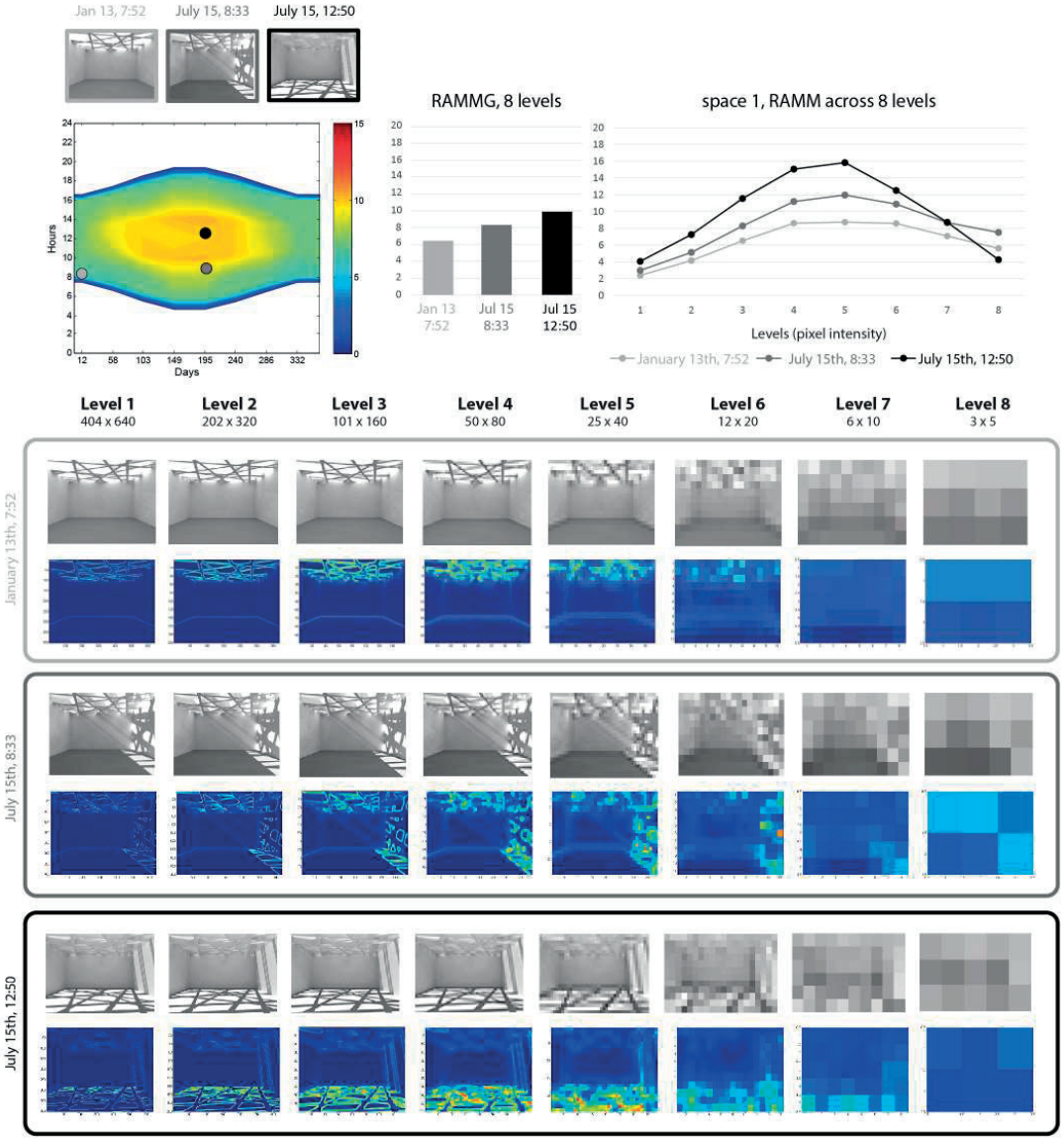


Figure 4.1 Showing the application of RAMM1 and RAMMG for an image with resolution 404 x 640. Contrast predictions for RAMMG are shown for each level in the pyramid subsampling as well as the average across all levels.

As a first proof-of-concept, the decision to use SC was based on an initial comparison of those measures that are most commonly found in the literature (Michelson and RMS) and the author's hypothesis that a localized neighbourhood measure could better predict compositional impacts of contrast on human perception. While chapter 3 found that SC better matched the author's intuition better than global measures like Michelson and RMS and was sensitive to changing sun positions across a time series, a broader analysis of related algorithms was not immediately feasible. This chapter introduces an expanded comparison of algorithms to reveal that a multi-level metric like RAMMG shows greater dependence to the author's intuitive contrast ranking (and therefore potentially better predictive capabilities) than a single level metric like SC or RAMM1, as was first presented.



Fifure 4.2 The application of RAMM across 8 subsampled levels for three dates (selected as the high, medium, and low instances of RAMMG from a time-series of 28 renderings). This analysis reveals the relative difference in RAMM predicitons at each subsampled level.

As mentioned in the chapter 3, SC is computed as the average difference between each pixel and it's 4 neighbouring pixels (equation 3.1). These localized contrast values are then summed across the image and divided by a hypothetical maximum contrast based on the image resolution (equations 3.2 - 3.3). It was initially thought that a ratio of the sum over the maximum would be more sensitive than the average of these values. In this chapter, however, we considered SC alongside RAMM1, which is computed by taking the average difference between each pixel and it's 8 neighbouring pixels using a weight applied to those pixels (equations 2.5 - 2.6) and then computing the average of those local values across the image (equation 2.4)(figure 4.1).

In addition to single level neighbourhood metrics like SC and RAMM1, this chapter includes multi-level measures like RAMMG (equation 2.3) which is computed using the same approach as RAMM1, but then applied to a pyramid of sub-sampled image resolutions, halved in each subsequent level using an anti-aliasing filter until reaching a minimum size (i.e. 3 x 5) which are then averaged across all levels (figure 4.1). The advantage of a multi-level measure like RAMMG is that it can account for micro and macro-scaled differences in local values across the image. For example, level 1 (404 x 640 pixels) will calculate many small localized pixel differences while level 4 (50 x 80) will calculate more mid-sized differences as the resolution becomes less sharp. The authors of this measure speculate that some levels of this measure might relate more to high-level, mid-level, and low-level vision parameters. Level 7 (6x10) and level 8 (3x5), for example, might relate to low-level contrast perception (Adelson, et al., 1984; Simone, et al., 2009).

Figure 4.2 shows the application of RAMM ($\overline{C_l}$) across three instances of time for each sub-sampled image resolution in space 1 (a case study introduced in the previous chapter). As was done in chapter 3 with SC, RAMMG was applied to 56 annual instances, allowing the authors to select three instances based on a relatively high, medium and low RAMMG prediction. The RAMM ($\overline{C_l}$) of each level was then plotted for each image to see how these values varied between time steps and across resolution levels. It is clear from the line graph in Figure 3.2 that some levels are more differentiated than others and that these vary depending on the resolution of that level. Appendix 1 (Figures A1.1 - A1.2) shows the same procedure applied to other spaces which vary in composition and per-level predications, drawing attention to the fact that each level within the RAMMG metric should be looked at in more depth to see if some are more important to contrast predications from the human visual system than others. This will be done in chapter 5, when subjective data from the online experiment will be analysed against a wide array of existing contrast measures.

The DOG measure (Tadmor & Tolhurst, 2000), introduced in section 2.4.2 takes the local differences between two bi-dimensional Gaussian filters with a centre and surround component (equations 2.7 - 2-11). These local differences are then averaged across the image (equation 2.12) As the combination of centre and surround radii can vary, the authors have assessed a range of radii combinations:

R_c=1, R_c=2 R_c=1, R_c=4 R_c=2, R_c=4 R_c=2, R_c=8 R_c=4, R_c=8 R_c=4, R_c=16

Figure 4.3 shows the application of DOG with R_c=1, R_c=2 to an input image (resolution 404 x 640) and the corresponding value for that resolution and radii combination.

The RSC measure (Simone, et al., 2009) is a combination of the DOG measure and the RAMMG measures mentioned above. The RSC measure computes the average DOG for each pixel across a pyramid of sub-sampled image resolution levels (equation 2.13), taking the average across those levels (equation 2.14). Figure 4.3 shows the application of RSC with R_c=1, R_c=2 to an input

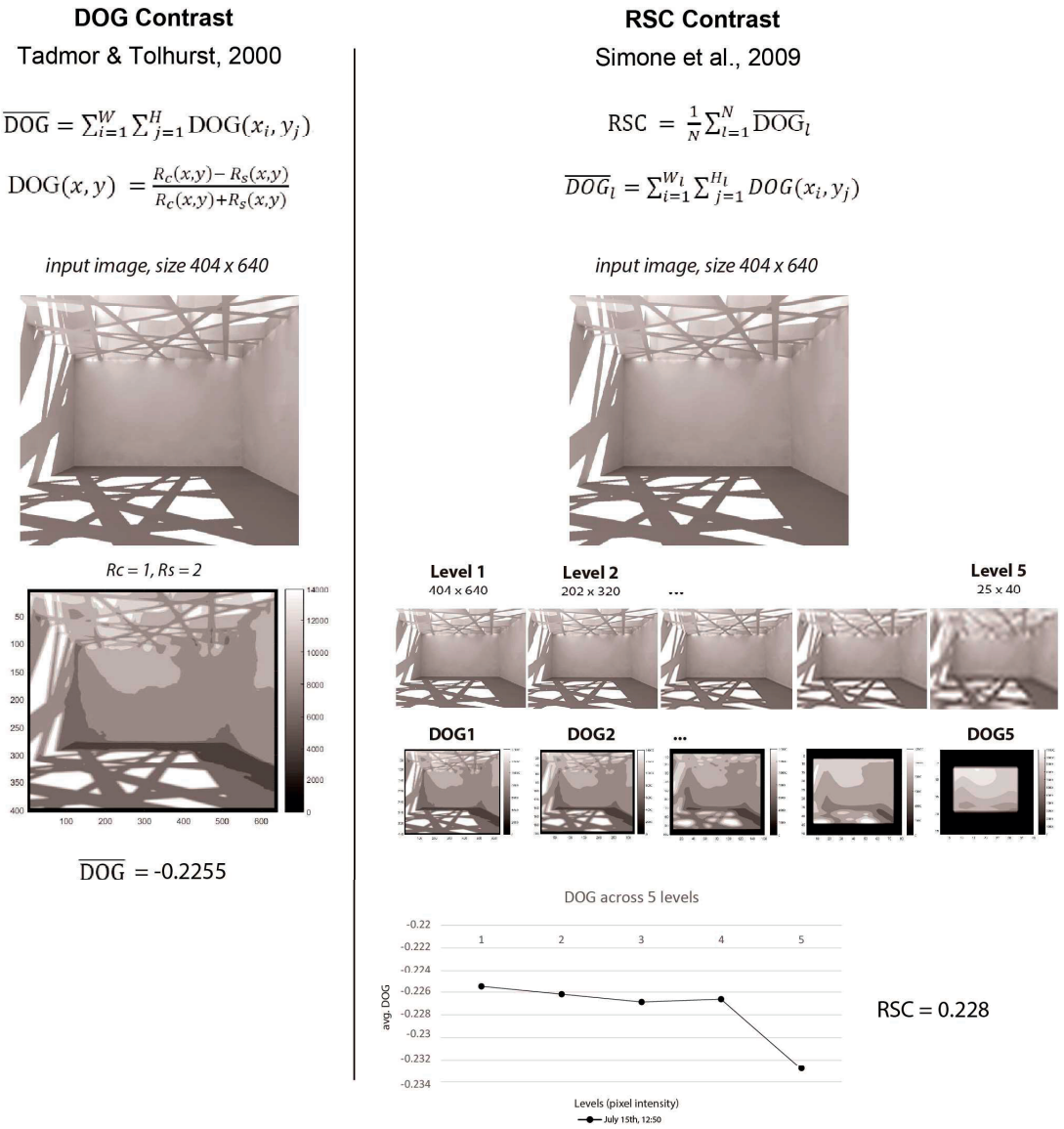


Figure 4.3 Showing the application of DOG and RSC contrast to a digital rendering of input resolution 404 x 640 pixels.

image (resolution 404 x 640), which is then sub-sampled 5 times to produce an average DOG across those levels. RSC computes the average DOG across fewer levels (in this case 4 and 5) that the RAMMG measure (8 levels) because the size of the Gaussian filter requires a minimum pixel radii and cannot compute across too small resolutions. For example, if the Gaussian filter of the surround is $3R_s$ and R_s=2, than the minimum resolution of the final level must be 13 x 13 (or larger). In this chapter we have looked at RSC for R_c=1, R_c=2 and compared the average across 4 and 5 levels as the authors in (Simone, et al., 2009) did not specify the number of steps as a function of the input radii combinations. In this case, there is a big difference between the RSC taken across 4 vs. 5 levels.

4.2 Comparison of Algorithms Using Case Study Renderings

Using SC, the algorithms mentioned in Section 5.1 (RAMM1, RAMMG, DOG, and RSC) and the global contrast measures RMS and Michelson, this section apply each metric to static renderings of the ten case studies presented in chapter 3 and compare their contrast predictions. To compare measures which vary in the unit and distribution of values, the authors applied each algorithm to the set of 10 static renderings and normalized the results from 0-1 using

$$X' = (X - X_{min})/(X_{max} - X_{min})$$

Measures like DOG that produce a negative value were first inverted and then normalized. These normalized values were then compared using Pearson correlation coefficient values to identify dependencies between measures. Figure 4.4 and Table 1 show Pearson values for each combination of selected measures, with values $r \geq 95$ highlighted in grey.

As was mentioned previously, the authors wanted to compare SC values to the first level RAMM prediction (RAMM1) with input resolution 404 x 640 pixels. In this analysis, we can see that predictions of contrast between these measures are very highly correlated (r =0.999), meaning that the average local value (equations 2.7 - 2.8) is redundant with the ratio of sum over hypothetical maximum (equations 3.2 - 3.3). While many measures show a high interdependence, the measures which appear to be the most highly correlated are $DOG_{RC=1, RS=2}$ to $DOG_{RC=1, RS=4}$ to $DOG_{RC=2, RS=4}$ ($r \geq 0.98$) and $RSC_{RC=1, RS=2}$ 4 level average to $DOG_{RC=2, RS=8}$ ($r = 0.98$).

Given these results, it appears that DOG measures (using the above-listed center to surround combinations) could be collapsed in any future analysis. While it is useful to understand which measures are highly correlated and may be redundant, the objective of the previous chapter was to compare contrast algorithms to the author's intuitive ranking. To compare this expanded set of algorithms using the same intuitive ranking, section 4.3 will look at the correlation between algorithms and rankings.

4.3 Comparison of Algorithms to Intuitive Ranking

The author acknowledges that her own intuitive ranking of perceived contrast is biased and any generalizable indicator must be compared to a broad population of subjects because humans vary in responses. For this reason, the experiments in Chapter 5 and 8 will compare contrast algorithms to a group of subjects and compare distributions of rankings to find those algorithms that agree with the broader population. Nevertheless, this section uses the same ranking presented in chapter 4 as a first step to make comparisons between intuition and quantitative prediction methods in an attempt to understand the relative difference between measures. Due to the high correlation between SC and RAMM1, we will collapse these two measures and only consider RAMM1 from here on.

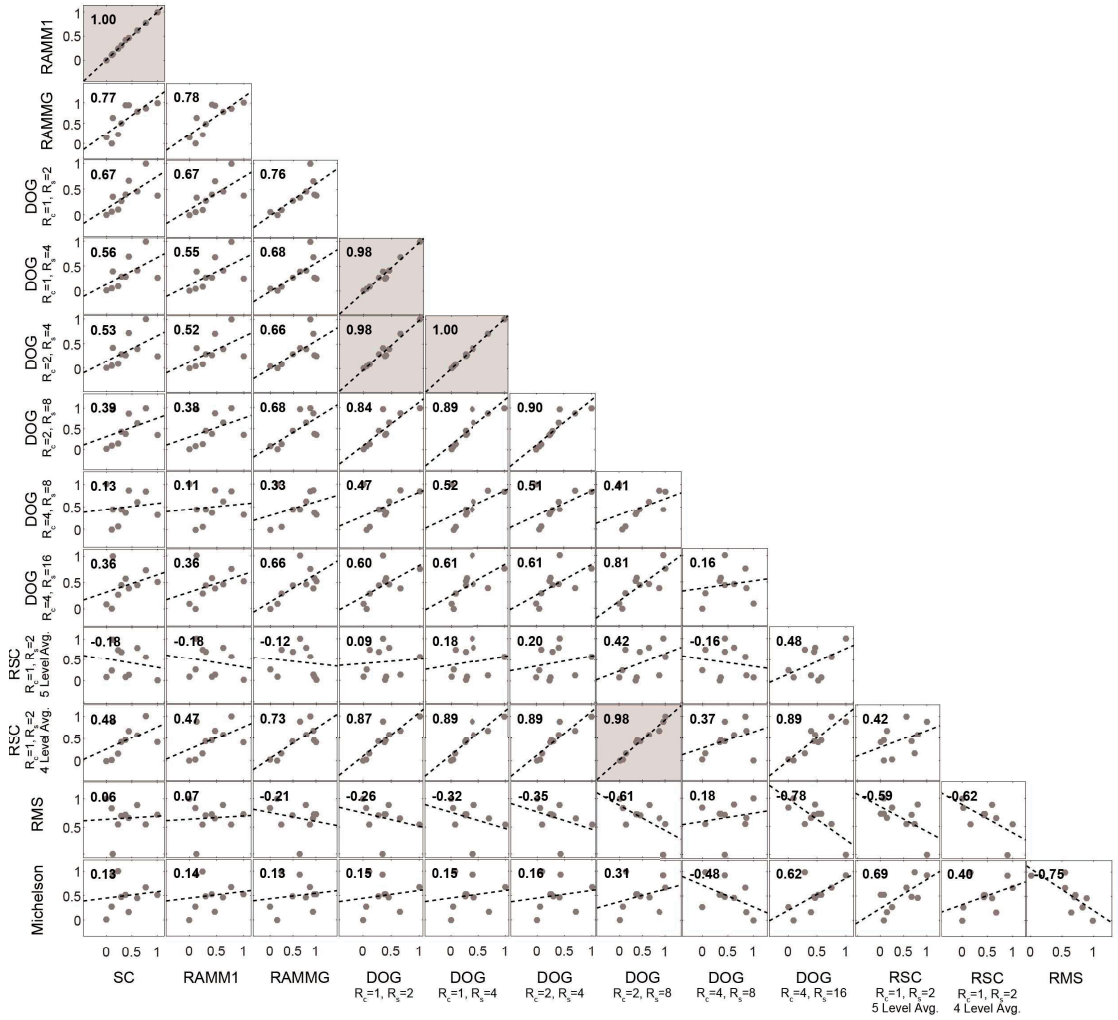


Figure 4.4 Pearson Correlation Coefficient values f (*r*) or all contrast algorithms (SC, RAMM1, RAMMG, DOG (in various combinations of R_c and R_s), RSC (4 level average and 5 level average), RMS, and Michelson.

Table 4.1 Showing Pearson Correlation Coefficient (*r*) values for all contrast algorithms (*r* values greater than 95 are highlighted in gray).

	SC	*RAMM1*	*RAMMG*	*DOG_12*	*DOG_14*	*DOG_24*	*DOG_28*	*DOG_48*	*DOG_416*	*RSC_12 (5 level)*	*RSC_12 (4 level)*	*RMS*
RAMM1	**0.999**											
RAMMG	0.767	0.778										
DOG_12	0.673	0.666	0.764									
DOG_14	0.558	0.549	0.682	**0.983**								
DOG_24	0.532	0.521	0.664	**0.975**	**0.999**							
DOG_28	0.388	0.385	0.677	0.845	0.891	0.900						
DOG_48	0.126	0.109	0.328	0.465	0.516	0.515	0.410					
DOG_416	0.358	0.359	0.665	0.604	0.608	0.611	0.814	0.158				
RSC_12 (5 level avg.)	-0.178	-0.180	-0.118	0.086	0.180	0.197	0.419	-0.156	0.481			
RSC_12 (4 level avg.)	0.477	0.474	0.734	0.868	0.890	0.893	**0.978**	0.371	0.888	0.418		
RMS	0.065	0.066	-0.211	-0.258	-0.324	-0.349	-0.609	0.179	-0.779	-0.587	-0.619	
Michelson	0.134	0.137	0.132	0.153	0.152	0.156	0.310	-0.481	0.615	0.692	0.403	-0.749

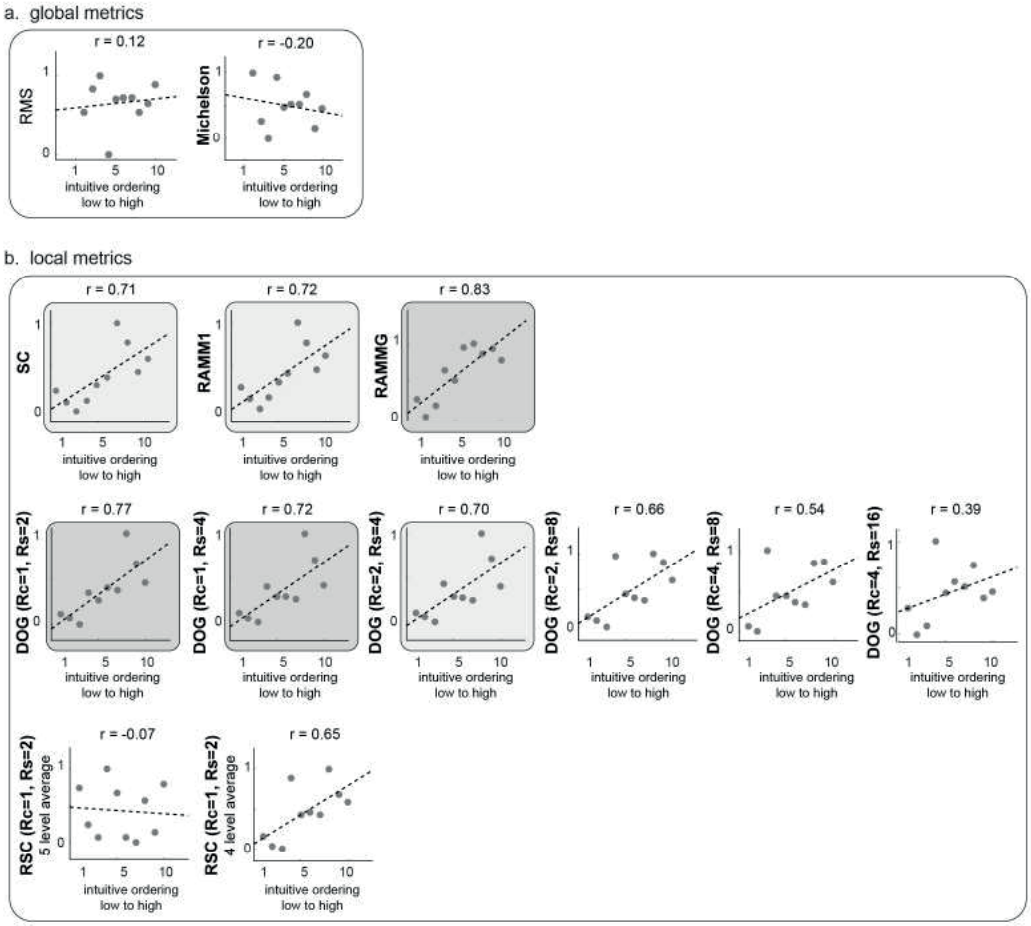


Figure 4.5 Pearson Correlation Coefficient values (*r*) between the author's intuitive ranking and each of the selected contrast algorithms.

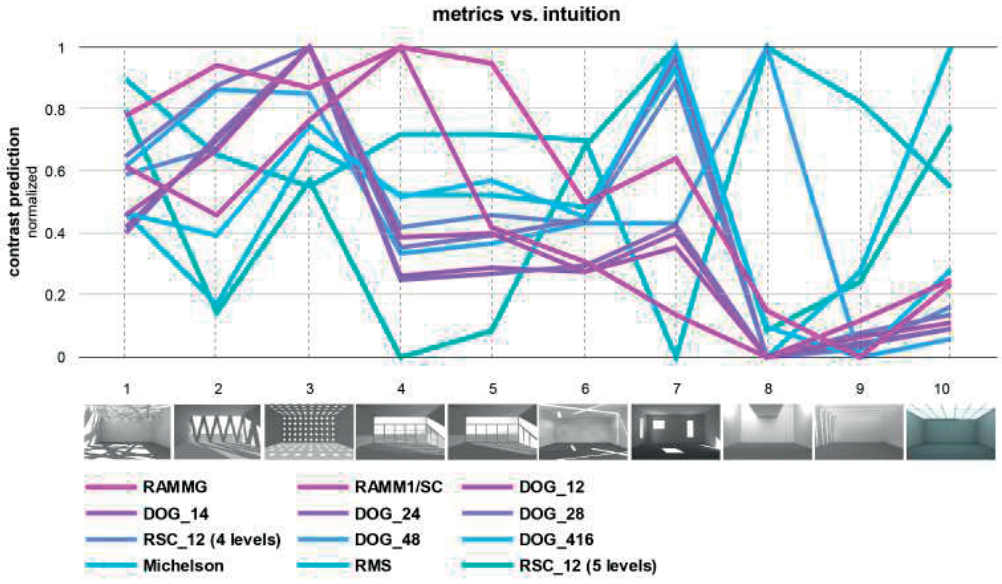


Figure 4.6 Graph showing each of the selected contrast algorithms plotted against the author's ranking of renderings from high to low (with space 1 receiving the maximum perceived contrast and 10 receiving the minimum).

Figure 4.5 shows Pearson values between the author's intuitive orderings (from low to high) and the normalized values for each algorithm. Figure 4.5a shows the correlation between intuitive rankings and global metrics *RMS* (r = 0.12) and *Michelson* (r = -0.20), while Figure 5.5b shows the correlation between rankings and local metrics *SC* (r = 0.71), *RAMM1* (r = 0.72), *RAMMG* (r = 0.83), $DOG_{RC=1,\ RS=2}$ (r = 0.77), $DOG_{RC=1,\ RS=4}$ (r = 0.72), $DOG_{RC=2,\ RS=4}$ (r = 0.70), $DOG_{RC=2,\ RS=8}$ (r = 0.66), $DOG_{RC=4,\ RS=8}$ (r = 0.54), $DOG_{RC=4,\ RS=16}$ (r = 0.39), $RSC_{RC=1,\ RS=2}$ *5-level* (r = -0.07), and $RSC_{RC=1,\ RS=2}$ *4-level* (r = 0.65). In this analysis, RAMMG shows the highest correlation, followed by $DOG_{RC=1,\ RS=2}$. Compared to the other measures, $DOG_{RC=1,\ RS=4}$, $DOG_{RC=2,\ RS=4}$, and RAMM1 were also highly correlated.

Figure 4.6 shows each metric plotted against the author's intuitive ranking to compare each algorithm in its prediction of contrast for each image. RAMMG calculates over 0.75 for images 1-4 and under 0.25 for images 8-10 (in normalized values) which is most consistent with the author's intuitive ranking. The DOG measures with small center to surround radii combinations (R_c=1 R_s=2, R_c=1 R_s=4, R_c=2 R_s=4) show similar relative values for each image (+/- 0.1), while they all seem to under predict the perceived contrast in images 1 and 4 (as compared to RAMMG). Images 1, 3, and 6 produce the least variation in predictions of contrast between algorithms, while images 4, 7, and 8 produce both the lowest and highest predictions of contrast between algorithms. This spread in prediction between algorithms makes the lack of consensus between methods quite clear. From this analysis, it becomes clear that the only way to determine which algorithm is the most effective for predicting contrast is to compare them to a larger population of subjective data.

4.4 Chapter Summary

This chapter introduced an analysis of existing contrast algorithms using a set of 10 abstract case study renderings, originally presented in chapter 3. The objectives of this analysis were two-fold: 1) to compare contrast predictions between existing algorithms and 2) to compare algorithm predictions to the author's intuitive ranking of contrast to identify those measures that best match. The findings reveal high correlations between several algorithms, including SC and RAMM1, and show that RAMMG and $DOG_{RC=1,\ RS=2}$ are the best predictors of contrast based on the author's intuitive ranking. While SC outperforms global measures like Michelson and RMS, as originally hypothesized, it appears that a multi-level measure or some combination of measures may prove more robust as predictors of contrast-bases visual effects.

Figure 4.7 shows a visual comparison between the author's intuitive ranking (1 being the lowest and 10 being the highest) and measures for RAMMG, RAMM1, and $DOG_{RC=1,\ RS=2}$ algorithms, which were shown to have the highest correlation in the previous section. As we can see in 5.7a, RAMMG predicts high relative values of contrast for spaces 1-5 (RAMMG≥0.75) and low relative values for spaces 8-10 (RAMMG ≥ 0.25). While a linear rank is not always maintained, an issue that will be discussed in more depth in the following chapter, there appears to be a

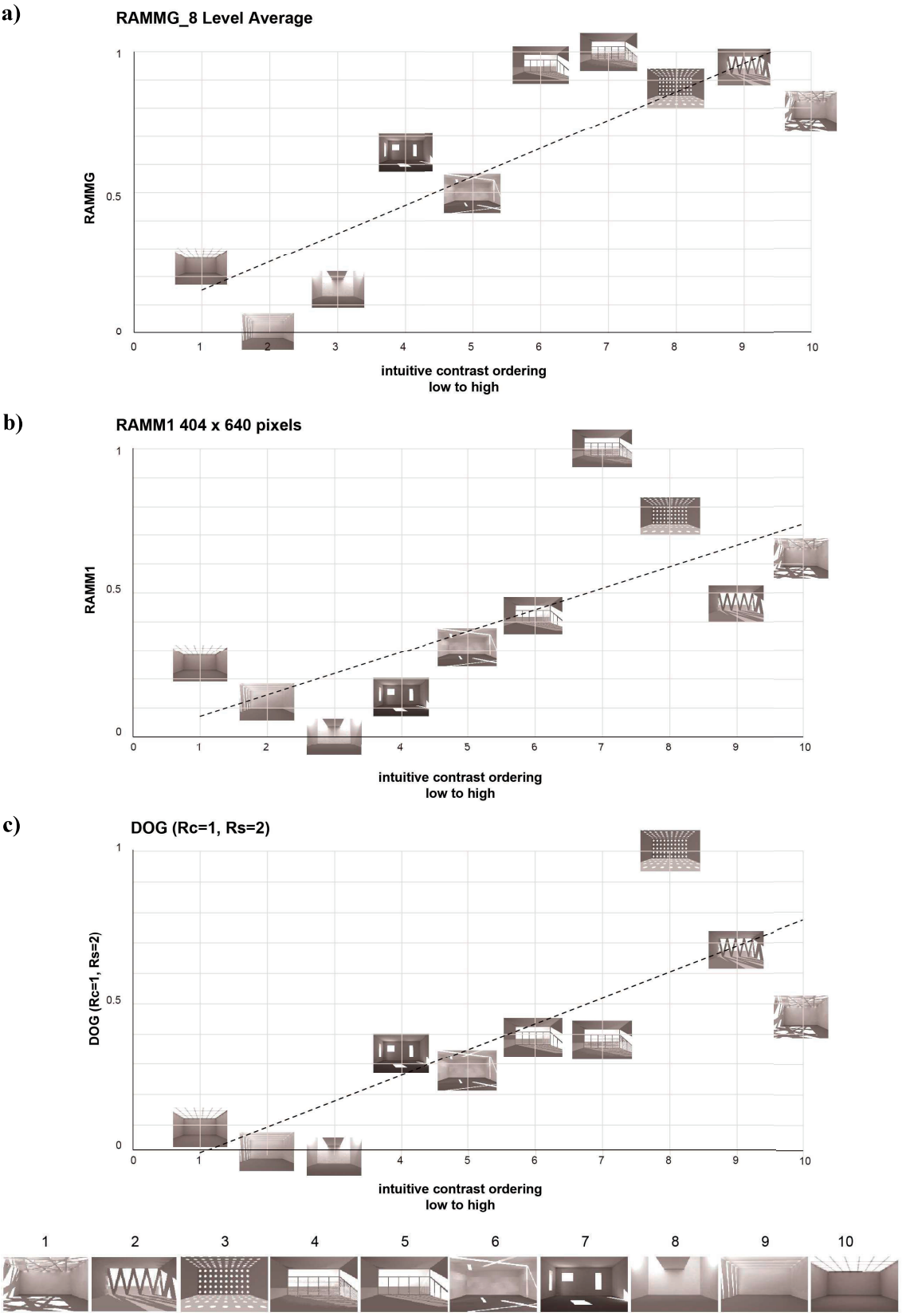


Figure 4.7 a) plot of RAMMG against the author's intutive ranking, b) plot of RAMM1 against the author's intuitive ranking, and c) plot of $DOG_{Rc=1, Rs=2}$ against the author's intuitive ranking for spaces 1-10.

threshold over which a relatively high prediction is achieved (matching the author's rank) and a threshold under which a relatively low prediction is achieved (matching the author's rank).

Figure 4.7b shows a comparison between the author's intuitive ranking and RAMM1 (found to be highly correlated and thus interchangeable with SC). In this case, we cannot see a clear set of thresholds for high and low predictions (as in the previous graph), with spaces 3 and 4 producing a relatively high RAMM1 and space 2 producing a relatively moderate RAMM1 as compared to the author's ranking. Using RAMM1, spaces and 4 achieve the highest contrast prediction, with spaces 1 and producing only a moderate contrast prediction despite institutive ordering.

In Figure 4.7c, $DOG_{Rc=1,\ Rs=2}$ achieves one the best linear fits for spaces 4-10, but appears to under predict the perceived contrast in space 1 and over predict the contrast in space 3. Give this singular intuitive comparison, which is admittedly limited, it appears that RAMMG outperforms SC/RAMM1 as a quantitative predictor for contrast in digital renderings. While the author's intuition alone cannot be generalized to a larger audience due to the subjective nature of the analysis, it provides a first comparison between quantitative measures. In the following chapter, we will compare subjective ratings of 2D renderings from an online survey to make a more informed validation of these algorithms in predicting specific contrast-driven visual effects.

5 A WEB-BASED EXPERIMENT ON VISUAL INTEREST USING 2D RENDERINGS

Building upon the proof-of-concept study in chapter 3 and the comparison of contrast-based metrics in chapter 4, this chapter will now transition from a purely quantitative application of image-based algorithms to an experimental approach, allowing for the comparison of quantitative metrics to subjective ratings. Using daylit architectural renderings to collected impressions of daylit in an online survey, this chapter will assess the predictive capability of existing contrast-based algorithms and propose a new model for predicting attributes of visual interest in 2D renderings. Where the previous chapter relied on renderings from simple 'shoebox' geometry models, this chapter will use existing architectural spaces to compare subjective ratings of more detailed scenes. To test the effects of sun position and architectural composition on attribute ratings, the online experiment presented here will use a series of hourly and monthly instances for a range of 9 spaces that vary in composition and spatial complexity.

In the first of two user-based studies in this thesis, the online survey in this chapter was distributed to a global population of subjects. This experiment was designed to test the effects of architectural composition and sun position on subjective responses for contrast, distribution, and excitement ratings and allow for the comparison of existing global and local contrast measures to these ratings. The objectives of this experiment are two-fold: 1) to estimate the impact of sky conditions and architectural spaces on subjective ratings of contrast-related characteristics in daylight composition, and 2) to compare the relationship between these subjective ratings and existing quantitative measurements for contrast. The first objective is to test whether the subjects reach a consensus on their ratings of contrast-based visual effects in architectural spaces and whether these ratings are sensitive to sunlight dynamics (sun positions). The second objective is to compare existing contrast measurements presented in Section 2.4 to subjective ratings to determine whether these quantitative metrics can be used as a model for predicting perceptual responses to daylight composition across a broad population (as opposed to the author's own intuition used in chapters 3 & 4).

5.1 Selection & Modeling of Architectural Case Studies

For this experiment, the authors modelled nine contemporary architectural spaces that display a range of contrast-based visual effects and were selected based on a representative interior view. To select the architectural spaces used in this initial study, the authors referred to the range of interior scenes presented in Chapter 1 and chose those examples that represent a unique typology of

daylight composition within a broader gradient of contrast-based effects (Rockcastle & Andersen, 2013). As can be observed in Figure 5.1, spaces were grouped into three contrast categories. For the high contrast group, the authors selected the Arab World Institute by Jean Nouvel (*arab*), the Zolleverein School by SANAA (*zoll*), and the Serpentine Pavilion by Toyo Ito (*serp*). The medium group contains the Neugebauer House by Richard Meier (*neug*), the Toldeo Glass Museum by SANAA (*toledo*), and the First Unitarian Church by Louis Kahn (*first*), all of which show a mix of direct and indirect daylight conditions. Finally, the low contrast group holds the Poli House by Pezo Von Ellrichshausen (*poli*), the Thermal Baths at Vals by Peter Zumthor (*vals*), and the Menil Gallery by Renzo Piano (*menil*). This group is characterized by more indirect daylight design.

Each of the selected spaces was modelled in Rhinoceros version 5 sr6 based on available building plan and section drawings, and exported to Radiance using the Diva 2.0 toolbar to produce HDR daylight renderings. Furniture and temporary artefacts were intentionally excluded from each space in an attempt to limit non-structural visual obstructions and minimize biases toward programmatic function. That being said, future work may investigate the role of program use in creating bias towards subjective lighting evaluations. To compress HDR images down to an acceptable range for conventional computer screens (0.5 to 200 cd/m^2) i.e. typical for personal tablet, laptop and desktop screens, the *pcond* mapping algorithm was used (Ward, et al., 1997). While tone-mapping produces a range of pixel values acceptable for conventional screen displays, the limitations associated with a compressed range of values are discussed in Section 2.3.4. As Cauwerts discuses in her experimental findings using HDR, LDR, and projection, a compressed range of luminance values can produce an acceptable surrogate for real spaces when testing subjective assessments of contrast and brightness (Cauwerts, 2013). The range of luminance values experienced in a real-world scene far exceed those that may be reproduced by an LDR or even HDR screen, but the advantage of wide-spread survey diffusion by means of conventional computer displays was seen to outweigh the marginal effect of display luminance range in this first user-based study. To assess this assumption further, an experiment conducted by Chamilothori et al. (2016) investigates the effect of different tone-mapping algorithms on subject impressions of digital images projected in immersive virtual reality and real space. This will be discussed in more depth in chapters 9 & 10. In parallel, the 3D user-based study presented in chapter 8 will introduce a method for gathering subjective assessments using tone-mapped images in immersive virtual reality. The jump from 80° x 60° renderings to immersive 360° renderings will provide further insight into the impacts of immersion and the limitations and/or impact of views in assessing daylight qualities related to perception.

5.2 Design of Experiment

In the 2D user-based study presented here, a repetitive 3 x 3 Semi-Latin Square was chosen to compare three variables – spaces, subject group, and sky - while limiting experimental fatigue by showing each subject 9 images, rather than the 27 which are required by a full factorial experiment.

The Semi-Latin Square allows for repetition (in the case of multiple variables within a given group) and nesting (with three architectural examples per sub category of high, medium, and low contrast – nine spaces in total). While the experiment is intended to study the effects of 'space' and 'sky type' on a subject's impression of contrast-based factors within a daylight rendering, there is a risk that subjective assessments of daylight composition repeated under multiple sky conditions will create a bias toward the architecture. To overcome this, a Latin Square experimental design uses three distinct subject groups. Each subject within a group is shown a single rendering of each of the nine spaces, under one of the three possible sun positions. With three distinct groups, we can then compare the effect of architectural composition within groups and the effect of sun position between groups. Figure 5.2 shows this logic. This proposed experimental design allows for a more compact survey and was intended to reduce subject fatigue while testing the effect of sky type between subject groups. More time-intensive surveys increase the drop-out rate amongst participants mid-response and this design allowed for completion at an average rate of 11 minutes (as computed by Survey Gizmo).

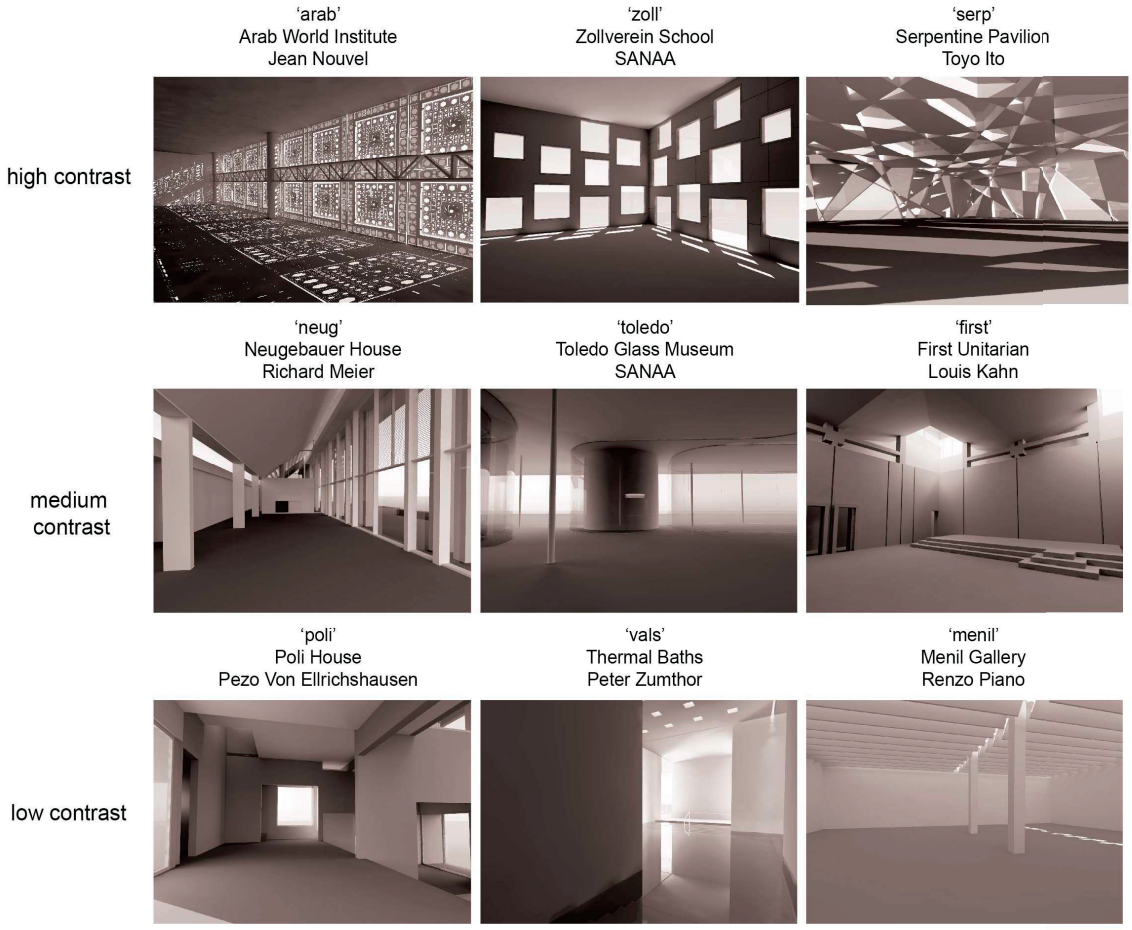


Figure 5.1 The nine architectural spaces selected for this experiment, grouped into high contrast, medium contrast, and low contrast categories.

To select the dates and times for each rendering within the study, half the year was divided (from the winter to summer solstice) into 28 moments, which represent symmetrical daily and monthly instances (Figure 5.3). Each of the nine architectural spaces was then rendered at each of the 28 moments and analyzed in MATLAB R2012b using the RAMMG contrast metric (Rizzi, et al., 2004), which was selected to represent the broader group of neighborhood metrics introduced in Chapter 4. From the assessment of RAMMG contrast across these 28 renderings, three images were selected: the highest (Sky 1), lowest (Sky 3), and mean (Sky 2) contrast composition for each space. The date and time for Sky 1, Sky 2, and Sky 3 therefore vary for each space as the RAMMG contrast assessment selected the highest, lowest, and mean compositions relative to the 28 renderings produced for each space. Based on the mean RAMMG contrast for each architectural space, the 9 spaces were ordered and divided into three sub-groups: high, medium, and low. Both global (Michelson and RMS) and a selection of local contrast metrics (RAMM1, RAMMG, DOG and RSC) were applied to the renderings in this study. DOG measurements are dependent on the radius of center and surround Gaussian filters with a range of possible combinations. To limit this range, a series of radii combinations were selected from r_c = 1 to 4 and r_s = 2 to 8, based on combinations chosen in the experiments conducted by (Tadmor & Tolhurst, 2000) and (Rizzi, et al., 2008) and further tested in chapter 4. Figure 5.3 Shows the results for each space and sky.

As metrics like RAMMG and RSC are applied to a series of levels depending on the original input resolution of the image, each level was also looked at independently. In this study, the original images were 1488 x 1024 pixels. We halved the resolution at each subsequent level, looking at 9 independent image levels for RAMMG (RAMM1, RAMM2, … , RAMM9), and 5-6 levels for RSC, depending on the center and surround radii.

From an initial comparison of these metrics in Figure 5.4, it is already apparent that global contrast metrics such as RMS produce less differentiated predictions of contrast between spaces and sky conditions than local measures RAMMG and DOG, despite the highly varied composition of pixels in each image. To find which of the metrics described above can be used to best predict perceptual visual effects, the contrast results will be applied to each image and then compared to semantic ratings collected from the online experiment.

5.3 Rendering & Tone-Mapping

When using images to collect subjective impressions of daylight related to perceptual factors such as brightness and contrast, it is essential that light levels are accurately captured or generated (in the case of renderings) and displayed with as broad a range of luminance levels as possible using proper tone-mapping algorithms calibrated for the specific display. In controlled laboratory

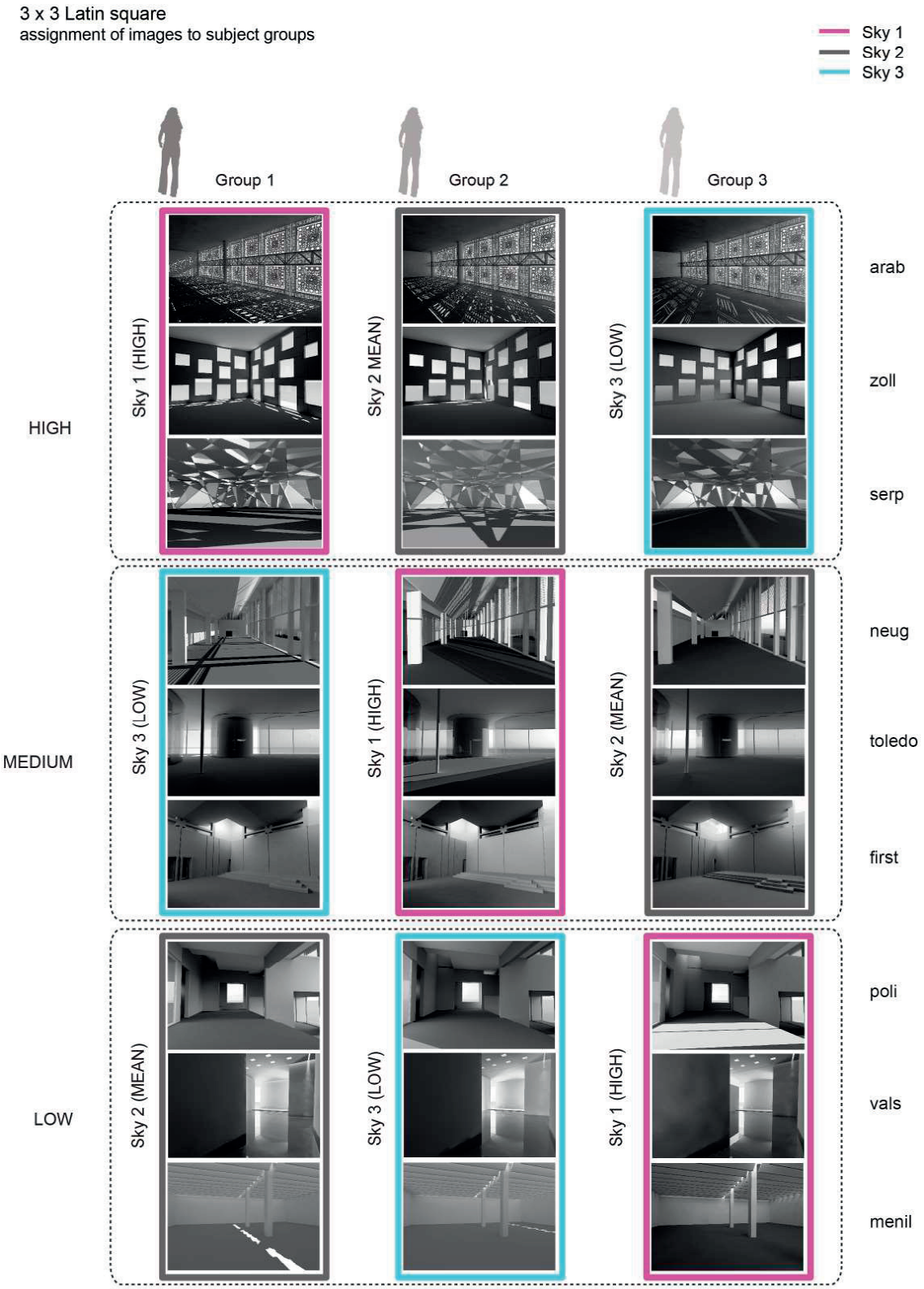


Figure 5.2 Shows the presentation of renderings to each of the subject groups based on the 3 x 3 Latin Square design of experiment. Each group sees all nine architectural spaces under at least one of the three contrast conditions (high, mean, and low) selected by the RAMMG metric.

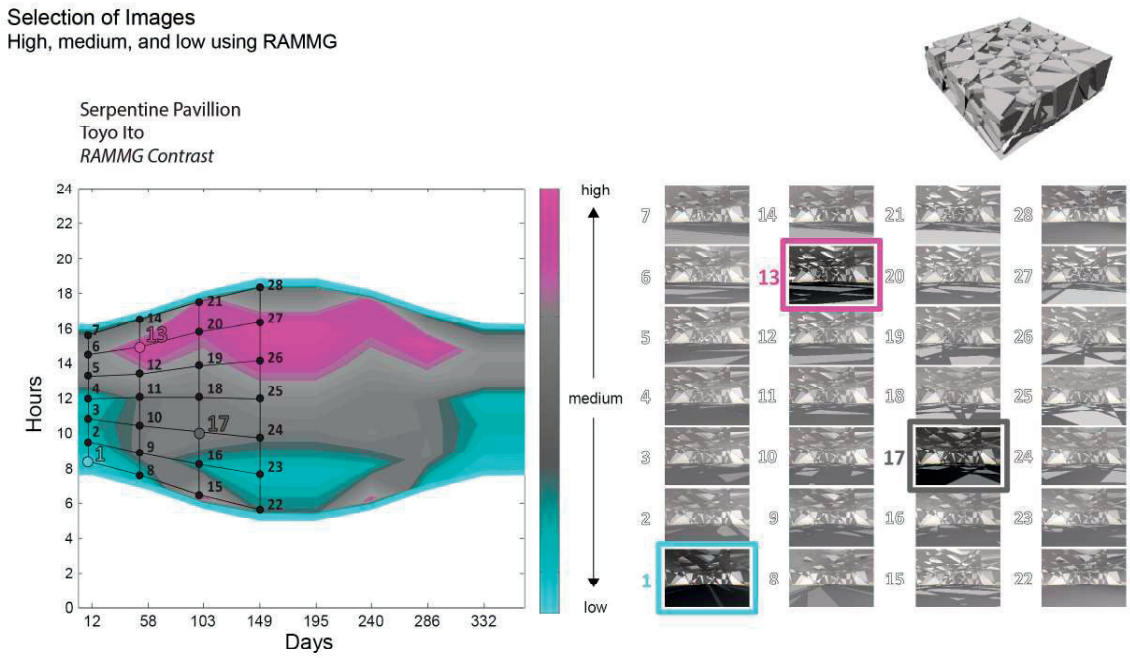


Figure 5.3 A temporal map of RAMMG results for 28 semi-annual instances showing high, medium, and low instances selected by the algorithm. These instances were then used in the latin square 3 x 3 design of experiment.

experiments, tone-mapped HDR images have been displayed to subjects using 2D or 3D projection, HDR displays, and conventional low dynamic range (LDR) displays. While there are now backlit HDR screens on the market which can display a wide range of luminance values, Cauwerts (2013) found that conventional LDR displays performed as well as HDR displays for many image attributes. In this study, there were some differences in ratings of pleasantness, distribution, and spaciousness between real world scenes and image displays, however it was concluded that conventional LDR displays of 200 cd/m^2 (with images tone mapped to 256 distinct luminance levels) could be used as a surrogate for real world spaces to conduct subjective assessments involving contrast and brightness (Cauwerts, 2013). For attributes related to distribution, she found that LDR displays even out-performed HDR displays. In (2013, 2015), Villa & Labayrade developed an online protocol to limit the impacts of uncontrolled experimental conditions (i.e. screen resolution, brightness, background, etc.) on the assessment of digital images for lighting quality research. Their study found that significant effects could be identified with 40 subjects despite systematic error due to uncontrolled conditions. Where controlled experiments are time-consuming, an online protocol allows for more widely distributed experiments.

There are a number of methods for creating and displaying interior images to assess subjective qualities of daylight (Section 2.3), each of which has its own set of advantages and limitations. While real spaces produce the most accurate impression of light for subjective assessment, experimental conditions in a real space are limited by the physical sky available on site at the time

and it is difficult to compare a range of spatial configurations or temporal conditions in an efficient manner. HDR renderings allow for a broad range of spatial and temporal lighting conditions, but are limited by the luminance output of the display device and must use tone-mapping algorithms to achieve an acceptable range for the specific screen and luminance range.

Due to the computational intensity of Radiance, the 28 renderings used to select highest, lowest, and mean contrast compositions were first rendered using low-quality ray-tracing and image parameters. After each set of 28 low-quality renderings was complete, RAMMG was used to select three images for each space (the high, mean, and low relative to all images in the set). These images were then re-rendered using high-quality parameters (-ps 2 -pt .05 -pj 0 -dj 0 -ds .15 -dt .05 -dc .75 -dr 3 -dp 512 -st .15 -ab 5 -aa .1 -ar 512 -ad 4096 -as 2048 -lr 8 -lw .005 -vv 60 -vh 80) with increased pixel resolution (1488 x 1024) and tone-mapped using the *pcond* operator (Ward et al., 1997). As a result, there is some variation in RAMMG values between the low quality images, which were used to select moments, and the high quality renderings, which were shown to subjects and analyzed with each of the available contrast metrics. Not surprisingly, the increase in computational intensity had an effect on RAMMG in both indirect and direct lighting conditions. This discrepancy does not, however, affect the comparison between quantitative calculations and subjective ratings, as the low-quality images were only used to select annual moments and only high quality images were shown to subjects at random.

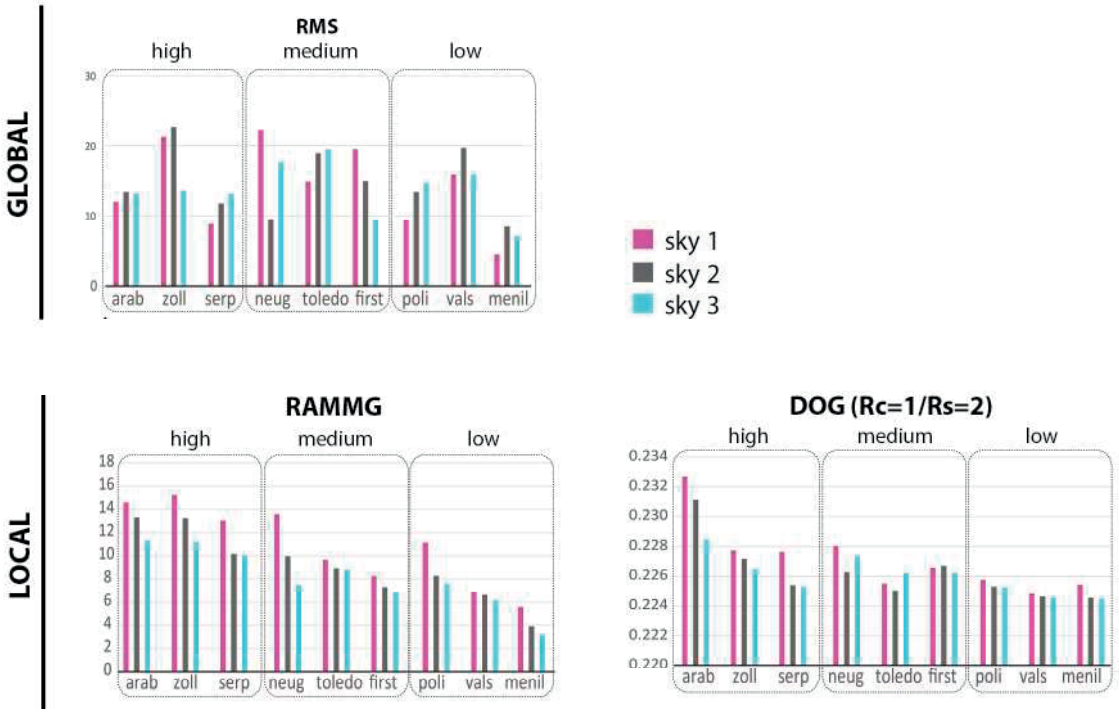


Figure 5.4 Results for each rendering (grouped by space and sky condition) for a selection of global and local contrast algorithms RMS, RAMMG, and $DOG_{RC=1, RS=2}$

5.4 Semantic-Differential Scales

For each image, subjects were asked to rate the daylight composition using the following seven point semantic differential scales:

low contrast – high contrast
uniform – non-uniform
unvaried – varied
diffuse – direct
simple – complex
calming – exciting
subdued – stimulating.

Flynn first introduced the use of semantic differential scales to gather subjective assessments of daylight quality in terms of visual clarity, spaciousness, evaluation, relaxation, social prominence, complexity, modifying influence, and spatial modifiers (Flynn, et al., 1979). Numerous studies thereafter have employed the use of these scales to conduct daylight quality research in real spaces and simulated or photographed views (Newsham, et al., 2002, 2005; Vogels, 2008; Demers, 2007). For the proposed study, the authors have focused on bi-polar semantic differential scales associated with complexity and spatial modifiers as well as visual interest. The selection of a seven-point scale was motivated by the use of bi-polar pairs to allow for a neutral rating at 4 and three ratings in either direction of the scale. In our analysis, we clustered subject ratings 1-3 and ratings 5-7 as being 'non-neutral' (with 4 as being neutral) and visualize the distribution of ratings on one side of the bi-polar scale or the other. Due to the limitations of Survey Gizmo in recording ordinal subject responses through the online survey format, bubble ratings were used instead of line scales. Pair-wise comparison was also considered, but rejected due to the number of rating pairs selected in our survey (seven pairs per image).

5.5 Pilot Study

To test the 3 x 3 Latin-Square design before a full initiation of the online survey, a small pilot study was conducted on printed paper (Rockcastle & Andersen, 2015). This pilot was composed of 5 female and 4 male researchers in building performance, with varying competencies in architectural design, computational, and civil engineering. Each subject was randomly assigned a survey (from three possible groups) and asked to rate each rendering (shown in Appendix ii) for six semantic differential ratings. Note that in this pilot, the scale 'sedating-stimulating' was used instead of 'subdued-stimulating' and then later changed for the online survey due to comments from the subjects that this word pair was confusing. Using this limited sample size of three subjects in each group, the authors conducted a 3-way ANOVA to test the effects of space, sky type, and group factors on each of the rating pairs. ANOVA results show that sky type had a significant effect on ratings of contrast, variation, complexity, excitement, and stimulation

while space (the selected architecture and view) had a significant effect on ratings of complexity, excitement, and stimulation. None of the three factors were shown to have a significant effect on ratings of uniformity, and were weaker for variation than the other four rating scales.

Based on the ANOVA results, we can see that the null hypothesis is always true for uniformity ratings under these experimental conditions, but can be rejected for some factors in the remaining five ratings, which show significant effects from at least one of the three factors. The effect of subject is not shown to be significant for any of the semantic rating scales, indicating that the experimental design used in this study is working, despite the limited population size.

To assess potential redundancies in the semantic rating scales selected for this study, Pearson Correlation Coefficient values (r) were computed between responses for each rating pair. 'sedating - stimulating' and 'calming exciting' ratings were found to be strongly correlated, with $r > 80\%$. While 'uniform – non-uniform' and unvaried – varied' had a $r > 50\%$, no other rating pairs showed significant correlation.

To understand the relationship between semantic ratings and quantitative measurements, a correlation analysis was conducted between mean subject responses per space and sky condition for each of the 9 subjects. In Figure 5.5, each of the four quantitative metrics were fit against each of the six semantic differential ratings to find which, if any, can serve as a prediction model for subjective assessments. Data points represent a mean rating for each of the 3 subjects who rated the same space and sky condition with error bars showing the standard deviation between subject ratings. It is important to note that PCC values are shown for the linear regression fit through *mean subject responses*. Most significant in this analysis are the r values $\geq 60\%$ for RAMMG – 'calming - exciting' and RAMMG – 'sedating – stimulating.' For all of the semantic rating scales, including 'low contrast – high contrast', global metrics such as RMS and Michelson show no significant trend.

What is perhaps most interesting about these initial results it that ratings of 'calming-exciting' and 'sedating - stimulating' are more strongly correlated to pyramidal sub-sampled metrics such as RAMMG than subjective ratings of 'contrast' for which the metrics were designed. This finding raises some interesting questions about subjective interpretations of the word 'contrast' and whether subjects are responding to micro or macro compositional effects within the images, which will be explored later in this chapter through the analysis of data in the Online survey.

5.6 Online Survey

The online survey designed for this experiment was created using Survey Gizmo (http://www.surveygizmo.com/) with a branch logic which allowed for random group assignment upon initiation of the survey. Each group was asked to respond to some basic demographic questions regarding geographic location and profession and then shown the nine architectural spaces at

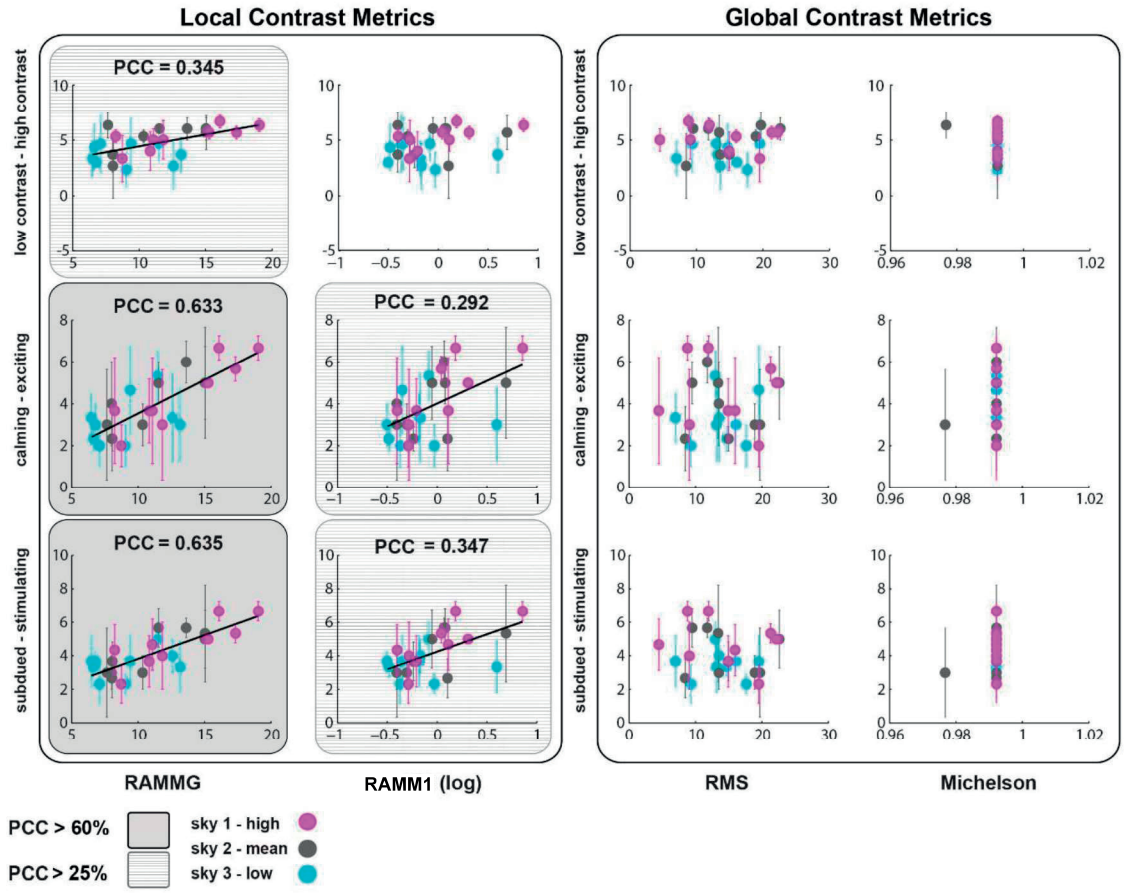


Figure 5.5 Results from the pilot study showing Pearson Correlation Coefficient values (PCC) mean subject responses for 3 rating scales and a selection of local and global contrast metrics. Pearson values over 0.60 are highlighted in dark grey (calming - exciting/subdued - stimulating and RAMMG).

random, under one of the three possible sky conditions. For example, Group 1 was shown three high contrast spaces under Sky 1, three medium contrast spaces under Sky 3, and the three low contrast spaces under Sky 2 (Figure 5.6). While we asked the participants not to initiate the survey on a smartphone, we allowed tablet, laptop, and desktop computers. Participants were asked to turn the brightness on their device to maximum, to ensure that the widest possible pixel range was observed.

5.6.1 Data Management

In total, there were 334 subjects who initiated the survey with 200 complete responses and 134 partially complete ones, which were discarded before the analysis. Interestingly, we did see a significant effect on responses from those subjects using tablets. These subjects (4.5%) were discarded as it was believed that this effect could be due to the smaller screen size (which forced subjects to manually zoom in and out to fully view each image) or the default button format which was automatically adjusted by Survey Gizmo on the tablet version, switching from a 7-point rating

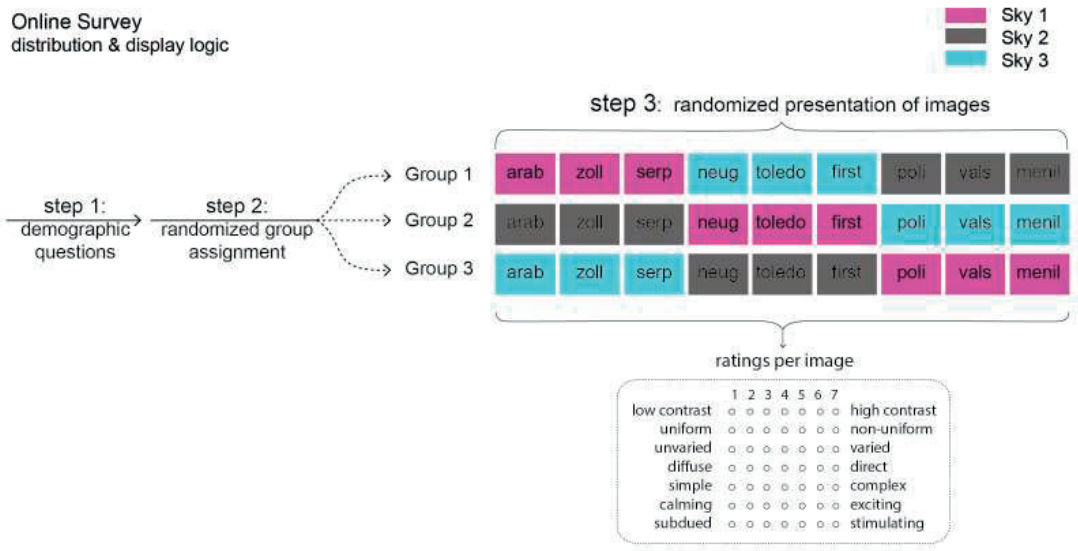


Figure 5.6 hows the online survey logic, where subjects are first introduced to basic demographic questions, after which they are randomly sorted into one of the three subject groups and asked to rate the images as they are presented in random order.

scale with selection bubbles to a sliding scale. There was no significant effect observed between subjects using a laptop or a desktop computer. Of the remaining 175 subjects, 96% selected their English language capacity as professional, bilingual, or native, with the remaining 4% responding with elementary or limited working proficiency. These subjects were also discarded. From the remaining 168 subjects, 64% were designers (architecture, landscape, urban, or interior), 36 % non-designers, with 55% reporting their expertise in lighting design as competent, proficient, or expert, and the remaining 45% claiming novice or beginner expertise. One subject was excluded from the analysis because 73% of responses were neutral. We normalized the responses (from 1 to 7) for five other subjects, as they did not use either extreme on the rating scale. The remaining 167 subjects were evenly distributed among the three groups (Group 1: 55, Group 2: 56, Group 3: 56).

5.6.2 Data Analysis

To test the significance of experimental factors on the data from each rating pair collected in the experiment, a 3-way ANOVA was used to test the effects of sky, space, and subject group. As the residuals for each rating pair was not normally distributed, a post-hoc analysis was conducted using the non-parametric Kruskal-Wallis test to determine the significance of each group within the factor under consideration. To analyze the relationship between subject ratings and existing contrast metrics, the Pearson's correlation coefficient was used. Using this coefficient, those combinations of rating-pair and contrast measurement with $r > 0.69$ ($p<0.0001$) were selected. An ordered logit model was then applied to fit the subject ratings to selected contrast measures, as the subjective ratings are ordinal response scales. For a given category c on the semantic differential

scale (1-7), the probability of each outcome is expressed as a non-linear function of the predictor x

$$\ln \theta_j = a_j - bx, \forall j = i, \dots, c \ . \tag{5.1}$$

where b is the corresponding effect parameter, a the cutoff point and θ_j the ratio of the probability that a response belongs to a category with a value less or equal to category j and the probability that this response belongs to a category with a value greater than category j.

$$\theta_j = \frac{p(y \leq j)}{p(y > j)} \ . \tag{5.2}$$

The cumulative predictive probabilities from the logit model for each rating are calculated as

$$P(y \leq j) = \frac{1}{1 + e^{a_j - bx}} \ . \tag{5.3}$$

5.7 Results of 2D User-Based Study

The description of results will be introduced in three parts. Firstly, the distribution of subject responses will be explored for each semantic rating pair collected during the Online experiment. Secondly, the significance of experimental factors on each rating pair will be analyzed in order to determine whether the effects were significant. Thirdly, the relation between subject ratings and quantitative contrast measures will be explored and a model will be introduced for predicting visual interest in daylit architectural renderings.

5.7.1 Analysis of Subjective Data

Figure 5.7 shows stacked bar plots with the distribution of subject responses for each level of the seven-point rating scale under each image presented in the experiment (organized by architectural space and sky type). Subject responses are sorted and clustered into gradients by tone gradient, with cyan demarcating ratings that fall on the left side of the scale (1-3) and magenta for ratings that fall on the right (5-7). Ratings in the middle (4), indicating neutrality or indifference toward either side of the semantic scale are shown as white void. For example, ratings 1, 2 & 3 on the contrast spectrum are shown in a gradient of cyan, with 1 being the lowest contrast rating. A rating of 4 was perceived as neither low nor high in contrast and shown as white void. Ratings 5, 6 & 7 are shown in magenta, with 7 as the highest possible contrast rating. The most frequent responses are displayed for each rating as percentage of subjects on the respective side of the rating scale. The dashed line shows where the median falls on the scale within each semantic pair. Using this graphic representation, we can see the relative distribution of subject responses for all ratings. Images with dominant bar length of magenta or cyan (over 50%) can be considered partial toward one side of the rating scale. Images with large areas of void can be considered neutral between each end of the rating scale, while images with an even distribution of cyan and magenta lacked consensus between subjects.

Figure 5.7 Stacked bar plot showing distribution of semantic differential scores for each pair-wise rating and image. The images are ordered row-wise for sky type and column-wise based on the average percentages in categories 5-7 for all ratings and sky types per space.

There is an obvious effect of sky type in some, but not all spaces – specifically those that see the largest variation in daylight composition due to sky type. *Zoll*, *neug*, and *vals* experience a noticeable shift in ratings between sky type 1, 2 and 3. The spaces with strongest subject consensus toward the cyan end of the rating scale (low contrast, uniform, unvaried, diffuse, simple, calming, subdued) are *first* and *menil*, while the magenta side of the rating scale (high contrast, non-uniform, varied, direct, complex, exciting, stimulating) was dominated by *serp* and *arab*.

While all rating scales were found to be significantly correlated to each other, the percentage of subjects that responded with ratings 5-7 for excitement and stimulation were the most highly correlated (r=0.98).

To isolate the sky parameter and visualize its impact on the distribution of responses for a selection of rating pairs, Figure 5.8 shows an overlay of distribution plots for sky 1 and sky 3. If we overlay the distribution of responses for sky type 1 and sky type 3 for low contrast-high contrast, calming-exciting, subdued-stimulating, and diffuse-direct, we can see a perceptible shift in ratings between those images rendered under the two sky types. If we consider the median response for each space under the two sky types, we can see that for most all spaces and rating pairs, the median rating is slightly lower for sky type 1 than sky type 3. While some of the spaces achieve a smaller differential between median ratings for the two sky types, it is clear from Figure 5.8 that spaces 2, 4, and 9 produce the most differentiated median response. While the median response can be useful, we can observe an even more noticeable shift in overall response distribution. Sky 1 (magenta) shows more responses in the far right side of the spectrum (5-7 in high contrast, exciting, stimulating, and direct) with sky 3 showing fewer responses in the high end of these rating scales and more towards the center and left. To further explore this trend quantitatively, Section 5.7.2 will look at a pair-wise comparison to establish the significance of sky conditions and space factors on ratings for each semantic-pair.

5.7.2 Effects of Experiment

The significance of experimental factors was evaluated using a 3-way ANOVA to test the effects of subject group, space, and sky type on each rating scale. While the ANOVA revealed a significant effect of both space and sky factors for all rating scales (p<0.01), the residuals were not normally distributed. A post-hoc analysis was conducted using the non-parametric Kruskal-Wallis test, to study pair-wise comparisons between levels in each factor under consideration within each rating. This test was run for both sky type and space group on each of the semantic rating pairs. This test revealed the effect of sky was significant on subject responses to all rating scales (p<0.01), except unvaried-varied. Figure 5.9 illustrates a pair-wise comparison between Sky 1 and Sky 3 which shows a significant effect (p<0.01) on ratings of contrast, uniformity, direct, complexity, excitement, and stimulation. Ratings of excitement and stimulation also show a significant effect (p<0.01) between Sky 1 and Sky 2, which suggest that these ratings were more sensitive to the range of sky types presented in this experiment. To test the effect of space, we grouped the examples into high, medium, and low based on the ordering in Figure 5.6. In this test, there was a significant effect of space between all groups in the factor (p<0.001) for all rating pairs. This

Figure 5.8 Overlaid distribution plots for sky 1 (magenta) and sky 3 (cyan) showing a shift in both the median response and the distribution of responses for most spaces and rating pairs.

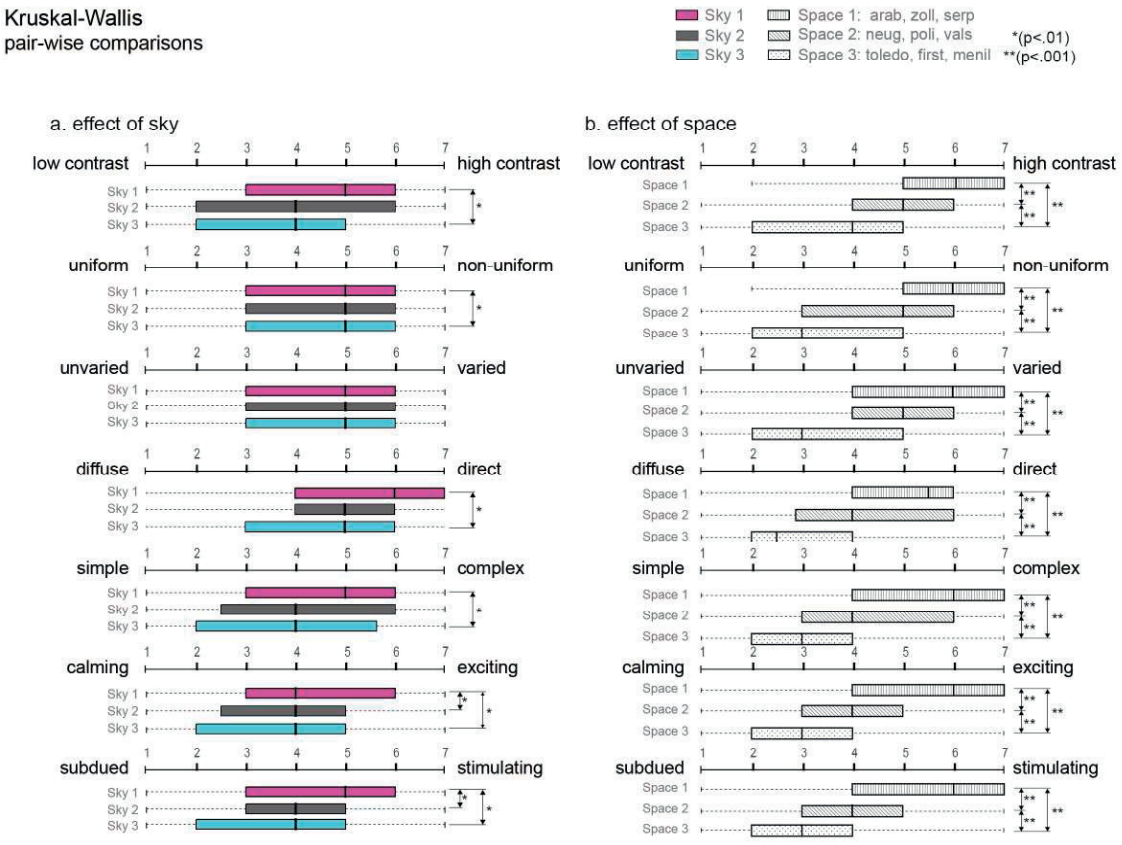


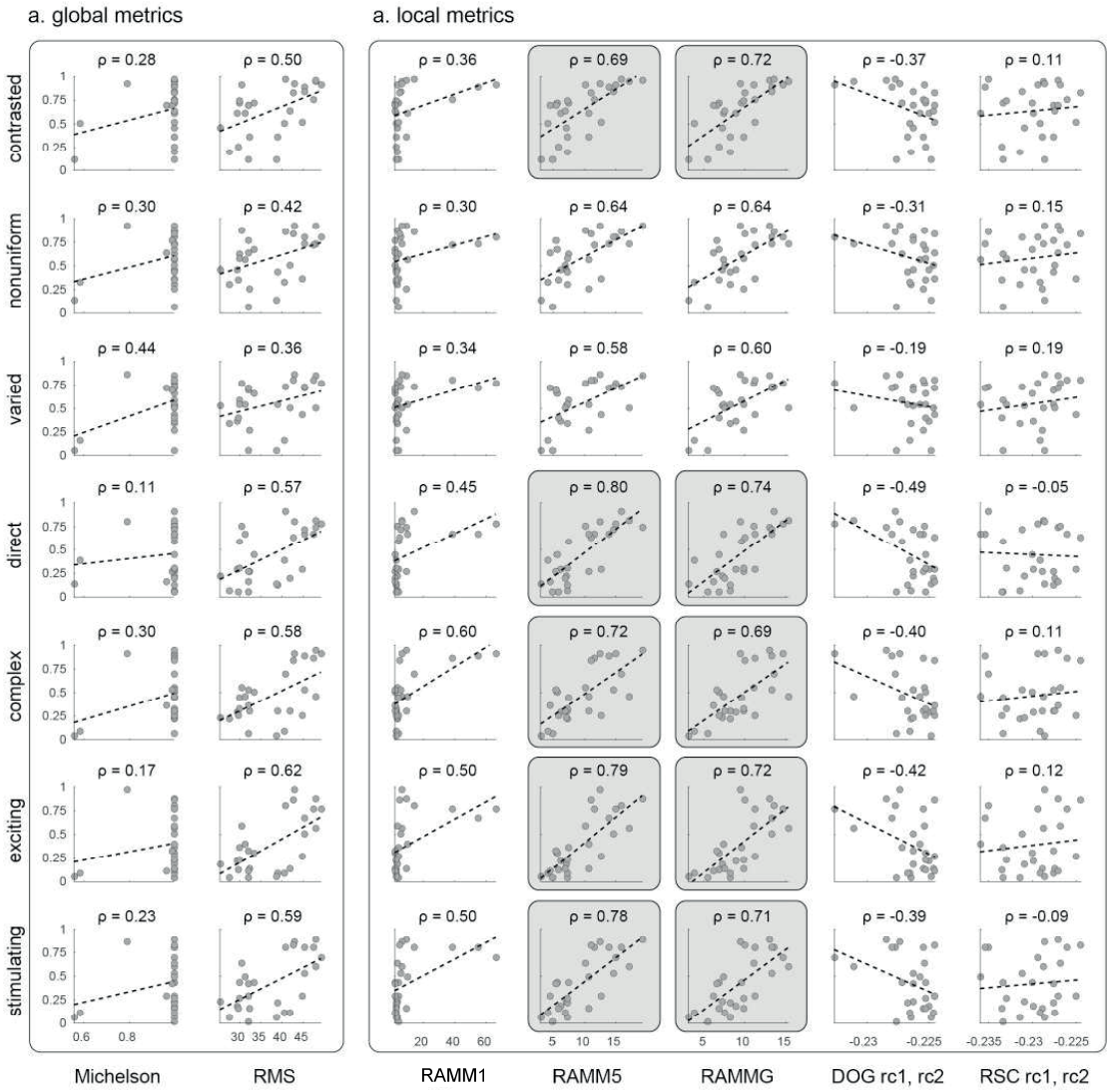
Figure 5.9 Shows box plots with a Kruskal-wallis pair-wise comparison for the effect (a) between sky type 1,2 and 3 and (b) between space groups 1,2 and 3 within each rating scale.

post-hoc analysis, in addition to the 3-way ANOVA would indicate that both space and sky factors had a significant effect on subject responses for all rating scales, except for variation. While the effect of space was significant on subject responses for all rating scales, the effect of sky was most significant on subject ratings of excitement and stimulation.

5.7.3 Subject Ratings & 2D Contrast Measures

To relate subject responses for each rating pair to the contrast measures described in Section 2.2, quantitative contrast results were compared to subject responses for each rating and image presented in the experiment. Figure 5.10 shows correlation values between the percentage of subjects that responded with ratings 5-7 for each image on each rating scale and the output for global and local contrast metrics. The Pearson correlation coefficient is shown for each rating/ metric combination, with highlighted combinations for significance $p<0.0001$.

In this experiment, ratings of contrast were found to have the strongest dependence with RAMMG ($r = 0.72$), but the RAMM of one image resolution level (RAMM5 - 64 x 93 pixels) in particular had the highest predictive fit to several other rating pairs. RAMM5 achieved the strongest linear



	Michelson	RMS	RAMM1	RAMM5	RAMMG	DOG $r_{c=1}\, r_{s=2}$	RSC $r_{c=1}\, r_{s=2}$
contrast	0.28	0.50	0.36	**0.69***	**0.72***	0.37	0.11
uniformity	0.30	0.42	0.30	0.64	0.64	0.31	0.15
variation	0.44	0.36	0.34	0.58	0.60	0.19	0.19
direct	0.11	0.57	0.45	**0.80***	**0.74***	0.49	0.05
complex	0.30	0.58	0.60	**0.72***	**0.69***	0.40	0.11
exciting	0.17	0.62	0.50	**0.79***	**0.72***	0.42	0.12
stimulating	0.23	0.59	0.50	**0.78***	**0.71***	0.39	0.09

*Rating pair and contrast measurements with $\rho \geq 0.70$ ($p < 0.0001$) were considered most significant.

Figure 5.10 Linear regression values between the percentage of subjects that responded with ratings 5-7 for each image on each rating scale and the output for global and local contrast metrics. Pearson values over 0.70 are highlighted in grey. Significant combinations are highlighted with an asteric in the table.

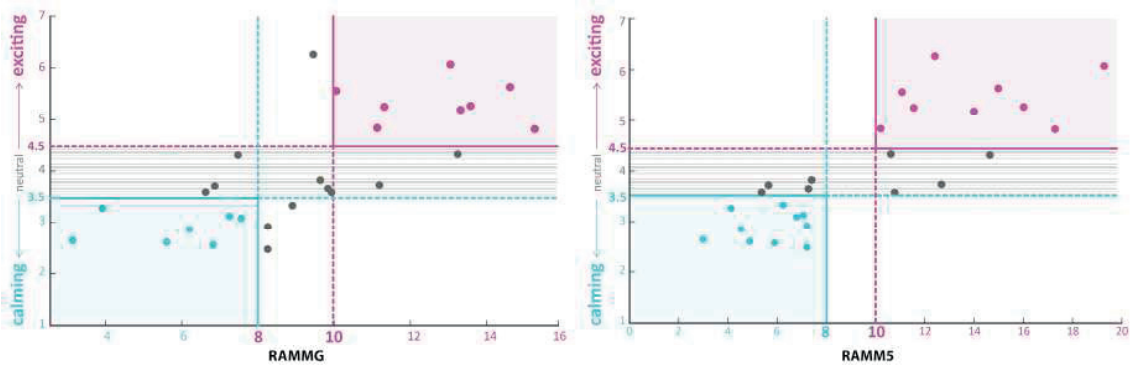


Figure 5.11 Shows a linear fit between RAMMG/RAMM5 values and median subject responses for calming - exciting per image considered in the online study. Lines drawn in magenta and cyan mark visible thresholds under and over which values appear to cluster on one end of the spectrum or another.

dependence to ratings of direct (r = 0.80), complex (r = 0.72), exciting (r = 0.79), and stimulating (r = 0.78). This finding suggests that the luminance resolution or radius of solid viewing angle may be related to subject impressions and tied to mid-levels of the visual system related to feature inference). Using Pearson correlation to pre-select contrast metrics as possible predictors of visual interest, we selected RAMM5, hereafter referred to as 'Modified Spatial Contrast' to study in more depth. RAMM5 or modified spatial contrast (mSC_5) is defined as,

$$mSC_5 = \frac{1}{WH}\sum_{i=1}^{W_5}\sum_{J=1}^{H_5} c_{i,j} \tag{5.4}$$

where $W_5 = W_{5-1}/2$ and $H_5 = H_{5-1}/2$ are the width and height of the image at level 5 (halved in each subsequent level from a starting resolution of 1488 x 1024 (W x H) and $c_{i,j}$ is the contrast of each pixel (as calculated in Equations 2.7 - 2.8).

5.7.4 Excitement Prediction Using a Logit Model

If we plot RAMMG and mSC_5 predictions against the median subject response for calming - exciting per image, as shown in Figure 5.11, we can see a moderate linear relationship. As RAMMG and mSC_5 values increase, so too do the subject ratings, crossing the neutral threshold around 8 and the exciting threshold around 10. This begins to suggest a set of thresholds under which ratings could be described in the calming spectrum and over which ratings can be considered in the exciting spectrum, but a linear fit appears weak on either side of the thresholds. Given the ordinal nature of subject responses, we care less about a linear response and rather more about a distribution of probable responses.

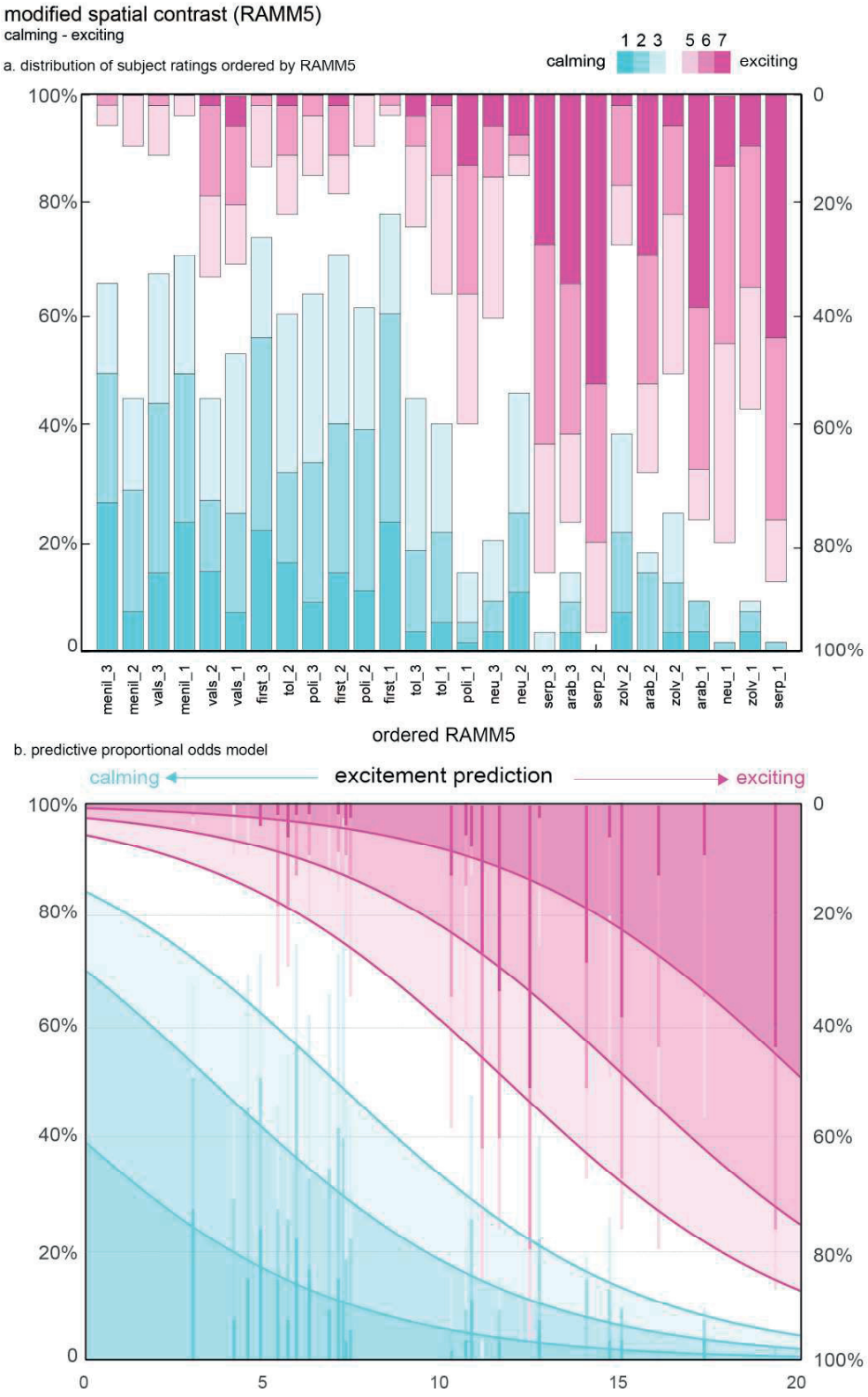


Figure 5.12 a) Distribution of subject ratings for calming – exciting per image ordered by RAMM5 / mSC_5 and (b) predictive proportional odds model where the bars show subject responses (binned into cyan for ratings 1-3 and magenta for ratings 5-7) for the 27 images located on the x axis at the calculated RAMM5. The curved lines show the predictive proportional odds through categories 1-6, where category 7 reaches 100%. The deviance of this fit was 9.36.

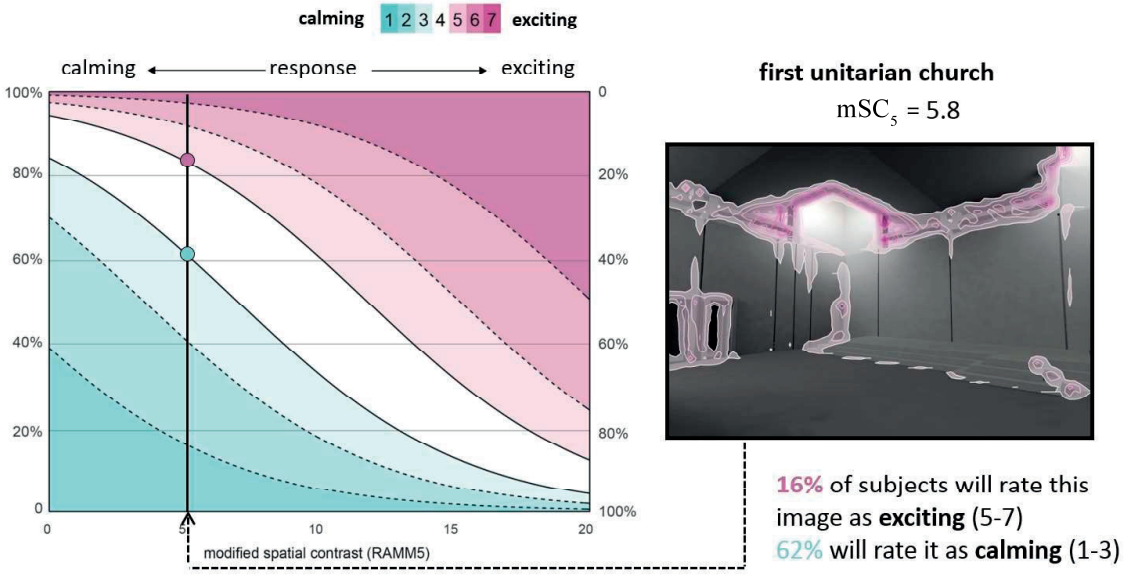


Figure 5.13 Logit model showing the percentage of subjects that would rate the chosen rendering of the First Unitarian Church as exciting (16%) versus calming (62%) using mSC_5.

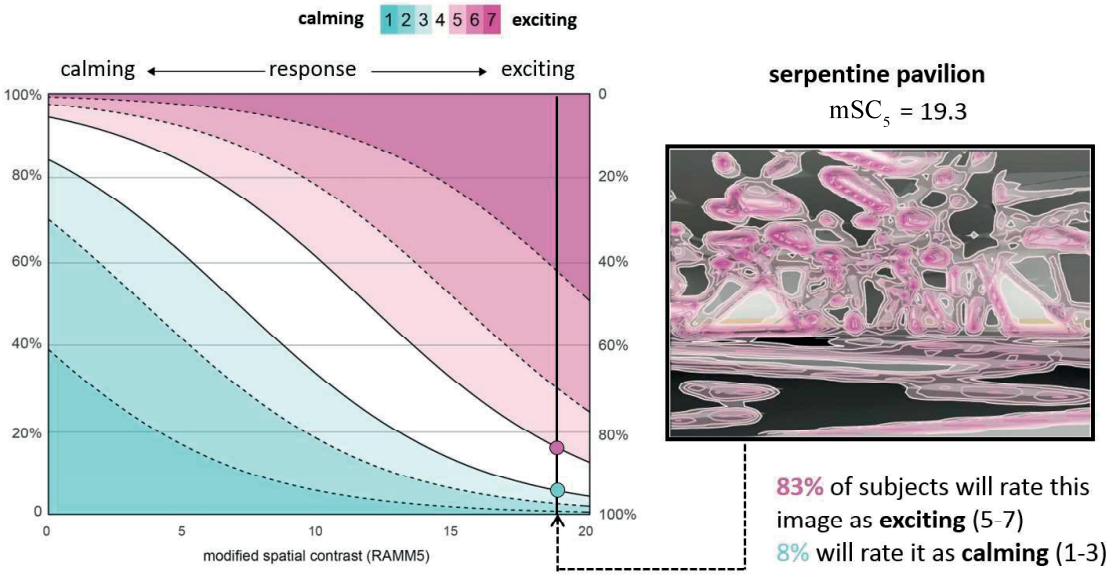


Figure 5.14 Logit model showing the percentage of subjects that would rate the chosen rendering of the Serpentine Pavillion as exciting (83%) versus calming (8%) using mSC_5.

As the subject responses in this experiment were gathered using an ordinal bi-polar rating scale (1-7), the authors selected a probabilistic rather than definitive or threshold-driven model for fitting quantitative and qualitative responses. As was be observed in Figure 5.8, the impact of sky condition on the *distribution* of responses was often more pronounced than the shift in the *median* response. Existing illumination metrics rely on target or threshold illuminance to describe light levels that promote optimal visual acuity for a given task. In daylight glare probability (DGP), discomfort is calculated as a percentage of people that would rate a given glare source as either disturbing or intolerable. This model was established by grouping equal sample sizes into 'classes' from the total number of cases and evaluating the percentage of disturbed subjects in each class (Wienold & Christoffersen, 2006). Where illumination is relatively objective, glare prediction must account for a range of subject responses to any given light condition. Unlike glare prediction, however, that relies on a uni-polar scale (from imperceptible to intolerable), the rating pairs used in this experiment are bi-polar (i.e. calming – exciting). That is to say that, unlike glare, neither side of the any rating pair is worse than the other.

Existing studies that have linked global contrast measures such as RMS to preference factors have found weak linear correlations and have not been robust enough to propose a predictive model for visual interest. The modified spatial contrast metric (mSC_5/RAMM5) and data from this experiment suggest that such a predictive model could exist and could indeed be robust enough to differentiate between architectural spaces that vary in daylight composition. An ordered logit model has been applied to fit the subjective ratings to mSC_5 (predictor x) in equation 5.1 using ordered logistic regression. Figure 5.12(a) shows the distribution of subject ratings ordered by mSC_5 and application of a predictive proportional odds model to subject ratings of calming – exciting in Figure 5.12(b). The plot was obtained by using equation 5.3 and the fitted parameters a = [-0.45; 0.84; 1.67; 2.82; 3.64; 4.81] and b = 0.24. This fit resulted in a deviance of 9.36 and was the lowest of any logistic regression analyses between mSC_5 and the 7 rating pairs. To evaluate the goodness of fit for this model, we also looked at the significance of each cumulative response (1-7). At level 3, the odds of achieving ratings 1-3 are significant at $p<0.05$.

When we group semantic differential ratings into two categories of calming (1-3) and exciting (5-7) as shown in Figure 5.12(a), Figure 5.12(b) shows that a mSC_5 of 13 (or more) triggers responses of excitement (ratings of 5, 6, or 7) for 63% of subjects, whereas a mSC_5 of 5 (or less) produces responses of calming (ratings of 1, 2, or 3) in 59% of subjects. While we cannot and should not define one threshold or target value to predict perceptual impressions of excitement given the subjective nature of such responses, this probabilistic model provides a first objective predictor for visual interest in daylit architecture.

Figures 5.13 and 5.14 show an overlay of mSC_5 algorithms on two renderings; the First Unitarian Church under sky 3 and the Serpentine Pavilion under sky 1, respectively. For the First Unitarian Church, the mSC_5 was computed at 6.8. Using the logit model described above, the subjective prediction is dominated by 62% calming (ratings 1-3), with 22% neutral (rating of 4) and 16%

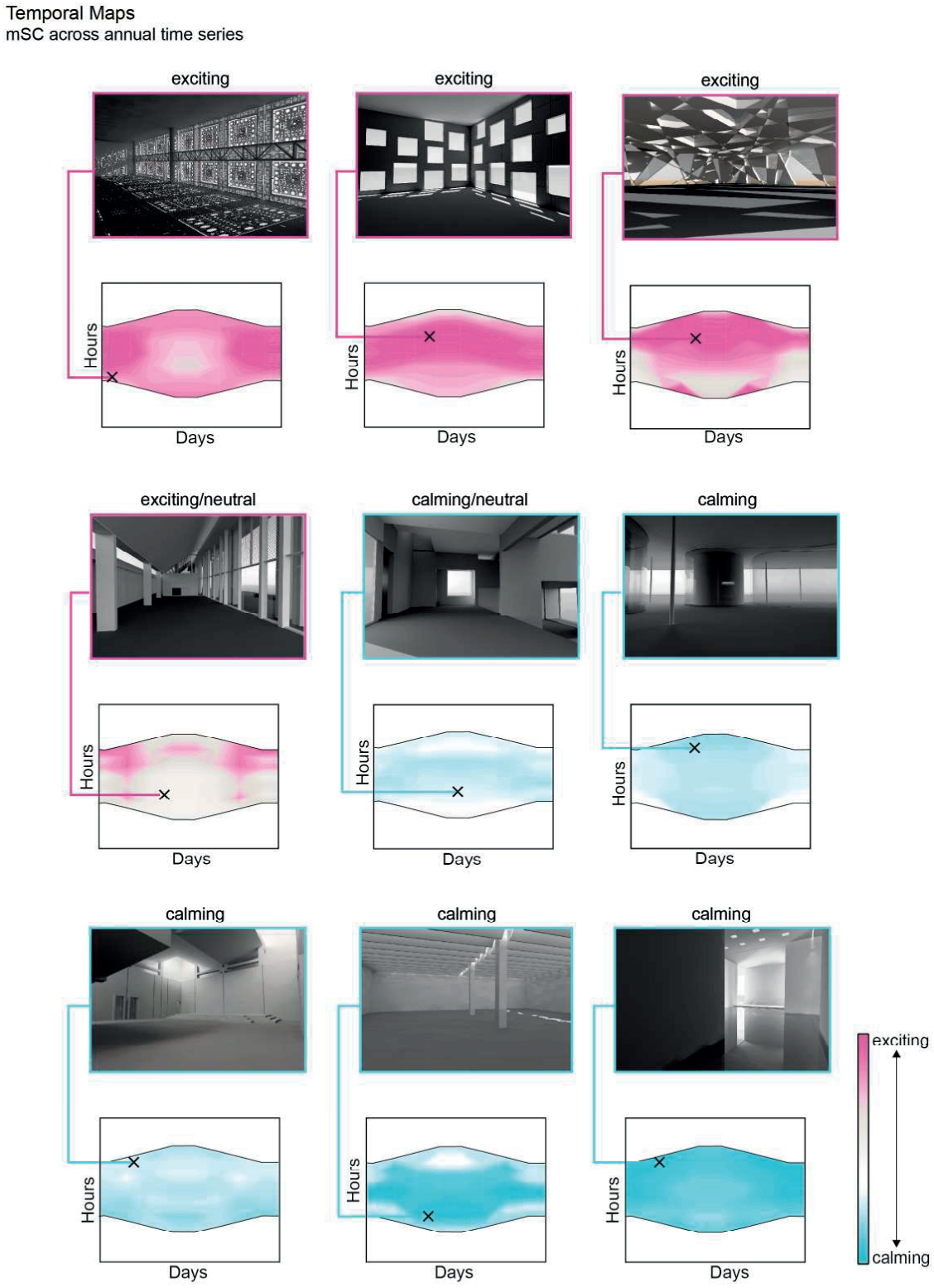


Figure 5.15 Temporal maps showing mSC_5 predictions for the time-series of renderings for each space. Colors represent instances of calm (in blue) to exciting (in pink), with neutral ratings shown in white.

exciting (ratings 5-7). For the Serpentine Pavilion, the mSC_5 was computed at 19.3 with the logit model predicting 83% exciting (ratings 5-7), 9% neutral (rating of 4) and only 8% calming (ratings 1-3).

This instantaneous application of mSC_5 to a single 2D rendering is useful in assessing a fixed view position and gaze direction under an established sky condition. As the algorithm highlights areas of the image where local mSC_5 values are higher, the authors hypothesize that these regions may predict a higher probability of visual interest, an assertion that will be further explored in the immersive 3D user study in Chapter 8 with the integration of head tracking data from the virtual reality headset.

A sceptic to this quantitative approach could argue that we are capable of assessing subjective impressions of daylight in a single rendering without the use of an algorithm, but the reality is that variations in individual responses make it difficult for any one person to objectively predict emotional responses for a larger population of people. The predictive model arising from the data analysis in this chapter proposes the first ever objective measure for predicting dynamic emotional impacts of daylight using simulated images. Moving beyond a single instance of time, this approach allows us to evaluate a series of hourly and/or annual renderings to see how these predictions vary over time under diverse sun positions and sky conditions (explored in more depth in the following chapters). This dynamic time-sensitive prediction would be much more valuable to architects and lighting designers who may not be capable of anticipating the temporally-induced impacts of sun and across a complex architectural composition.

Figure 5.15 shows the application of modified spatial contrast to each of the spaces included in the experiment, rendered across 56 annual instances and mapped temporally to show an annual prediction for impressions of calming and exciting. In this figure, results are ordered by the average modified spatial contrast across all 56 annual instances with values in magenta showing point-in-time predictions of excitement and cyan showing predictions of calming. As the experimental results have shown, there are modified spatial contrast values which can be identified as neutral, triggering consensus in neither impressions of calming or exciting, but for spaces where a large degree of variability occurs, this annual prediction can illustrate specific times of the day and year when subjects are likely to identify the space on one side of the spectrum or the other. The darker magenta values represent a stronger probability that exciting daylight compositions will be observed at that moment.

In the top row of Figure 5.15, we can see that all three spaces; ‘arab’, ‘zoll’, and ‘serp’ vary between exciting and neutral, with no instances achieving a calming prediction. What is interesting about these temporal maps is the distribution of predicted values for excitement, which vary depending on the space and view direction. Predictions of excitement are higher in the winter months for ‘serp,’ when the sun angle is lower and penetrates the screen on the façade, casting a complex composition of sunlight and shadow on the interior.

In the middle row of Figure 5.15, we can see a higher frequency of neutral responses. Peaks in the exciting spectrum are achieved in the morning and late afternoon during winter months in the Neugebauer house when sun is presumably lower and creates more intensive shadow patterns within the field-of-view. The other two spaces are predicted to be slightly calming or neutral for the entire year.

The bottom row of Figure 5.15, especially the Menil Gallery and Thermal Baths, are predicted in the calming spectrum, with neutral instances in the mornings and late afternoons. These annual predictions can be useful in identifying times of the day and year when sunny sky conditions may have a more pronounced impact on visual perception and emotional response. They also show an inherent variability of daylight in these spaces, where shifting sun angles produce temporally sensitive visual effects which the mSC_5 algorithm is sensitive enough to predict.

5.8 Outlook

The authors acknowledge that this single view position, while commonly used in architectural practice to communicate occupant experience through interior renderings, is limited in its ability to capture the immersive reality of space. To further validate this approach, the next chapter will introduce an experiment using an expanded set of view parameters. To limit potential error due to screen size, brightness, and tone-mapping, these forthcoming experiments will be conducted under controlled laboratory conditions using screen technologies with an extended view (virtual reality) and controlled luminance range. A study by Sarey Khanie, et al. (2015) in visual comfort found a significant effect of luminance distribution on gaze patterns within the field-of-view and suggests the need for a more sophisticated approach to both the application and assessment of perceptual lighting measures. Further work is needed to integrate knowledge of the visual system as our finding that some pixel resolutions are better predictors of perceptual effects than others suggests a possible psychophysical connection.

5.9 Chapter Summary

Building upon the authors' preliminary simulation-based studies which attempt to measure the compositional and temporal dynamics of daylight, this chapter introduces the first data set which confirms a link between quantitative contrast measures, human perceptions of daylight composition in digital renderings and their varied effects over time. This chapter resulted in the following findings: A 3-way ANOVA found a significant effect of space and sky factors on all rating scales, but since the residuals were not normally distributed, a host-hoc Kruskal-Wallis test was used to study comparisons between levels in each factor under consideration for each rating pair. In this test, the effect of sky (when grouped between sky 1, sky 2, and sky 3) was significant (p<0.01) for all rating pairs except for unvaried – varied. Two or more levels in the sky factor were significant for ratings of calming – exciting and subdued – stimulating.

The effect of space (when grouped by RAMMG into a high, medium, and low category) was also significant on all levels and all ratings pairs. To relate the various contrast measures presented in Phase 1 to the subjective data presented in Phase 2, linear and logistic regression analyses were conducted. In a linear regression analysis using Pearson Correlation Coefficient values, RAMMG and mSC_5 (the 5th sub-sampling level in the RAMMG, also referred to as RAMM5) were found to have the highest dependence to all rating scales (for the percentage of subjects that responded 5-7 for each image). This finding supports the preliminary comparative findings on Phase 1 and were similar to the author's own intuitive ranking. These values were the most highly correlated between mSC_5 and ratings of diffuse – direct, calming – exciting, and subdued to stimulating ($r > 0.75$).

Using the results of this analysis to determine which rating pairs show the highest linear correlation to selected algorithms, logistic regression was then conducted between the distribution of subject responses for each rating pair and the mSC_5 algorithm. The proportional odds model from this regression analysis found the lowest deviance between ratings of calming – exciting. Using this proportional odds model, we can predict the percentage of subjects that would rate a given image 1-7 based on its corresponding mSC_5. With this model, mSC_5 was then applied to the semi-annual time series of renderings produced for each of the nine architectural case studies to show how perceptions of calming - exciting may be predicted over time. The use of mSC_5 and this proportional odds model provides the first ever simulation-based method for predicting daylight-driven excitement in rendered images, showing temporal impacts of sunlight on emotional indicators in architecture.

In the following chapter, the mSC_5 model presented here will be applied to an eye-level view position and used to evaluate an occupant's field-of-view across a range of view directions, annual instances, and sky conditions. The assessment of visual interest will then be implemented alongside a novel health-based metric presented in Chapter 7 to illustrate the potential for holistic evaluation of daylight performance in architecture from a human perspective.

6 A HUMAN-CENTERED APPROACH TO EVALUATING VISUAL INTEREST

Building upon the 2D user study and prediction model introduced in the last chapter, we will now propose an immersive 3D approach to predicting Spatial Contrast across a 360° view range using hemispherical view parameters. To emulate an occupant's field-of-view, this chapter will exemplify a human-centric approach to visual interest prediction, using two building case studies to evaluate daylight composition from a fixed view position across an immersive range of view directions. This full 360° range will create incremental coverage of daylight distribution across the field-of-view, using High Dynamic Range (HDR) renderings with a wide fisheye specification as input for the modified spatial contrast (mSC_5) metric proposed in the previous chapter (Equation 5.4). From the fixed view position selected for this analysis, the time of day and day of year will be varied under both clear and overcast skies to evaluate the selected case studies for temporal and climate-driven dynamics. This occupant-driven approach does not seek to replace existing methods for evaluating illumination, but rather propose a human-centric approach for analyzing those aspects of daylight that are dependent on daylight we perceive at eye-level. This model will help shift our attention from spatially-dependent (as in task plane illumination or fixed field-of-view comfort) to human-centric performance assessment methods which account for an occupants' visual immersion in daylit architectural space.

6.1 A Visually-Immersive Approach

When we transition from a fixed view position such as the 80° x 60° rectangular rendering used in Chapter 5 to a human-centric approach for predicting visual interest, three main challenges emerge: 1) how can we use image based metrics like mSC_5 (developed for 2D images) to predict visual interest across a 3-dimensional space and 2) how can we quantify the variation in this prediction based on dynamic environmental conditions and view direction (i.e. across the entire 360° view range). Finally, 3) how can these predictions across a subject's field-of-view, over time, and across sky conditions help contribute to a more holistic assessment of the indoor environment as it impacts programmatic use and design intent? In short, what does this prediction tell us about the occupant's experience and the role of daylight in creating emotionally infused, spatially diverse, and temporally dynamic ambiances in architecture? While this chapter proposes a method of application based on the model we developed from data in the 2D online survey, chapter 8 will delve even deeper into this question using data from immersive 3D renderings in virtual reality.

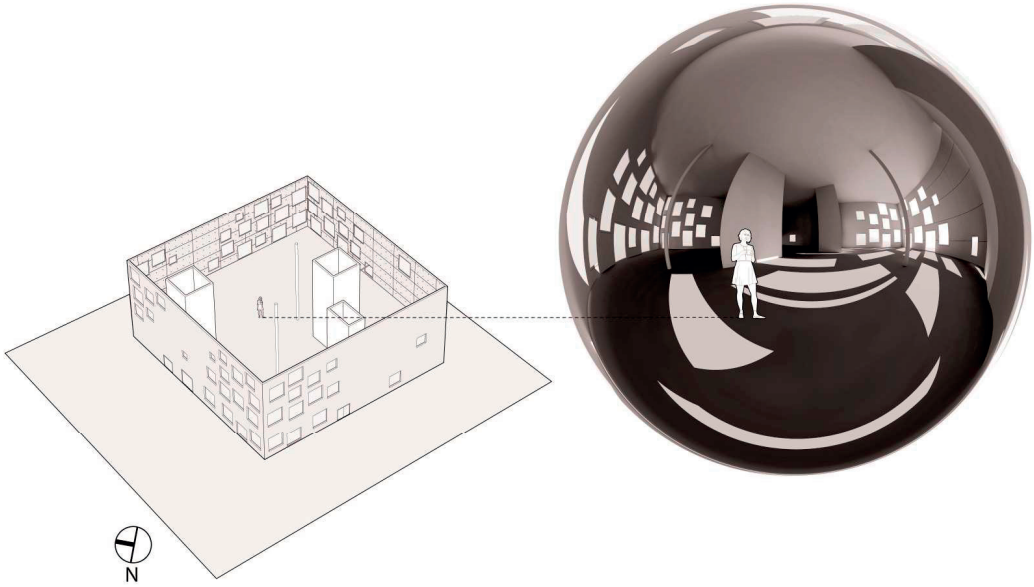


Figure 6.1 A 360° fisheye rendering, created in Radiance from a specified view position (vp), from which an analysis of visual interest may be computed for a series of

As mentioned many times already in this thesis, natural lighting conditions vary frequently based on temporal changes in the indoor environment and an occupant's visual perception of indoor space is influenced by the diverse temporal dynamics of light and shadow over time. 2D renderings have been used in this thesis to assess the perceptual impacts of daylight in architecture using a quantitative model mSC_5 (introduced in Chapter 5) derived from subjective responses, but a single rendering (taken at one moment of time, under a single sky condition) cannot provide adequate information to evaluate the range of temporally induced effects that may occur over time. To address this, mSC_5 was applied to a series of annual renderings to visualize changes in its prediction temporally, but even with these time series of renderings, we cannot compare the relative perceptual impacts of these effects *throughout* the entire view range of an occupant in the architectural space under consideration. To really understand the impact of daylight *throughout* an occupant's perception of space, we must look to apply mSC_5 across an immersive field-of-view. Figure 6.1 shows a 360° fisheye rendering, produced from a given view position (vp). Using the *pinterp* function in radiance, a series of hemispherical fisheye renderings can then be extracted from this single 360° rendering to cover an incremental range of view directions. An immersive application of mSC_5 to this range of view directions will allow us to understand how visual interest varies as a result of diverse architectural conditions within the field-of-view.

6.2 Application of mSC_5 to Visually-Immersive Renderings

Simulation is a powerful tool for evaluating performance dynamics and a robust image-based measurement for predicting subjective impressions of daylight composition on emotional human-responses can help designers to understand where (within their current view and throughout their immediate surroundings), and when (over time), the effects of sunlight and shadow are likely

to produce specific desired or undesired responses. To predict occupant impressions of visual interest, a quantitative model was developed from subjective ratings of visual interest in daylight renderings gathered through the online survey in chapter 6. The survey used in that experiment asked participants to rank a selection of renderings: nine architectural spaces under three sunny sky conditions. Using seven-point semantic differential scales, subject ratings were then gathered for impressions of: low contrast – high contrast, uniform – non-uniform, unvaried – varied, diffuse – direct, simple – complex, calming – exciting, subdued – stimulating. The ordinal responses collected from the semantic scales listed above were then compared to a range of existing and modified contrast algorithms to find which, if any, could be used to predict subject ratings.

We found that for an image (of input resolution 1488 x 1024), the 5th pixel sub-sampling level (N=5) in RAMMG, a local neighborhood contrast measure developed by Rizzi at al. (Rizzi, et al., 2008), could predict subjective impressions of calming – exciting using an ordered logit model. This rating pair was selected due to its goodness of fit. This modified algorithm, will be hereafter referred to as modified Spatial Contrast mSC_5. The cumulative predictive probabilities for each rating on the scale 1-7, from calming to excitement, best fit the equation:

$$P(y \leq j) = \left(1 + e^{a_j - b \times mSC_5}\right)^{-1}, \forall j = i, \ldots, 6\,, \tag{6.1}$$

where $b = 0.24$ is the corresponding effect parameter and a = [-0.45; 0.84; 1.67; 2.82; 3.64; 4.81] the cutoff points. At level 3, the odds of achieving ratings 1-3 were significant at $p<0.05$. Rather than predicting the probability of a subject's response at each level, this chapter will focus on a pair of thresholds for predicting impressions of 'calming' or 'exciting' derived from the ordered logit model presented in Section 5.7.4 and computed in Section 6.4.

To assess the immersive field-of-view representative of a human subject in space, we have used a 360° spherical rendering to generate a series of eighteen 180° hemispherical renderings in 20° radial increments. Using these hemispherical renderings, we can then compute mSC_5 in each of the selected view directions and see how the excitement prediction varies as the architectural composition changes across the field-of-view. While any number of view directions may be used, we selected 20° increments to ensure a good resolution of coverage with some overlap in scenes to create a smooth radial prediction.

Figure 6.2 shows the steps used to predict visual interest across the hemispherical fisheye renderings used in this immersive approach. For each hemispherical image, the matrix of luminance values is first sub-sampled by halving the resolution in 5 subsequent steps. This reduced map of luminance values is then used to compute local variations $c_{i,j}$ in brightness between each pixel and it's eight neighbors, using a weight α applied to each neighbor, element wise. The average $c_{i,j}$ taken across all resulting values of the matrix is then calculated to produce the resulting mSC_5. The integration of these thresholds into an immersive analysis allows us to see, for the first time ever, the potential effect of view direction on predictions of visual interest within an architectural space. Whereas the

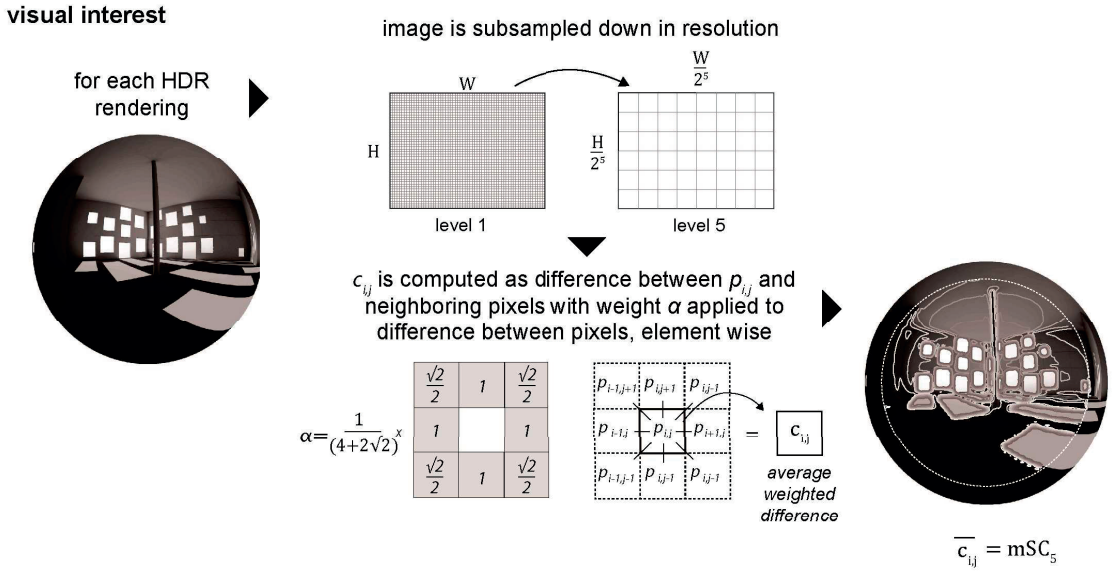


Figure 6.2 For each HDR rendering, the image is subsampled down. Local luminance variations are then computed for each pixel and its eight surrounding neighbors to create a matrix of local variations in compositional brightness. From this new matrix, the average mSC_5 is then computed.

previous chapters focused on predicting impressions using a 2D rectangular rendering with a fixed view direction, this approach allows for the immersive assessment across an entire 360° view range, revealing the effects of orientation on architectural composition for predicting impressions of excitement and calm.

Using this immersive analysis approach, we begin to move beyond conventional spatial methods and address perceptual responses to daylight as seen from the perspective of an occupant. The case studies selected to demonstrate this approach will be introduced in the following section together with the simulation parameters selected for analysis. In Section 6.4, threshold values chosen for the different performance indicators will be identified, with assumptions to the maximum and minimum acceptable values explained in more depth.

6.3 Simulation work flow

To assess the impacts of daylight on predictions of visual interest across an occupant's immersive field-of-view, the following work flow has been established. Within a given 3D digital model, a view position is selected to represent an occupant's gaze at eye level. From this view position, a series of spherical renderings are generated for each time and date selected for analysis. To generate the spherical renderings for analysis, Radiance was used to produce a 360° x 360° angular fisheye image as well as a '.spec' file which contains information about the view position and an index of each pixel in relation to the rendered view direction. These spherical hdr renderings and spec files were then used to generate a set of hemispherical fisheye images from a series

of 18 radial view directions using the *pinterp* function in Radiance. With this function, each hemisphere is 'unrolled' into a new view direction resulting in series of 18 images, each centered on 20° radial increments from North. This work flow is shown in Figure 6.3 and enables the rapid production of images for any number of view directions from a fixed view position in space. Using a conventional approach with independently rendered angular hemispherical images in each view direction, we would need to produce a new ambient file for each times. For that reason, the spherical approach saves significant computational time and also allows future flexibility in analysis as we are able to unroll additional view directions using the same spherical hdr and spec file. Each hemispherical image is then analysed to evaluate daylight impact using mSC_5 with the threshold values described in Section 6.4.

6.3.1 Simulation Setup and Case Study Buildings

To illustrate this applied method, two buildings were selected for the application of mSC_5 in order to compare predictions between interiors which vary in the symmetry and complexity of daylight distribution. The first building we considered was SANAA's Zollverein School of Management in Essen, Germany (latitude: 51.5°N, longitude: 7°E) shown in Figure 6.4. This space was selected in large part due to its unique distribution of windows in each of the four cardinal directions. This 35-meter cube was completed in 2006 and contains of a series of vertically stacked single, double, and triple height educational and mixed-use spaces. This case study was selected as a

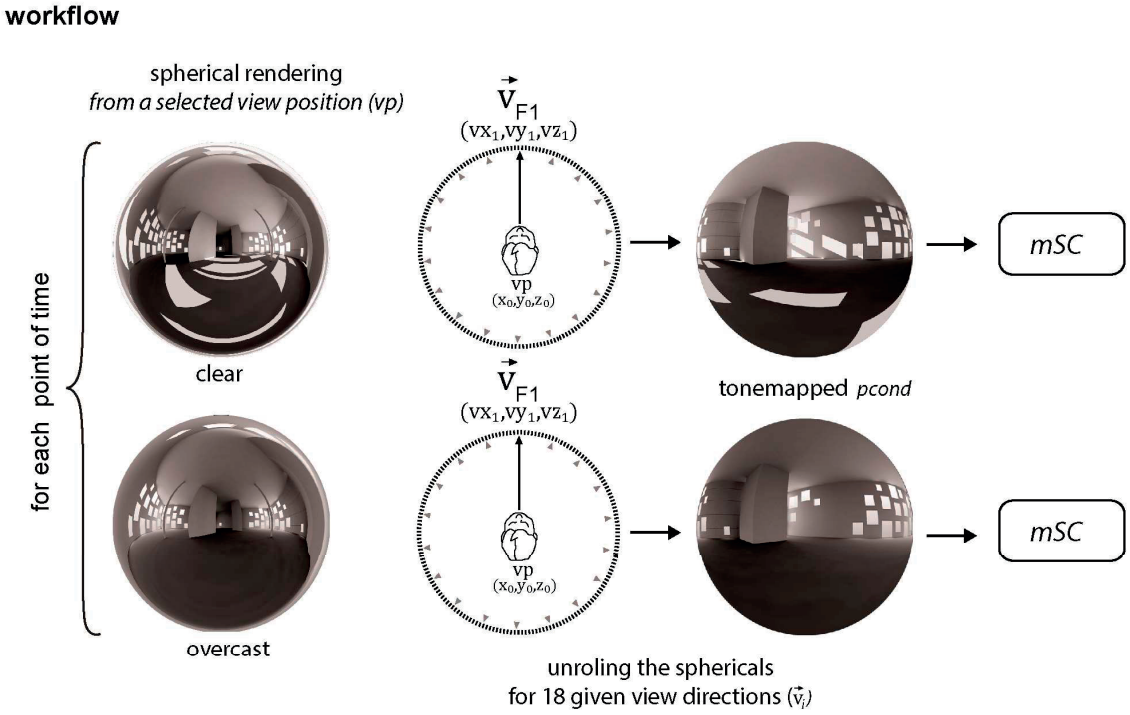


Figure 6.3 For each view position (vp) during each time step, a clear and overcast 360° fisheye rendering is produced, tone-mapped, and unrolled in a series of eighteen view directions (in 20° polar increments). Each view direction produces a new 180° hemispherical fisheye image.

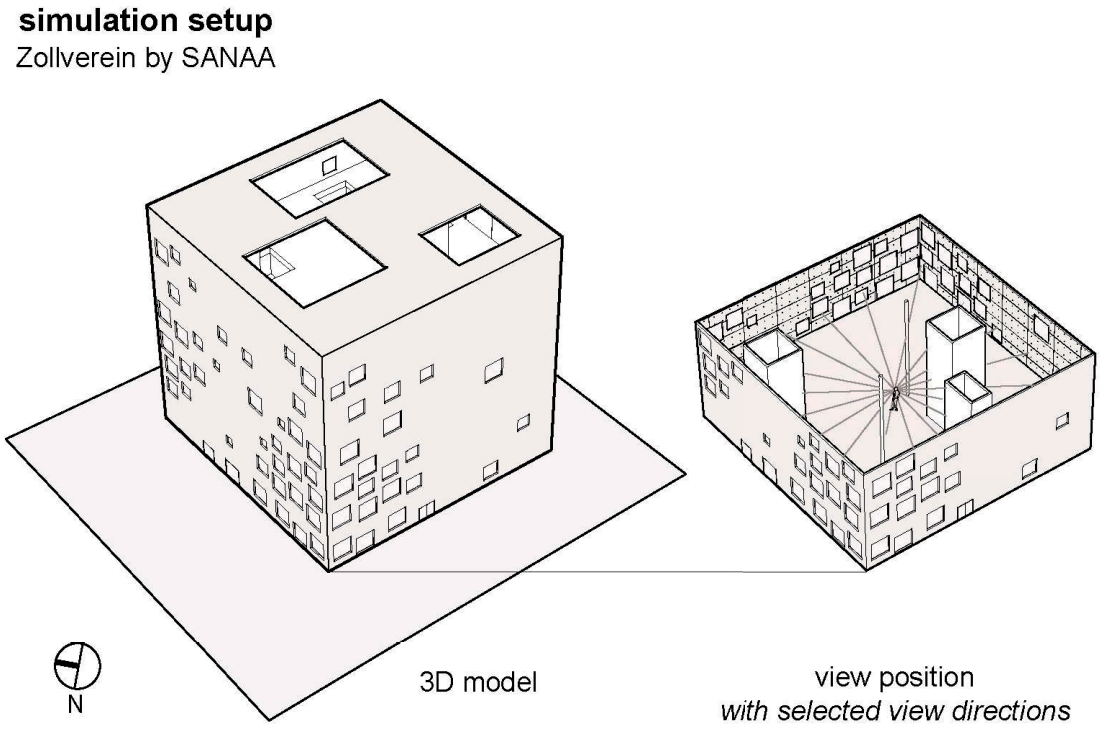


Figure 6.4 The simulation setup: (a) 3D model of the case study building: SANAA's Zollverein School of Management in Essen, Germany, (b) The view position is located in the center of the space with eighteen view directions evenly distributed radially at 20° increments to cover the entire visual scene.

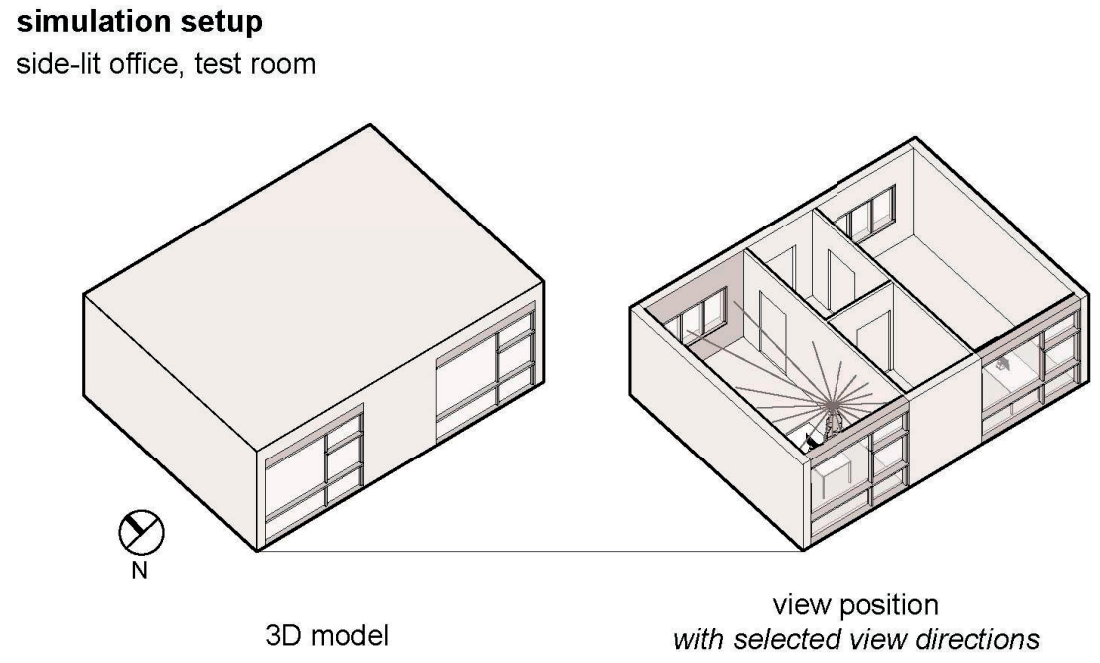


Figure 6.5 The simulation setup: (a) 3D model of the case study building: A side-lit office test space in Fribourg, Switzerland, (b) The view position is located at eye level while sitting looking west towards the computer screen with eighteen view directions evenly distributed radially at 20° increments to cover the entire visual scene.

representative example of contemporary architecture where daylight and an understanding of its dynamics have played an important role in its design. Using an architectural case study where daylight considerations were central to the architect's design intent allows us to discuss the role of architecture in choreographing human responses to light stimulus within an interior space. The façade is composed of asymmetrically distributed square windows, which filter sunlight and provide framed views of the surrounding landscape. The authors have selected a central view position in a triple-height mixed-use space (first floor level), with eye-level 1.67 meters from the floor, shown in Figure 6.4.

The second case study selected for this application was an office-like test room, designed for field studies in Fribourg, Switzerland (latitude: 46.8°N, longitude: 7.14°E), as developed for the Smart Living Lab (http://www.smartlivinglab.ch) and shown in Figure 6.5. This test room was chosen because it represents a more normative interior spatial configuration and façade design. In addition to its more muted architectural complexity, this case study is also smaller in scale, demonstrating that the immersive analysis proposed in this chapter is relative to the occupant's visual field and can be applied to any view position in any space. The fixed view position selected for the test room is not centrally located as in the Zollverein, but rather located at eye-level for someone seated (1.16 m from the floor) in front of the desk in the South East corner of the room (Figure 6.5). This seated view position allows us to assess the impacts of daylight in a space where its distribution is asymmetrical and typical of many side-lit office environments where indoor occupants spend a significant number of daytime hours.

For both case studies, eighteen default view directions have been used to generate hemispherical renderings in 20° even radial increments to cover the entire visual field. Using 6 daily moments, starting at 8h and running in 2-hour increments until 18h, we have chosen to assess the typical work-study period for an academic or office building. These 6 daily moments are repeated across 4 semi-annual days, representative of a symmetrical half year (Figure 6.6). This approach is slightly different from the symmetrical hourly distribution of instances used in Chapter 5. In this case, we used an even time step for each day as a parallel set of analyses described in (Amundadottir, et al, 2017). This decision was due in part to the holistic integration of a time-based dosing model developed to predict non-visual health effects (Amundadottir, et al., 2016). For each point in time, a spherical rendering was generated under CIE clear and CIE overcast sky models using the rendering parameters and material definitions in Tables 6.1 and 6.2, respectively.

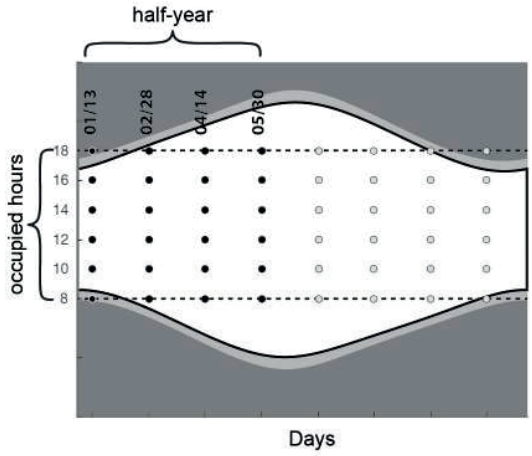


Figure 6.6 Dates and times used for annual analysis, taken in 2 hour increments from 8am to 6pm.

Table 6.1 – Non-default rendering parameters used as an input to *rcontrib* in Radiance

dt	dj	ds	ab	aa	ar	ad	as	lr	lw	pj	ps	pt
0.05	0	0.15	3	0.1	512	4096	2048	8	5e-3	2	0.05	0

Table 6.2 – A list of Radiance materials used in this simulations study.

Materials	Description
GenericInteriorWall_50	Diffuse reflectance of 50%
GenericCeiling_70	Diffuse reflectance of 70%
GenericFloor_20	Diffuse reflectance of 20%
Glazing_DoublePane_Clear_80	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%; SHGC = 0.72 ; U-Value = 2.71 W/m^{2}K

This shift from using only clear sky renderings in Chapters 3-5 to using both clear and overcast sky models allows us to see the relative impacts of weather as well as the relative effects of time for each sky type.

6.4 Application of mSC_5 to Hemispherical Views

As described in Section 6.2, the mSC_5 model used for application in this chapter was derived from the online experiment in Chapter 5, which compared subjective ratings of 2D tone-mapped renderings to a range of contrast algorithms and found that one modified algorithm could be used to predict ratings of calming - exciting (Rockcastle, et al., 2016). From the distribution of ratings across the renderings in this experiment, the authors used an ordered logit model to predict the probability of subject responses at each level, from 1–7. From this analysis, it was found that cumulative responses were significant at level 3, i.e. the odds of achieving a rating of 1-3 were significant at $p<0.05$. For the purpose of this paper and from the finding in this analysis, we decided to group levels into either calming, neutral, or exciting. Using this model, two thresholds were identified for predicting occupant impressions of calming (j=3) and exciting (j=4). In the range 0-20 mSC_5, we obtain these thresholds by solving for mSC_5 in equation 6.2 and by setting the predictive probability to 50%, which gives us:

$$mSC_5 = (ln\ 1 - a_j)/b\ , \tag{6.2}$$

In this case, j=3 represents the accumulation of levels 1-3 and j=4 represented the accumulation of levels 1-4. Predictions for ratings 5-7 are those represented by any value above the j=4 threshold. As seen in Figure 6.7, the obtained threshold values are 6.96 (j=3) and 11.75 (j=4) respectively for calming and exciting. mSC_5 values between 6.96 and 11.75 are therefore established as neutral as this range represents subjective ratings in the middle of the bi-polar scale. Although these thresholds were derived from subjective data on 2D images with a horizontal and vertical view range of 80° x 60° (respectively), the mSC_5 for each image is computed as an average of local values taken across the composition. As such, the same thresholds for average mSC_5 can theoretically be applied to images with a different set of view parameters, such as the 180° fisheye selected for this paper because the average and not the sum of values is used to determine subject

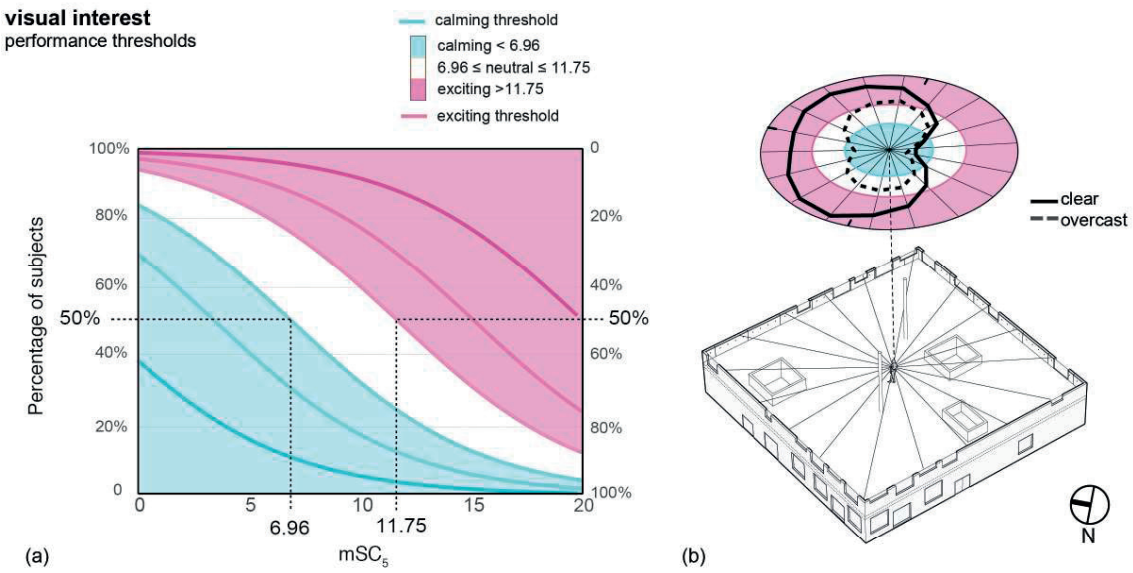


Figure 6.7 Calming and exciting responses are determined using the thresholds at 6.96 (j=3) and 11.75 (j=4), where 50% or more of the population is predicted to rate that rendering on one side of the scale. A point-in-time overlay of results for clear and overcast skies is shown to the right so as to visualize the impacts of the case study and view directions on predicted responses.

responses. Further work is needed to understand the precise view range most appropriate for an immersive perception study, taking into account the degradation in acuity and contrast sensitivity in the periphery of our vision.

In the case of this study, 180° HDR fisheye images were imported into matlab and sub-sampled down to the 5th pixel resolution. Due to the edge condition of the circle and the process of sub-sampling, a 5° degree perimeter was lost on each image and mSC_5 was therefore calculated on the remaining 170° degree fisheye. Since an appropriate view range has not been established for use in subjective assessments of visual interest, the authors have accepted this slightly reduced view angle pending further refinement through studies that address the impacts of interest perception on the human visual system. Further work in the development of this model will look at subject predictions gathered from across an immersive image to further refine the prediction model and its relationship to immersive viewing environments. This will be presented in Chapter 8.

6.5 Results

The preliminary visulization method shown in Figure 6.7 and used in this chapter relies on a angular plot of connected mSC_5 values (in solid line for clear sky and dotted line for overcast) between view directions. By coloring the area within the graph where valued are under the

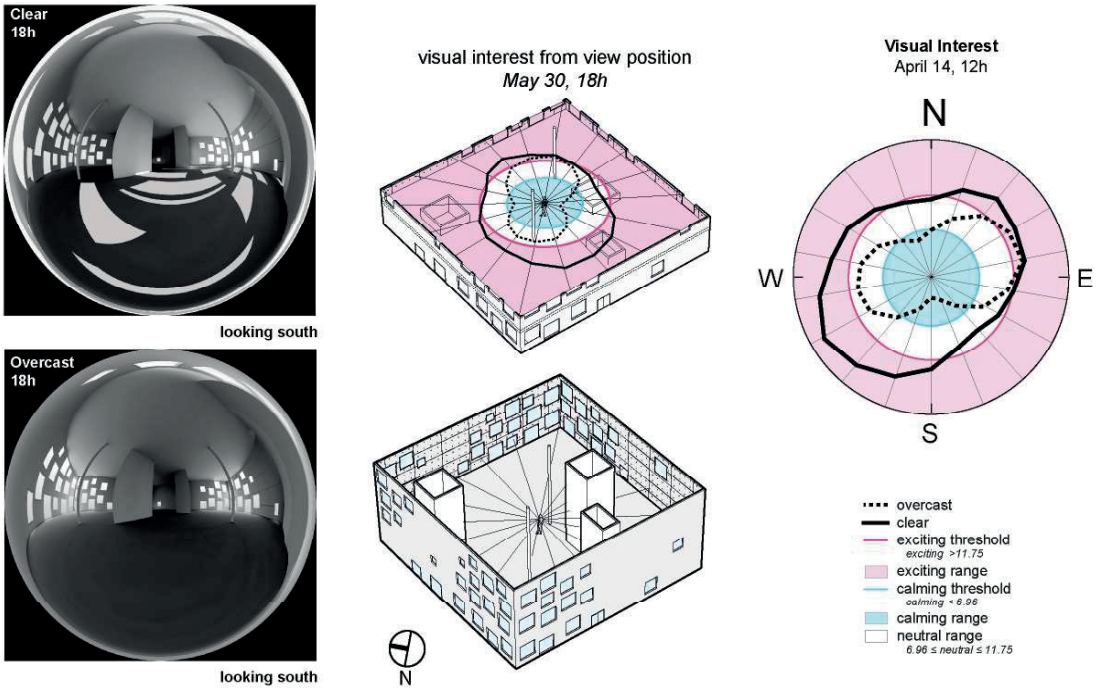


Figure 6.8 Shows 360° hemispherical renderings for clear and overcast skies on May 30 at 18h in the Zollverein School of Management. mSC_5 results across all 18 radial view directions are overlaid within the space with calming and exciting thresholds highlighted (in cyan and magenta, repectively). Clear sky results are shown in a solid line, with overcast dotted.

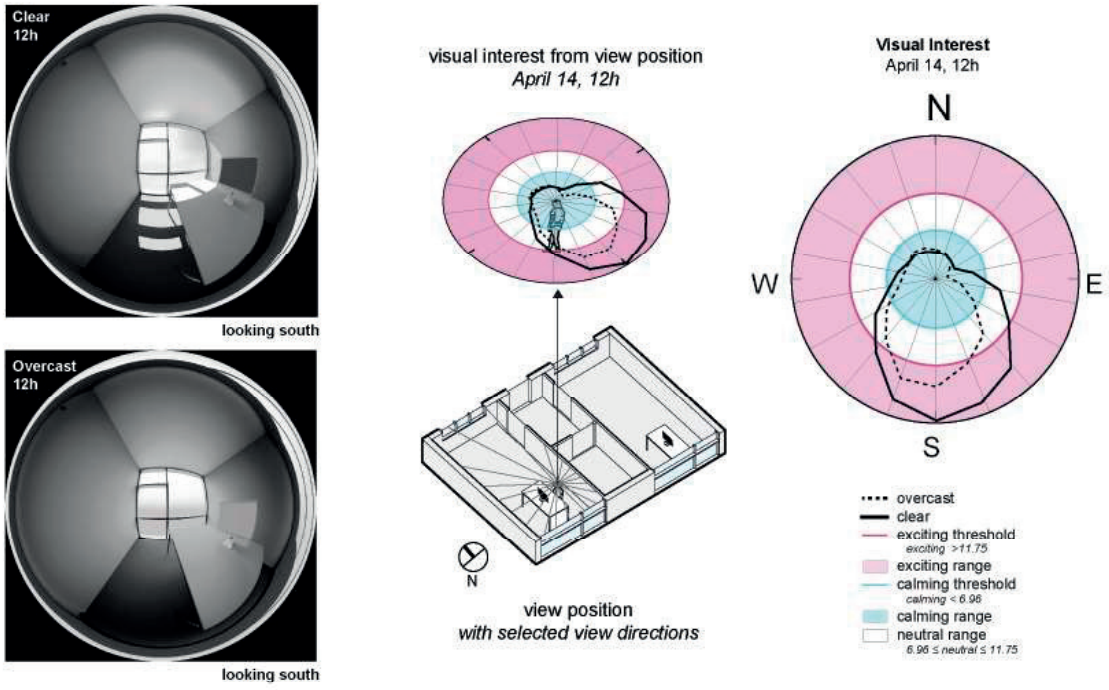


Figure 6.9 Shows an example of 360° hemispherical renderings for clear and overcast skies on April 14 at 12h in the Fribour Test Room. mSC_5 results across all 18 radial view directions are overlaid within the space with calming and exciting thresholds highlighted (in cyan and magenta, repectively). Clear sky results are shown in a solid line, with overcast dotted.

calming threshold (cyan) or over the exciting threshold (magenta), we can see which view directions and sky conditions fall within these ranges and how they change as we move our view direction around the scene. In this section, mSC_5 will be applied to the case studies introduced in Section 6.3 and results will be represented using the angular plots described above to enable the comparison of results for each sky condition and view direction across the full horizontal view range. Angular plots were used as a way to visualize predictions with a directional dependency as they may be overlaid easily within the architectural plan. The results will be compared over time, to discuss the role of temporal dynamics, and between spaces to discuss how the interior layout, view position, and façade can contribute to occupant predictions of excitement and calm.

Figures 6.8 and 6.9 show an example of results for each case study selected for analysis in this chapter. On the left, you can see a pair of 360° hemispherical fisheye renderings for two time-steps (May 30, 18h - Zollverein School and April 14, 12h - Test Room) under both clear and overcast skies, with the resulting mSC_5 values and calming/exciting thresholds shown in cyan and magenta on the right. When the prediction data is overlaid within the geometry models, it becomes possible to see the spatial impact of architectural elements on excitement predictions. The view directions facing windows/more complex openings achieve more exciting predictions under clear sky conditions are significantly reduced under overcast skies for both spaces.

6.5.1 Case Study 1: Zollverein School of Management

The radial plots shown in Figure 6.10 illustrate the results for mSC_5, plotted across each of the 18 radial view directions in 2-hour increments from 8h to 18h on each of the four selected dates. As 8h on January 13 falls before sunrise and 18h falls beyond sunset, these instances have been excluded for assessment of visual interest in this case study. While it seems intuitive to begin the analysis using a start time after sunrise, the intention was to analyse the space during regularly occupied hours for the programmatic use of each case study and therefore 8am was maintained. This decision was also made based on a parallel application of non-visual health potential, which relies on even time steps to generate a dosing profile related to ocular light exposure.

The results in Figures 6.10 show a clear sensitivity to the composition of indoor space, with predictions of excitement generally higher towards the East and West corners, where there is an increased frequency of window openings resulting in more direct sunlight penetration. There is also a visible impact of sky condition, as overcast scenes are almost always predicted to be less exciting and more neutral or calming than the same instance rendered under a clear sky scene. Results from the 2D experiment in Chapter 5 did not account for overcast sky conditions as the three moments considered for each space were all taken under sunny skies. These results are particularly interesting as they show a potentially significant range in perceptual response generated in the same view direction from sky conditions alone. This will be explored further in Chapter 8.

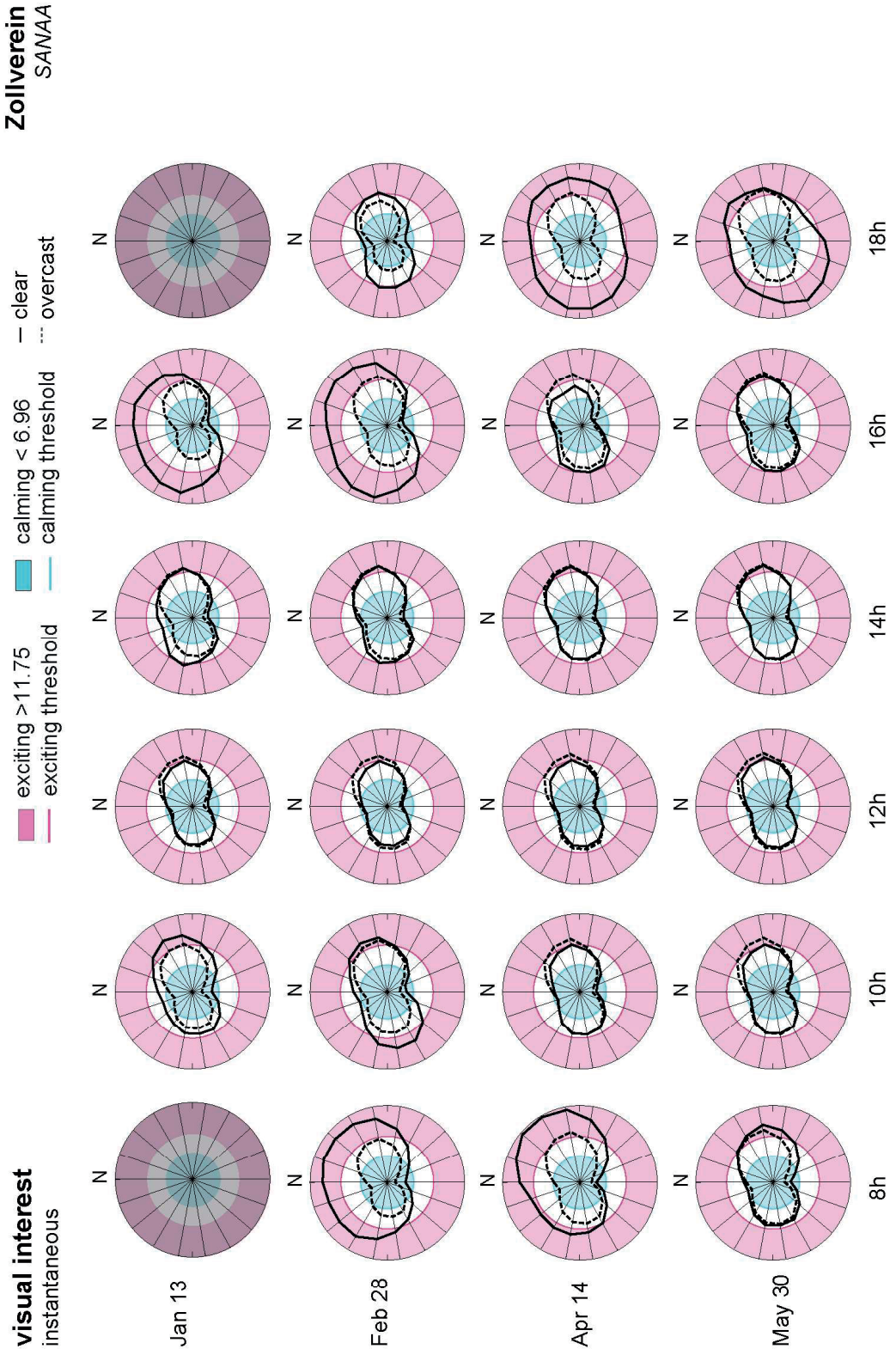


Figure 6.10 Radial plots showing results for of visual interest (mSC_5 values) for the Zollverein School of Management across 4 daily and 6 hourly instances with exciting (magenta) and calming (cyan) thresholds obtained from the experimental results in Chapter 5.

While it is interesting to see the difference between sky conditions, it is equally informative to see how dynamic sun positions affect predictions of visual interest over time. If we look at the radial plot for February 28 at 8h, we can see that the overcast sky condition remains calming or neutral in all view directions, while the clear sky condition results in an exciting response throughout those view directions oriented North. The radial plot for April 14 at 8h shows a similar trend, with exciting responses higher towards the Northeast, where sun angles have penetrated further in through the windows distributed towards sunrise. An opposite response can be seen on May 30 at 18h, where the most exciting view directions are predicted towards the Southwest, where sunset angles have a greater effect in the field-of-view.

6.5.2 Case Study 2: Test Room

The radial plots in Figure 6.11 show the results for mSC_5 across each of same 18 radial view directions, this time applied to the Fribourg Test Room. Where the distribution of results favoured the exciting threshold in East and West orientations for the Zollverein case study, the distribution of results for the Fribourg test room show an asymmetrical distribution of exciting view directions for both clear and overcast sky conditions facing South. While there is a clear pattern in predictions regardless of sky condition, there is also a clear impact of sky, particularly during January 13 and February 28. This finding makes sense when you consider the South-facing façade, the low sun angles at that time of year and the impact of direct sunlight through the facade in those view directions. While excitement predictions breach the 11.75 threshold for 8 out of 18 view directions under clear skies at 10h on January 13, they only apply to 3 view directions at the same instance under overcast skies. This translates to a significantly smaller view angle which may induce a specific perceptual response as compared to the clear sky conditions.

6.6 Discussion

The impact of view position and an occupant's access to stimulating view directions can be discussed with some interest using the results of this case study as there is a clear difference between predictions in each view direction. If a subject is sitting close to an interior wall, which is the case in this analysis, little excitement can be expected in the direction of that wall due to the compressed depth of view and limited complexity of space. As a result, subjects seated throughout a given space will experience a range of possible variations in visual interest depending on position and view direction. The impacts of this knowledge could influence designers to locate certain occupants or tasks in specific locations to provoke feelings of excitement of calm that are appropriate to that task, occupant, and location.

While the impact of sky condition can vary an occupants' perceptual impressions of space at any given moment, the variability of those impressions over time are more stable under overcast skies than they would be under a clear sky. While this may seem intuitively obvious, it is essential to understand the impact of climate on our perception of daylight in the indoor environment.

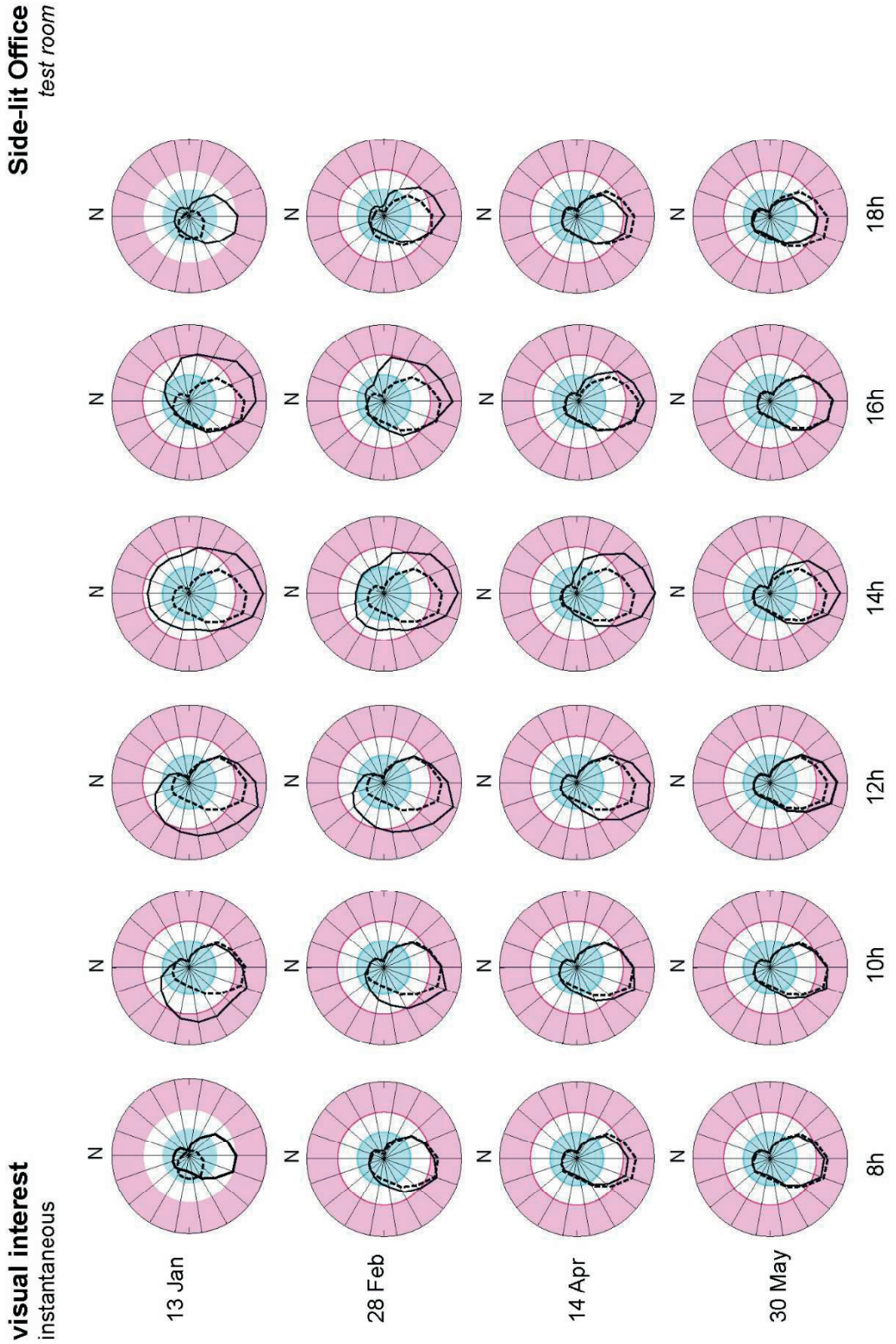


Figure 6.11 Radial plots showing results for of visual interest (mSC_5 values) for the side-lit Fribourg test room across 4 daily and 6 hourly instances with exciting (magenta) and calming (cyan) thresholds obtained from the experimental results in Chapter 6.

Overcast sky conditions, while generating some variation in brightness and distribution over time, produce much more stable interior lighting conditions. The composition of façade elements and a subject's distance from those elements may still trigger impressions of excitement under an overcast sky, such as can be seen in the side-lit office space, but those predictions are always generally weaker than the same instance under a clear sky. The role of climate and dominant sky conditions must therefore be considered when analysing an occupant's perceived impressions of calm and excitement as the presence of absence of sunlight plays an important role in revealing or neutralizing visual effects.

6.7 Chapter Summary

This chapter introduced a method for assessing visual interest across a 360° immersive field-of-view from a fixed position in space. Using the 2D algorithm and predictive model (for calming – exciting) presented in chapter 5, a method for applying this algorithm was adapted to a 180° hemispherical fisheye image and applied to a series of eighteen view directions, taken in 20° angular increments. To understand the dynamics of time and sky condition, these eighteen view directions were then assessed under two sky types (clear and overcast) across 2-hour increments for 4 days which represent a symmetrical half-year. An overlay of clear and overcast skies, plotted for each hourly instance across all eighteen view directions illustrate the impacts of view direction, architectural composition, and sky on predicted impressions of calm and excitement. A comparison of results for two different architectural case studies, the Zollverein School of Management and a side-lit office space demonstrate how the application of mSC_5 can be used to compare spaces that vary in architectural composition.

This multi-view-direction analysis offers a substantial contribution towards the application of human-centric performance models, but a single view position in space cannot sufficiently represent an occupant's immersive experience within architecture. Our dynamic user behaviour requires that multiple view positions be considered to account for the multiplicity of conditions that can be experienced throughout a building between occupants and over time. To move from a single to a multi-view position approach, Chapter 8 will introduce a platform that streamlines inputs for mSC_5 across an array of view positions, sky conditions, and time-series steps. The visual interest model presented in Chapters 5-6 will be integrated alongside a novel metric for health potential to exemplify the potential for a more holistic assessment of human-centric daylight performance. The move from a single to a multi-view position analysis necessitates further development in the visualization of performance data due to the sheer quantity of data. This will be explored and discussed in the following chapter.

7 A SPATIALLY-IMMERSIVE SIMULATION-BASED WORKFLOW

Where the previous chapter offered a method for assessing perceptual performance across sun positions and sky conditions from a single, fixed view position, the work flow presented in this chapter will introduce an annual simulation-based platform for assessing daylight across a series of view positions, exemplified through a multi-level architectural case study. To evaluate daylight characteristics related to human perception and emotion in architecture, the application of visual interest models must be considered across a dynamic set of environmental and behavioural inputs. We experience architecture through inhabitation and while we might occupy some positions longer than others (in the case of a fixed work station), it is important to evaluate the diversity of *space* as well as the diversity of view and sky conditions. The added dimension of space increases the amount of computed data and requires a new framework to visualize performance predictions.

To illustrate the potential for a holistic, human-centric, and spatially diverse evaluation method, this chapter will apply mSC_5 alongside a novel non-visual response model developed in parallel by members of the LIPID research group. Through the integration of mathematical models used to predict visual interest (introduced in the previous two chapters) and non-visual health potential (Amundadottir, 2016), this chapter introduces an automated work flow to assess an array of view positions (located at eye level) under varied sky conditions and across multiple view directions. Building upon the work completed by Kleindienst, et al. (2008) and Andersen et al. (2013) which relies on annual goal-based performance assessments using an optimised rendering engine called OpitX™ (https://developer.nvidia.com/optix), the platform presented in this chapter is supported by Radiance with a user-friendly interface used to set rendering/analysis parameters and call modular performance models. While Radiance can be slower and require greater computational intensity than OptiX™ based simulations, its robust validation makes it a more reliable tool for daylight research. By providing a user-friendly interface with a robust simulation engine, this novel workflow allows users to define location, sky, and time-based parameters, generate batch scripts in Radiance to run simulations across any number of eye-level sensor points, and then analyse those points using mathematical models developed to measure visual interest and health potential. Through an integration of immersive analysis techniques (as described in the previous chapter), this approach allows for a spatial (across many sensor points) and occupant centric (across many view-directions from those eye-level sensors) analysis of daylight through a modular simulation platform. Using a new method of visualization presented in section 7.1.3, results will be illustrated across space, between skies, and over time.

CHAPTER 7: SPATIALLY-IMMERSIVE WORKFLOW

7.1 Development of an Integrated Platform

As stated above, the primary motivations for development of this modular platform are 1) to integrate perceptual and health-related modules at the building and occupant scales, 2) to offer a simple interface that guides users step-by-step through the simulation setup resulting in an automated work flow and, 3) to provide spatial and temporal visualizations that illustrate the effects of daylight and architecture on human-centric performance predictions. Using Radiance as the background simulation engine, this platform implements performance modules for visual interest (mSC_5) and human health potential (nvR_D) (Amundadottir, 2016).

7.1.1 *Simulation Workflow*

This work flow is based on the concept of a time-series which has been simulated for a single, centralized position (as described in the previous chapter) and is here implemented and generalized to the simulation of multiple viewpoints in space over any number of desired moments and sky conditions. While integrating perceptual and health-based performance modules in a manner that is accessible for less experienced users, the platform remains flexible for more advanced users who want to override defaults and create more controlled or atypical simulations. The user is guided through this work flow step-by-step using a simple interface that calls/runs Radiance to generate illuminance and luminance-based outputs. Instead of running the simulation locally, the platform can generate shell scripts for running computationally demanding projects on an external server. One of the advantages of using this interface is more efficient data management, where results are stored in an organized data folder system. To speed up simulation times, the generation and storage of shared ambient light files makes the rendering of many view positions possible without a linear increase in rendering time. Using a method of extracting multiple views from a single 360° image, we were able to use one rendered scene to generate many 180° fisheye images across a desired range of view directions, as described in Chapter 6.

The simulation work flow is illustrated in Figure 7.1 and outlined below:

a. Load a 3D model (OBJ, DXF, 3DS, SKP) and convert it to RAD material and geometry files. Alternatively, RAD and material files can be loaded directly. Preview the 3D model, allowing users to navigate through the scene and verify file translation.

b. Select the site location, weather and moment distribution desired for analysis and create OCT files (one per moment and sky condition).

c. Define the position of analysis sensor points by reading in a text files or export of point locations from a geometry modeling platform like Rhino.

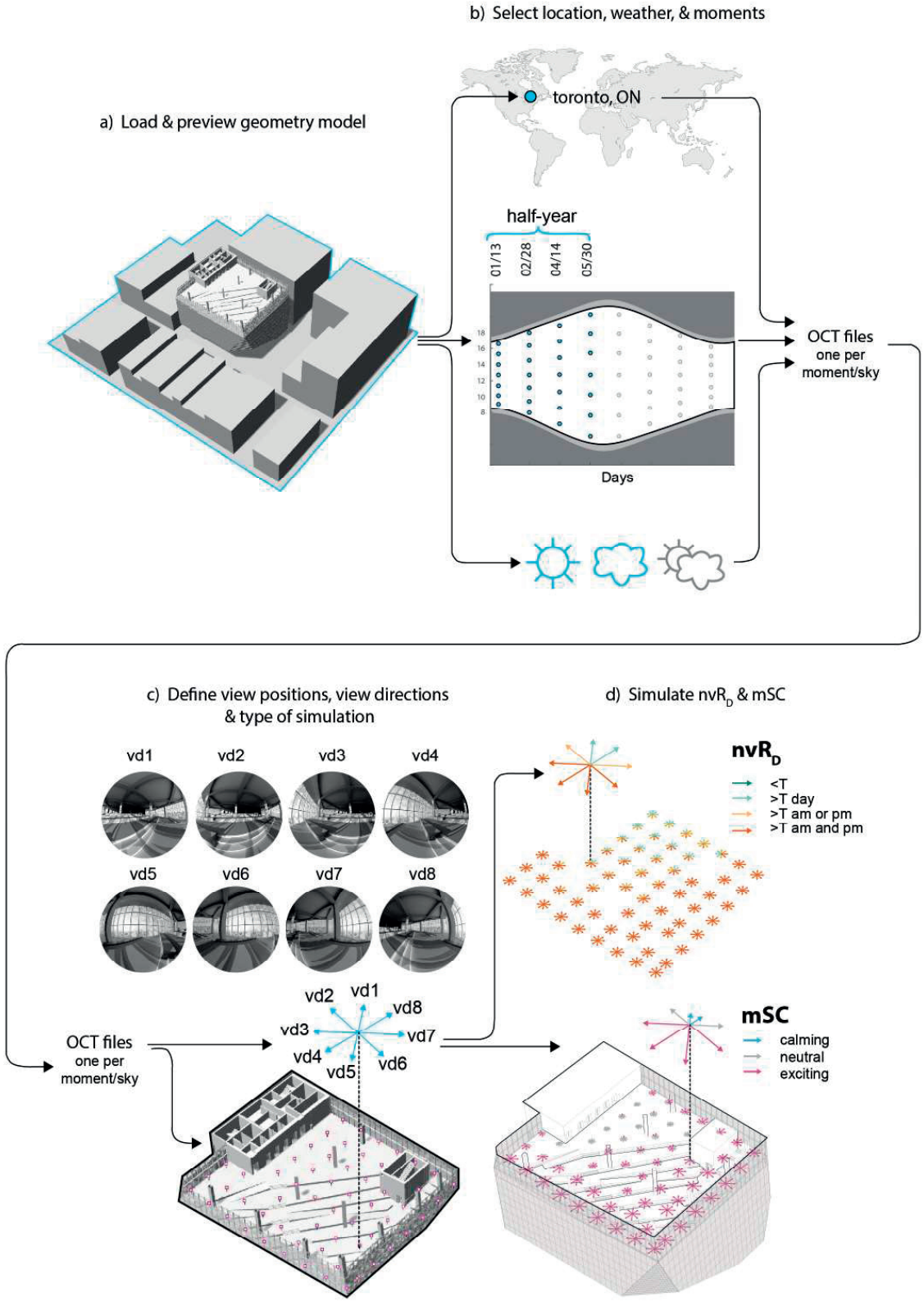


Figure 7.1 Workflow diagram showing simulation steps from 1) loading and previewing a geometry or RAD model to b)selecting dates and times to c) defining view directions and type of analysis to d) simulating health and visual interest modules.

d. Run the simulation locally (time intensive) or create the scripts to run simulations externally (time optimized) to generate illuminance and/or luminance outputs. Compute desired performance modules like mSC_5 and nvR_D from the obtained illuminance/ luminance results and visualize the results spatially and temporally.

7.1.2 Performance Modules

While a detailed mathematical description of the mSC_5 model is introduced in the previous chapter, the nvR_D model is described in depth in Amundadottir (2016) and Amundadottir, et al. (2016). The following section will skip a more in depth description of these models and instead outline the implementation of performance modules into the proposed simulation platform.

Predicting Visual Interest in Renderings

The online experiment introduced in chapter 5 compared a broad range of image-based algorithms to find that RAMMG and a modified version of that algorithm, called modified Spatial Contrast (mSC_5), were highly correlated to subject ratings of visual interest on bi-polar scales from calming – exciting and subdued – stimulating. Using logistic regression, the authors fit mSC_5 predictions to the distribution of subject responses for each rendering to compute the percentage of subjects that would rate that image in the calming, neutral, or exciting spectrum.

In this chapter, the mSC_5 algorithm is computed on hemispherical images derived from spherical HDR renderings obtained using *rpict* > *pinterp* in Radiance (described in Chapter 6). Each Radiance picture is then independently tone-mapped and compressed using *pcond* > *ra_bmp* to provide the appropriate image-based input for computing mSC_5. Based on the computed mSC_5 value, the fitted model then returns a prediction of visual interest. From this fitted logistic function, two thresholds were determined to predict perceptions of calm and excitement in a majority percentage of the surveyed population. The application of the mSC_5 algorithm and threshold predictions of calm and excitement are described in depth in chapters 5 and 6.

Predicting Non-Visual Health Potential

Shifting from visual to non-visual effects of light, there are several factors that have been linked with health-related performance indicators, such as the quantity of light we receive over time and the duration/ timing of that exposure. Increased daytime exposure to bright light (<1000 lx) has been positively associated with sleep quality by Hubalek (2010) and shown beneficial effects on alertness and vitality by Smolders, et at. (2013). Light exposure does not always induce positive effects, however, as night-time exposure of dim light (>100 lx) can shift the circadian clock and disturb other behavioural/physiological processes such as melatonin production (Zeitzer, 2000). The estimation of a healthy daily light dose poses many challenges, in part due to the relative novelty of research in this areas. Non-visual responses must be evaluated as a dynamic system,

which adapt to intensity, wavelength, duration, history, and timing of light received at the eye and ultimately at the retina. Based on recent work by Amundadottir (2016), the work flow presented in this chapter applies a novel model, called non-visual direct-response (nvR_D). This model evaluates the non-visual health potential of light by integrating its underlying photo biological properties on responses in humans.

The outcome of this model is evaluated over a 24-hour day and gives a cumulative response (RD). The cumulative response RD is mainly sensitive to total light intensity and duration of light exposure, therefore providing a measure of daily light dose, which is independent of circadian timing. At any time during the day, it is possible to return an intermediate value if, for example, it is relevant to evaluate a spatial position for shorter periods of time. This is especially relevant in buildings, where our occupation may span anywhere from a couple of minutes to several hours or days. Setting minimum targets or performance goals is necessary to evaluate non-visual health potential as we would like to know how much daylight we need, within a given space, from a given position and view, to achieve a desired health potential. Given a reference profile of an ideal light exposure, the nvR_D model can produce a target for evaluating performance. The nvR_D model is described in depth in (Amundadottir, 2016) (Amundadottir, et al., 2016).

7.1.3 Visualization of Results

The mSC_5 and nvR_d performance modules take luminance and/or illuminance data as input for a series of view positions within a geometry model and present the results in a format that illustrates whether performance thresholds have been met across each view direction over an established time series. The production of simultaneously spatial and temporal data helps to inform designers, engineers, and building operators about *where* and *when* daylight might affect human perceptual and non-visual responses in the built environment. While a preliminary (static) visualization of results for both mSC_5 and nvR_D is provided for a select architectural case study in this chapter, future development of this platform will allow users to interact with results across view positions and over time. An overall picture of average daily performance can be just as useful as a detailed assessment of hourly performance or an instantaneous prediction across view directions. For the demonstration of this simulation-based work flow, we selected an architectural case study that presents a multitude of interior daylight conditions due to its patterned glass façade, varied interior layout, and mix of programmatic uses.

7.2 Selection of Architectural Case Study

The Ryerson Student Learning Center (SLC) in Toronto, designed by Snøhetta and Zeidler Partnership (opened to the public in 2012), contains a mix of programmatic functions. The building is surrounded by a dense urban context, ranging in height from 3 to 9 stories (Figure 7.2). As an expansion to the neighboring library, the SLC houses student work/study spaces, staff offices, and collaboration spaces – both open and enclosed. The integrated daylight design concept harnesses

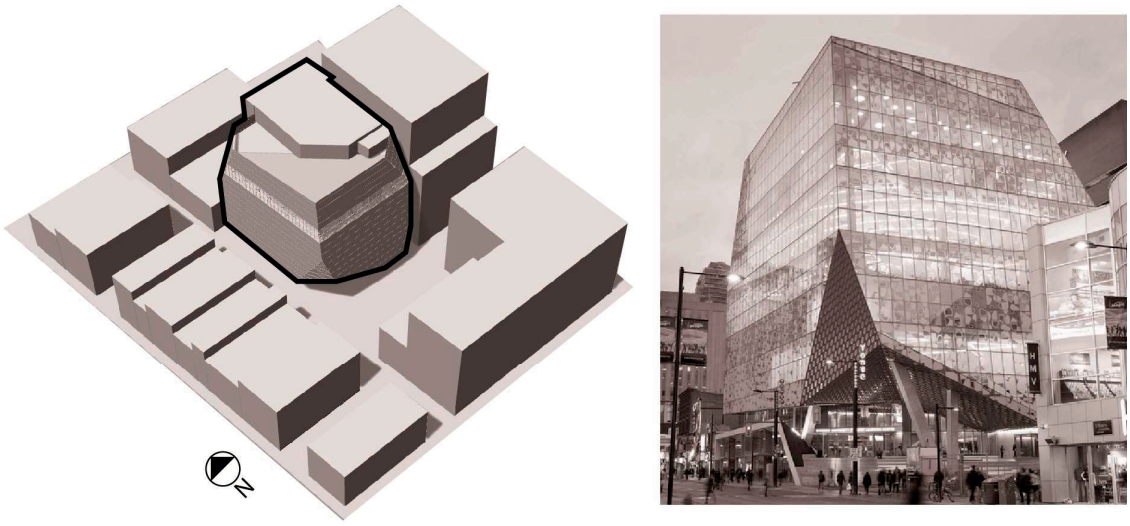


Figure 7.2 Showing the Rhinoceros model and surrounding urban conext on the left and a photograph of the exterior facade, courtesy of the architect Snøhetta.

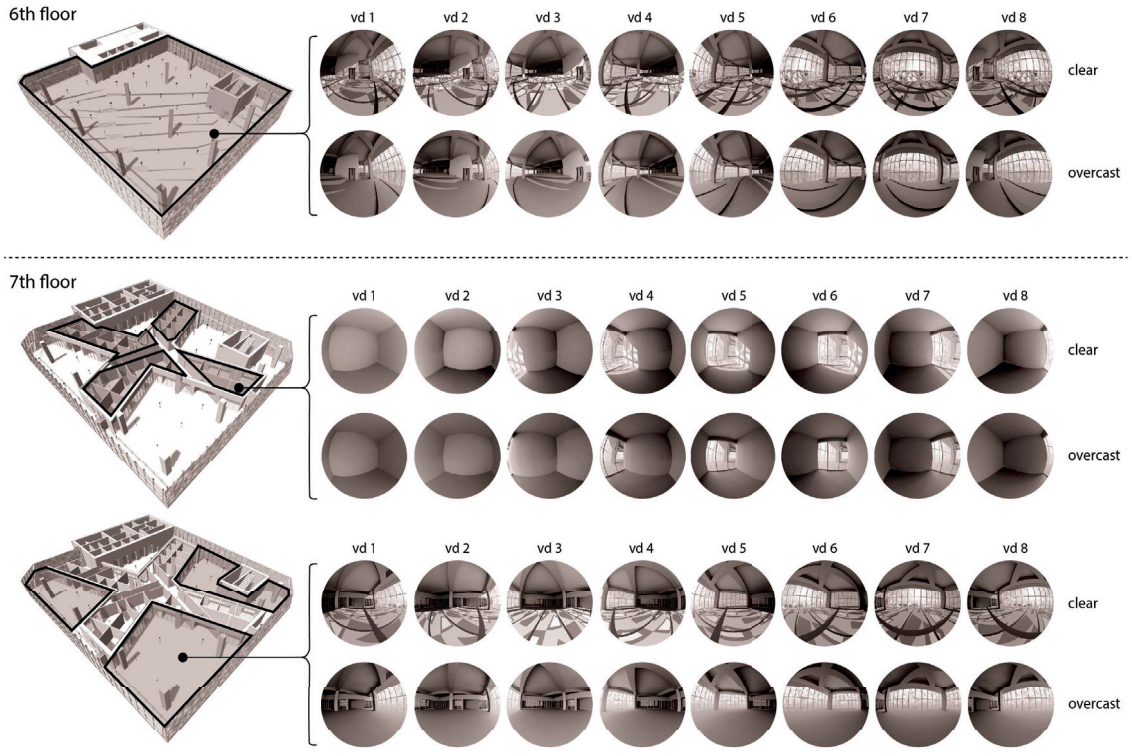


Figure 7.3 Showing a snapshot of the interior for two modelled floors with a sample of hemispherical renderings for a selection of points. For each sensor point, a time-series of instances were rendered under each of two sky conditions: clear and overcast.

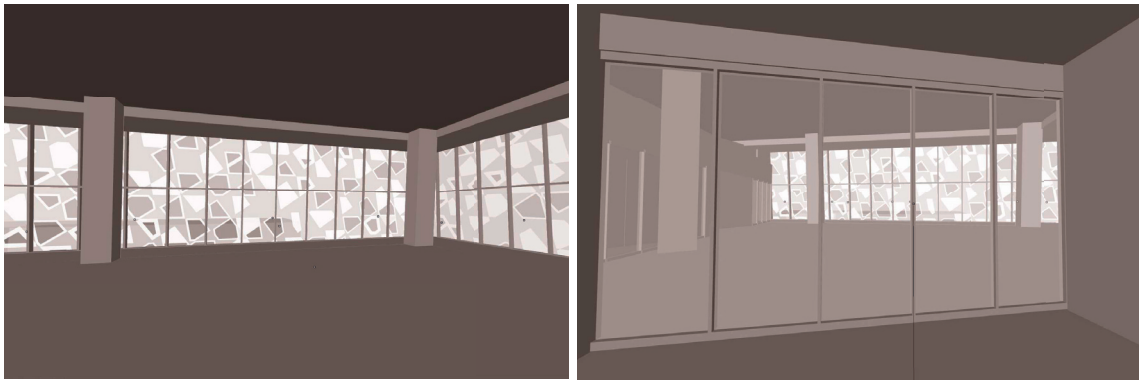


Figure 7.4 Showing two interior views in the 7th floor Rhinoceros model with re-constructed facade geometry and glass components.

natural illumination using high ceilings and open spaces with translucent frit to minimize direct sunlight and diffuse daylight deep into the study spaces. The large fritted pattern on the exterior glass façade produces strong visual effects under direct sunlight. This case study is relevant for the exemplification of our workflow due to the mix of interior typologies within the same exterior building envelope. The 6th and 7th floors offer vastly different interior architectures (from a fully open and sloped plan to semi-enclosed work rooms) which allow for the direct comparison of visual interest and non-visual health modules for the same sky and time series. While illuminance sensors can pick up localized changes in this pattern of direct and diffused sunlight on a horizontal plane, an immersive eye-level analysis is needed to determine the impacts of this pattern on visual impressions and health responses.

7.2.1 Selection of Spaces for Comparative Study

The Ryerson SLC contains 9 floors, each of which houses a mix of student and staff-related programs. To demonstrate the impacts of architectural form, orientation, and façade patterns on human-centric daylight performance, we decided to simulate a series of points across the 6th and 7th floors, which are dramatically different in plan. The 6th floor is composed of informal open study space across a series of ramped floor levels, stepping down in elevation towards the Southwest corner. Using a 5-meter spacing, the floor plan was divided into a grid of 59 points, offset a minimum of 2 meters from the façade and avoiding enclosed interior spaces like circulation cores or structural elements like columns. Figure 7.3 shows the layout of space, distribution of sensor points, and sample of rendered view directions from a select view position.

The 7th floor is split into a series of three open study spaces, flanked by two rows of enclosed study rooms, as seen in Figure 7.3. All enclosed study rooms look out onto the open study spaces through an interior glazed wall. While many of these spaces require electric light at all times of the day due to insufficient natural illumination deep within the floor plate, we decided to simulate only the daylight to evaluate the carbon-neutral potential for interest and health-related lighting

performance. Using the same spacing as for the 6th floor, with the addition of points at the center of each enclosed study room, we established a grid of 76 points.

7.2.2 Translation of Geometry Model into RAD Format

The geometry of our selected case study was received from the architects in DXF format (exported from Revit) and imported into Rhinoceros. After re-grouping layers by material definition, the glass façade was rebuilt to make frit patterns into individual glass objects so that patterns could be read for their material transmission values (Figure 7.4). We defined a 5 m grid of points at eye level (1.21 m from the floor while sitting) and exported those point locations as a text file.

7.2.3 Selection of Moments & Views

Using the DIVA toolbar, we exported our model and material definitions in the RAD format and imported them into the platform. We selected 28 semi-annual instances (7 hourly instances on each of January 13, February 28, April 14, and May 30) to give us a symmetrical half year. A finer grain or less frequent time-series could be selected depending on the aim of the analysis. This time series was developed from the Lightsolve method which uses 56 symmetrical instances to interpolate an annual temporal profile of illuminance-based data (Kleindienst, et. al., 2008). We chose 8 view directions per view position (135 points between the 6th and 7th floors) to conduct a series of illuminance and luminance-based simulations as inputs for our mSC_5 and nvR_D models.

7.2.4 Interpretation of Performance Modules

For each view position and view direction (60, 480 in total = 135 view positions x 8 view directions x 28 instances x 2 sky conditions) we applied the mSC_5 algorithm to rendered hemispherical images. For the same view positions and directions, the cumulative response R_D was also calculated using effective irradiance (Amundadottir, 2016).

Visual Interest

Based on the model proposed in Chapters 5 & 6, two threshold values were used to categorize results of the mSC_5 algorithm into three categories (see Figure 7.5):

1. Calming, $mSC_5 < 6.96$ (shown in cyan)
2. Neutral, $mSC_5 > 6.96, < 11.75$ (shown in grey)
3. Exciting, $mSC_5 > 11.75$ (shown in magenta)

Results in this chapter are presented as both the average daily mSC_5 achieved across each view direction and also as instantaneous moments. The average daily mSC_5 allows the user to visualize a compact overview of the perceptual impact on each view direction, while the instantaneous hourly results allow him/her or see how those results vary over time. In general, the work presented in this thesis has been critical of daily averages as they collapse hourly variations in

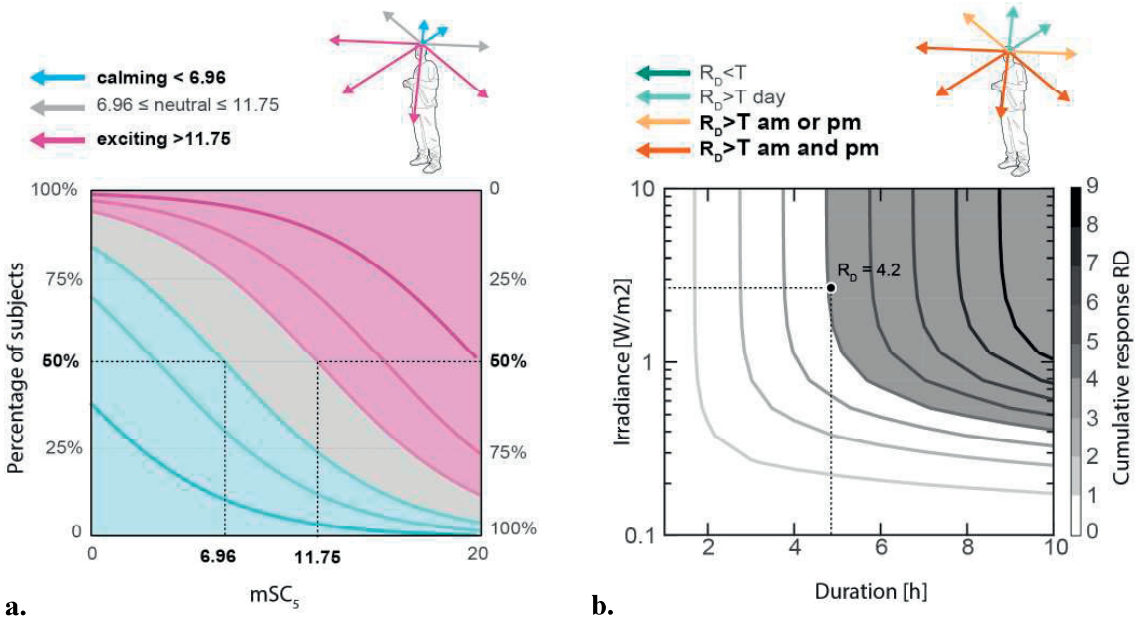


Figure 7.5 Diagram showing the thresholds used to drive performance for each metric with a) mSC_5, which can be applied instantaneously or as a daily average and b) R_D, which is applied in 4 categories, based on duration and time of day.

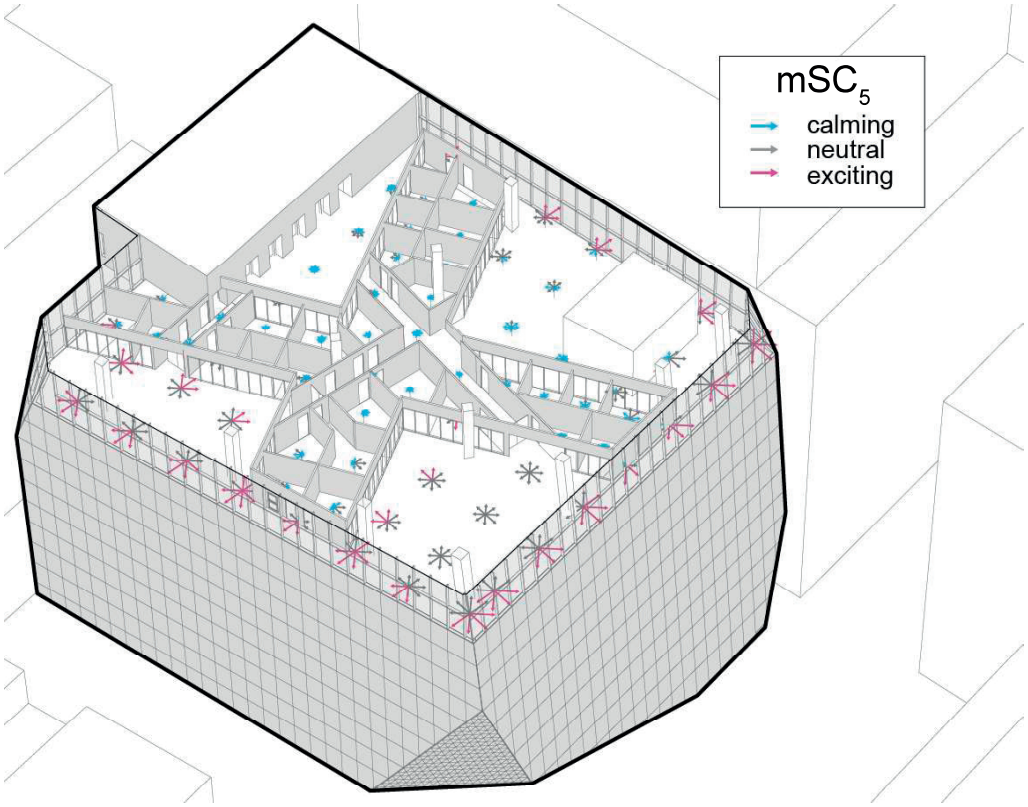
daylight, the dynamics of which are important to understand, but can still be useful when you want an overview of which view directions achieve generally higher or lower values over time. Results for each view direction are shown using coloured arrows, with the length determined by the output of the mSC_5 model and the color (cyan, grey, and magenta) determined by the thresholds listed above.

Non-visual Health Potential

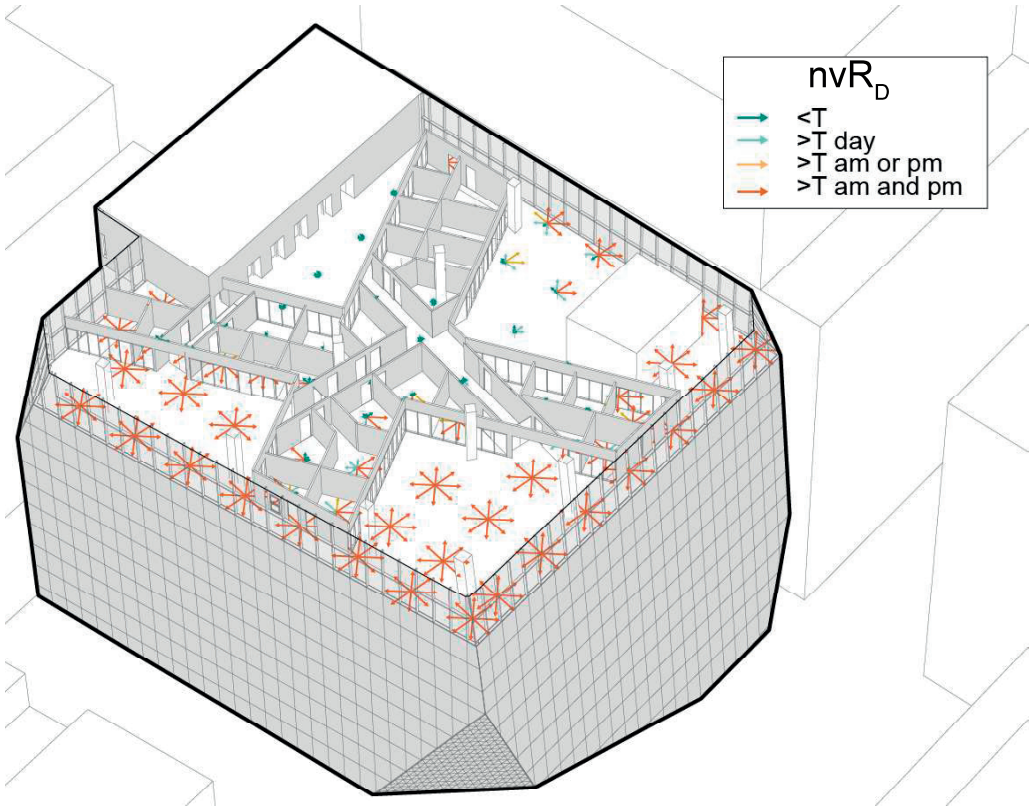
As described in Amundadottir (2016), a threshold value of T = 4.2 was used to categorize the resulting daily cumulative responses R_D into four categories (see Figure 8.5):

1. Not achieved during the day, $R_D < T$ (poor – shown in dark green)
2. Achieved during the day, $R_D > T$ (fair – shown in light green)
3. Achieved am or pm, $R_D > T$ (good – shown in yellow)
4. Achieved both am and pm, $R_D > T$ (excellent – shown in orange)

Achieving $R_D > T$ over the period of full day is not necessarily considered sufficient. The goal of T = 4.2 can be achieved during mornings or afternoons only if the duration of the solar day is sufficient, since the nvR_D model depends on duration. By binning the results into am and pm the user can better understand the influence of time of day and how it can affect the accumulation of dose received. The binning of the data should be adjusted to every case study depending on specific program use. Results are shown with colored arrows in each view direction, with length indicating the magnitude of R_D and the color indicating the threshold it falls into (dark green, light green, yellow, or orange). The desired performance is to achieve $R_D > T$ before noon (am) or/and after noon (pm).



a.



b.

Figure 7.6 Axonometric views with overlaid results for a) mSC_5 on April 14 under clear sky condtions and b) nvR_D under clear sky conditions.

7.3 Results

The results of this exemplary analysis are shown at both building scale and occupant scale. Figure 7.6 shows two axonometric views of the 7th floor with a) average daily mSC_5 and b) daily dose predictions for nvR_D on April 14 under clear skies. The aim of showing daily averages is to give a synthetic summary of the relative spatial impacts on performance. We can immediately see that sensor points within the enclosed interior study spaces receive much lower predictions of mSC_5 and nvR_D as compared to those points on the perimeter, adjacent to the glass façade. Figure 7.7 shows a zoom in for one view position on the 7th floor (in a closed private study room) with

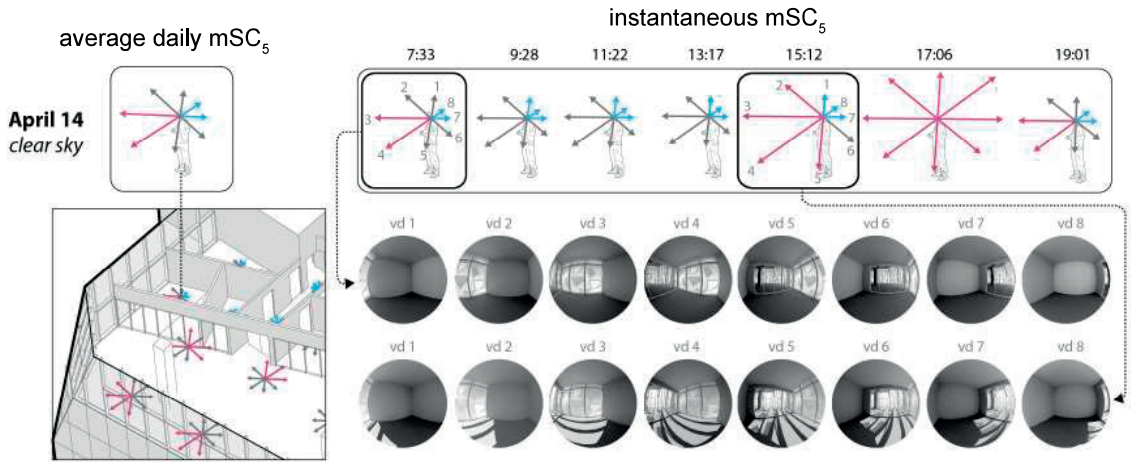
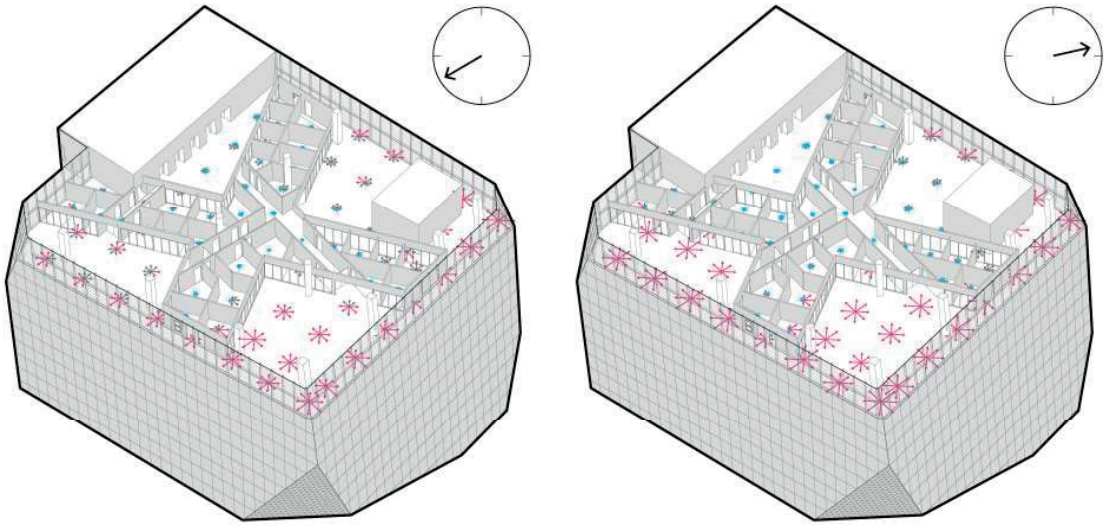


Figure 7.7 Instantaneous mSC_5 results plotted across 8 view directions and 7 hourly moments on April 14 under clear sky conditions.



a. **b.**

Figure 7.8 Instantaneous mSC_5 results plotted across all view positions for January 13 in a) the morning and b) the afternoon.

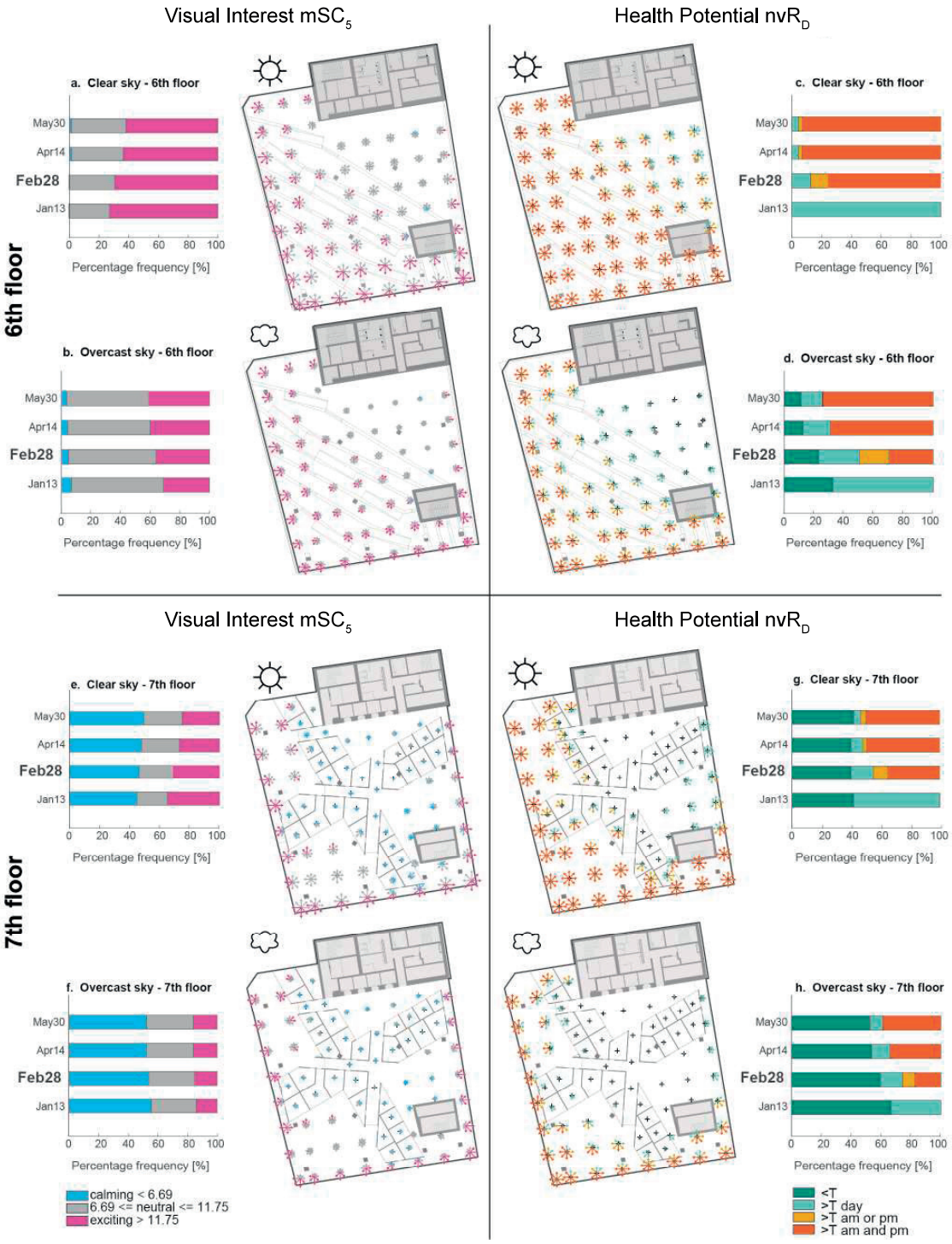


Figure 7.9 6th floor average daily mSC_5 on Feb 28 under a) clear and b) overcast sky conditions and nvR_D under c) clear and d) overcast sky conditions. 7th floor average daily mSC_5 on Feb 28 under e) clear and f) overcast sky conditions and nvR_D under g) clear and h) overcast sky conditions.

daily average and instantaneous predictions of mSC_5 across the day. This allows the user to see when, during the day, specific view directions provide strong visual effects and how they might vary over time. Figure 7.8 shows average daily mSC_5 and daily dose predictions for nvR_D on February 28 for both the 6th and 7th floors under clear and overcast sky conditions overlaid in plan. Figure 8.8 also shows a frequency distribution of model predictions for each floor under clear and overcast sky conditions for each of the 4 days included in our analysis.

7.3.1 Visual Interest mSC_5

If we look at the results for February 28, the 6th floor shows mostly exciting predictions for average daily mSC_5 under clear sky conditions (Figure 7.9a), with a slight shift toward more neutral predictions under overcast skies (Figure 7.9b). The 7th floor shows a significant shift towards neutral and/or calming predictions under both sky conditions (Figures 7.9e & 7.9f), with the peripheral open-study spaces achieving much more exciting predictions than the closed study rooms on the interior.

While we can draw certain spatial conclusions from the overall daily averages in Figures 7.6 and 7.9, some of which are fairly intuitive, Figures 7.7-7.8 show a more nuanced overview of the dynamic hourly experience within a single view position. In figure 7.7 (April 14, clear sky), the prediction for excitement shift dramatically across each view direction depending on the time of day and resulting sun position. Figure 7.8 shows two instantaneous predictions for a) the morning and b) afternoon on January 13 across all points on the 7th floor. In this comparison, the spaces along the West facade see a noticeable increase in exciting predictions in the afternoon instance as compared to the morning instance. These hourly and side-by-side instances illustrate the variable nature of perceptual performance as shifting sun positions from one moment to the next can alter the evaluation of excitement in our visual field.

Depending on the intended program use and qualitative ambiance in an architectural space, hourly performance predictions can be useful for the designer to know *when*, over the course of the day and year and *where*, within the field-of-view, impacts of daylight are likely to alter an occupant's emotional state. In some locations and instances, the strength and dynamics of these effects may be celebrated, while in others they may be seen as undesirable. While this thesis does not propose a set of guidelines for when, where, and for whom predicted visual effects may be desired, future work must address the link between visual interest, programmatic use and design intent.

7.3.2 Health Potential nvR_D

Figure 7.9 (c,d,g & h) shows the daily cumulative response R_D across each viewpoint in plan on February 28. The results for the 6th floor show excellent performance throughout the space under clear sky (Figure 7.9c). Lower light intensities caused by overcast sky conditions show reduced performance in view directions facing North and East, where viewpoints did not exceed

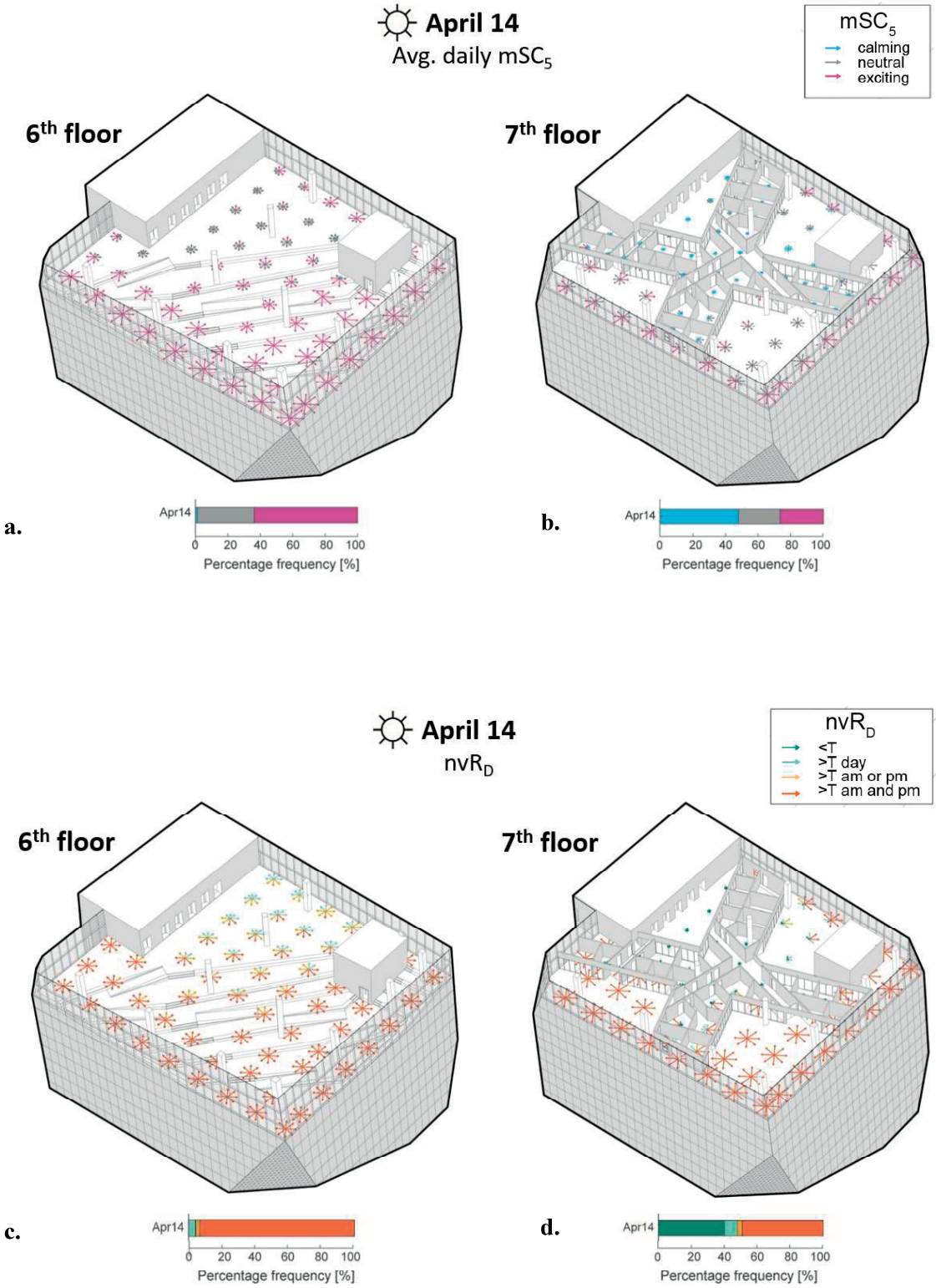


Figure 7.10 mSC_5 results plotted across all view positions under clear sky conditions on April 14 for the a) 6th floor and b) 7th floor. nvR_D results plotted across all view positions under clear skies for April 14 on the c) 6th floor and d) 7th floor.

the desired performance (Figure 7.9d). During solar hours for each of the 4 days the percentage frequency of the 4 performance categories (Section 7.2.4) is counted for all viewpoints in the scene and displayed using stacked bar graph.

The results for the 6th floor under clear sky conditions are similar for most days of the year except during the darkest winter months, where the number of solar hours limits the overall health potential. As expected, the performance under overcast skies is reduced. Interestingly, the category of achieving T = 4.2 during am or pm increases on February 28, which means that achieving non-visual health potential is more sensitive to timing of occupation for overcast than clear skies around this time of year. This can be explored in more details by analyzing different time periods of the day.

The results for the 7th floor are less spatially homogeneous compared to the 6th floor, which is explained by the partition of the 7th floor into open and enclosed study rooms. Less than half of the simulated points belong to open study spaces (30/76) resulting in a lower overall performance for the 7th floor. Under clear sky conditions on February 28 good and excellent performance is achieved for 45% of the viewpoints (Figure 7.9g), which reduces to 22% under overcast conditions (Figure 7.9h). As seen in Figures 7.9g and 7.9h, enclosed study rooms receive much less light than open study spaces. Under overcast sky conditions almost none of the viewpoints in enclosed spaces (ca. 1%) achieve the desired performance, while it is achieved under clear sky conditions in the enclosed rooms facing South and West.

7.4 The Impact of Architecture

Figure 10 illustrates a side-by-side comparison of average daily mSC_5 and nvR_D for the 6th and 7th floors on April 14 under clear sky conditions. The Ryerson Student Learning Center presents an interesting case study opportunity for human-centric daylight analyses, as the variation in interior space typologies between floors allows us to compare the impact of architectural design directly. In the frequency plots for Figure 7.10a & b, we can see a substantial shift in the frequency of points that fall above the exciting threshold, from over 60% on the 6th floor to just over 25% on the 7th floor. The small, semi-enclosed study rooms that flank each of the open corners on the 7th floor produce mostly calming predictions for mSC_5, whereas only a few view directions on the 6th floor ever achieve a calming prediction. Depending on the intended use and design intent for these study spaces, daylight is can be expected to produce a very different ambiances for the occupants on each floor.

Figure 7.10c & d also show a dramatic shift in health potential. The frequency of points which achieve the desired R_D for both morning and afternoon is over 90% on the 6th floor and only 50% on the 7th floor. If the occupants rely on daylight alone, many of the enclosed study spaces on the 7th floor will not achieve the recommended R_D. The enclosed rooms on the North and

East facades achieve particularly low performance values for nvR_D where sunlight is blocked by the adjacent building and view positions deep within the floor plate are not exposed to adequate daylight levels.

Coupling frequency plots alongside a plan (Figure 7.9) or axonometric (Figure 7.10) view for each performance model allows for a quantitative and simultaneously spatial overview. Where the radial plots introduced in Chapter 6 allow for a detailed understanding of performance across multiple view directions, this method of visualization was difficult to read across many view positions. The colored arrow diagrams developed in this chapter allow for a more synthetic overview of performance across many view positions, where the output of each metric is abstracted to a color and a length for each view direction. Future work is needed to

7.5 Chapter Summary

This chapter has introduced an automated work flow to implement two human-centric performance modules in daylight perception and health potential across a series of view positions and view directions within an architectural case study. Developed to automate the process of simulating human-centric performance predictions across a range of view positions and view directions within a given geometry model, the proposed platform creates a flexible platform for simulation with a new method of visualizing results. While past studies have implemented a similar work flow on a single view position in space (as described in Chapter 6), this chapter offers a more flexible protocol, allowing for the analysis of many view positions in a computationally efficient work flow. While the speed of simulations is significantly improved through shared ambient files, data storage is minimized by keeping only the illuminance and luminance data necessary for the computation of desired performance modules. Visualizations of the data presented in this chapter, while preliminary, provide both spatial and user-centric insights regarding the daylighting performance of visual interest and non-visual health potential. Ongoing work will seek to further develop the visualization strategy presented here to create an immersive user-experience.

Strategies for synthesizing and communicating complex information across many points, instances, performance modules, and sky conditions is a subject on ongoing conversation in the development of this work. As both the content of the modules, their integration alongside other performance metrics, and immersive viewing techniques are implemented further into software packages, the work presented in this chapter and Chapter 6 open an exciting new avenue of applied research.

8 A VISUALLY-IMMERSIVE EXPERIMENT ON VISUAL INTEREST USING VIRTUAL REALITY

The final phase of this thesis will present the results of an experiment where subjective impressions of visual interest were collected in immersive daylit scenes using an Oculus Rift CV1 virtual reality headset. In chapter 6, hemispherical fisheye renderings were produced in 20° radial increments across a 360° view range and then analysed using an adapted mSC_5 prediction algorithm developed in Chapter 5. In chapter 7, that same prediction model was applied to an array of view positions throughout an architectural case study alongside a model for non-visual health potential to exemplify the potential for *spatially-immersive* evaluations across many view directions. In this chapter, we will introduce a *visually-immersive* experimental approach using a virtual reality headset where subjects are divided into two groups and asked to rate either hemispherical or fully immersive renderings under varied sky conditions.

In this experiment, angular fisheye HDR renderings *-vta* were produced using a 180° and a 360° image format to collect subjective evaluations of daylit scenes from two subject groups (one that saw the 180° scenes and one that saw the 360° scenes). To test the effect of sky condition and view direction from a fixed view position within the architectural scenes, each subject, regardless of group, saw a random presentation of sky and view conditions for each space. Subjective ratings for a series of semantic uni-polar scales were then collected and compared to the algorithms presented in Chapters 4 & 5. The aim of this experiment was to test whether the image-based algorithms like mSC_5 could be used to predict subjective impressions of visual interest in visually-immersive scenes. Does this work with a hemispherical image format? A cube-face format (an image type used to project 3D scenes in virtual reality)?

Using the projected 360° and 180° renderings, this experiment allowed for the gathering of two data sets: 1) responses to qualitative daylight characteristics from within a semi-immersive 180° view range and 2) subjective responses to qualitative daylight characteristics from within an immersive 360° view range. In addition to more realistic visual immersion, the use of virtual reality also allowed for the extraction of head-tracking data, providing additional insight into how people perceive different parts of the immersive scenes. Subjective ratings collected from within these two data sets were then compared to quantitative predictors to validate the use of different image-based algorithms in predicting impressions of visual interest across space and under varied sky conditions.

8.1 An Immersive Display Approach

For the experiments introduced in this chapter, the Oculus Rift CV1 VR headset will be used to display visually-immersive architectural scenes. This headset has a 110° field of view display, using OLED panels with a resolution of 1080×1200 pixels per eye and a refresh rate of 90 Hz. The maximum measured luminance of the display is 80 cd/m2 for a white scene (RGB 255, 255, 255). Although the luminance range of the display is expected to be much broader, this measurement is explained by the low persistence mode of the display, which turns the pixels off between frames to alleviate motion blur, resulting to a lower perceived luminance in the duration of a second.

Renderings for this experiment were generated using Radiance and tone-mapped to the measured luminance range of Oculus Rift CV1 with calibrated gamma correction (0 – 80 cd/m2, 1.8 respectively) using *pcond* (Ward, 1997) with the contrast sensitivity function *-s*. In Section 7.1, the author explained the process of extracting an angular fisheye image from a spherical rendering, based on user-defined view parameters using *pinterp*. We used a similar approach in this chapter to build scenes for the Oculus Rift. To project scenes in the Oculus Rift CV1, we extracted a series of six rectangular renderings *-vh* 90 *-vw* 90° (facing north, south, east, west, top, and down) from a single 360° angular fisheye *-vta* rendering. These six 'cube faces' were then combined in Unity, the software selected for use with Oculus, to create a seamless and visually-immersive virtual scene, allowing subjects to freely explore the daylit space in all view directions.

8.2 Simulation Workflow

The following section will introduce the selection of case studies and the creation of renderings used in this experiment. Eight architectural spaces, selected to represent a range of architectural designers, spatial qualities, and interior daylight conditions were modelled in Rhinoceros and rendered in Radiance to generate 360° angular fisheye HDR renderings across a 28 step semi-annual time series. The modified spatial contrast (mSC_5, otherwise known as RAMM5) algorithm was then adapted to the 'cube face' format and applied to this rendered time-series in order to select instances of relative high and low mSC_5 under both clear and overcast sky conditions. Using the adapted 'cube face' mSC_5 algorithm to select high and low instances of excitement over time, the author then used the hemispherical adaptation of mSC_5 (introduced in Chapter 6, Figure 6.1) to evaluate a range of 18 view directions and select the directions that correspond to the highest and lowest mSC_5 values from within the full 360° field-of-view. Using those view directions, 'cube face' projections were then extracted using *pinterp* to project seamless 180° or 360° scenes in the Oculus Rift CV1 virtual reality headset. Subjects in the 360° group entered each space in the high mSC_5 view direction and were then allowed to explore the entire scene under either clear or overcast skies (randomized for each space). Subjects in the 180° group entered each space in either the high or the low mSC_5 view direction (under clear or overcast skies) and were then allowed to explore the semi-immersive scenes.

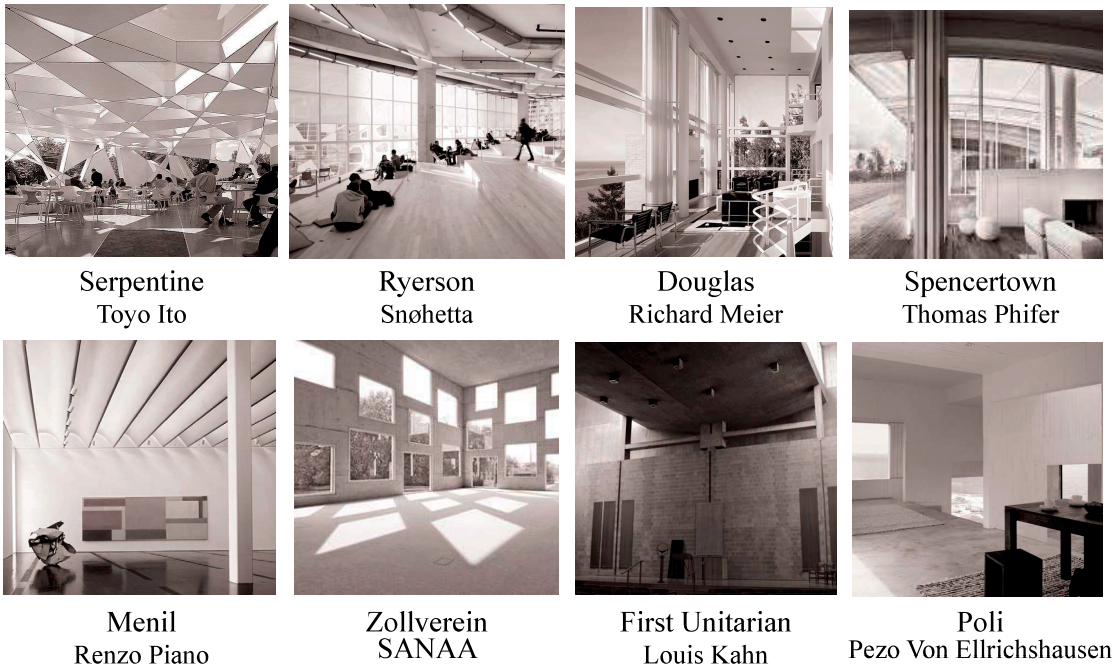


Figure 8.1 Eight architectural spaces selected for use in the experiment: the Serpentine Pavilion by Toyo Ito, Ryerson Student Learning Center by Snøhetta, Douglas Residence by Richard Meier, Spencertown House by Thomas Phifer, Menil Gallery by Renzo Piano, Zollverein School of Management by SANAA, First Unitarian Church by Louis Kahn, and Poli House by Pezo Von Ellrichshausen. See the image credits page.

8.2.1 Selection of Case studies

For this experiment, a range of architectural spaces were selected based on their internal daylight composition, from direct and exaggerated sunlight penetration to diffuse and uniform daylight conditions. For the selection of spaces, the author considered a range of conditions: daylight distribution (direct, diffuse, varied), architectural style, latitude, and program use. Regarding daylight composition, spaces were selected to cover a range of typically high and low contrast daylight conditions. The final selection of spaces for this experiment is shown in Figure 8.1. Selected spaces include the Douglas Residence by Richard Meier, the Serpentine Pavilion by Toyo Ito, the Ryerson Student Learning Center by Snøhetta, the Spencertown Residence by Thomas Phifer, the Zollverein School of Management by SANAA, the Poli House by Pezo von Ellrichshausen, the Menil Gallery by Renzo Piano, and the First Unitarian Church by Louis Kahn.

8.2.2 360° HDR Renderings

All selected case studies were modelled in Rhinoceros to a consistent level of detail for structure, façade and fenestration components, interior partitions, and fixed elements such as railings. Figure 8.2 shows a photograph of the Zollverein School of Management by SANAA on the left and a 3D model on the right, built in Rhinoceros with separate material layers for floor, wall, structure, and glazing components. Removable interior artefacts such as furniture, lighting

Figure 8.2 Showing a photograph of the Zollverein School of Management by SANAA on the left, with a 3D Rhinoceros model on the right. Digital models were generated based on available documentation and were created to represent visibly accurate scenes. See the image credits page.

components, and people were intentionally excluded to minimize elements that were not part of the built architecture. Material textures and fine surface details were also excluded to economize on modelling and rendering time. While some spaces were easy to render with a highly accurate level of detail, others would have required a serious time investment, and thus a consistent rather than photo-realistic level of detail was considered a priority by for this study. For those models that we received in part or in full by the architects (such as the Ryerson Student Learning Center by Snøhetta), components that were not considered 'structural,' such as suspended light fixtures, were removed and facade components were re-built to allow for accurate material assignments (in this case, the exterior facade and interior glass partitions were fully remodelled).

A central view position was established in each space, equal distance from exterior walls (if possible, otherwise centered within the largest unobstructed area of the space) and at eye level (1.65 meters from the floor) to represent a human perspective while standing. Geometry models were then exported as Radiance files using the Diva-for-Rhino toolbar, which provides outputs for rad geometry and material files. Material selections were made based on default reflectance values for wall, double glazed window, floor, ceiling, and fixed components, except where those elements were clearly higher or lower in reflectance (such as the Spencertown residence where walls, structure, and ceiling elements are all painted in the same high- reflectance paint). In this case, a 70 percent ceiling reflectance was applied to all those elements uniformly. A list of these material descriptions can be seen in Table 8.1.

The selected architectural case studies were rendered in two phases using the Radiance lighting simulation software (Ward, 1994). In the first phase, each scene was rendered at an intermediate level of accuracy (Table 8.2) across 28 symmetrical semi-annual instances (Figure 8.3b) under clear sky conditions using a 360° angular fisheye view setting (*-vta*) as can be seen in Figure 8.3a These 28 moments were adopted from the Lightsolve method developed by Kleindienst et al.

(2008), introduced in Chapters 4 and 5. As both clear and overcast conditions were simulated, the symmetrical sun path allows us to get a representative series of moments from only half of the 56 instances originally proposed by Kleindienst et al. (2008).

Table 8.1 – A list of Radiance materials used in each simulation

Elements	Material Description
Zollverein School of Management	
Walls	Diffuse reflectance of 50%
Ceiling	Diffuse reflectance of 70%
Floor	Diffuse reflectance of 20%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
Ryerson Student Learning Center	
Walls	Diffuse reflectance of 70%
Ceiling	Diffuse reflectance of 50%
Floor Light wood	Diffuse reflectance of 50%
Floor Carpet	Diffuse reflectance of 20%
Glazing Clear	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
Glazing Frit 1	Translucent glazing with visual transmittance of 70%; visual transmissivity of 90%; specular transmissivity of 90%
Glazing Frit 2	Translucent glazing with visual transmittance of 70%; visual transmissivity of 60%; specular transmissivity of 30%
Serpentine Pavilion	
Steel Structure & Panels	Diffuse reflectance of 70%
Floor	Diffuse reflectance of 35%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
Douglas & Menil	
Walls, Ceiling, Mullions	Diffuse reflectance of 70%
Light Wood Floor	Diffuse reflectance of 50%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
Spencertown Residence	
Structural Walls & Columns	Diffuse reflectance of 70%
Floor & Dark Wood Walls	Diffuse reflectance of 20%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
Poli House	
Painted Walls & Ceiling	Diffuse reflectance of 70%
Mullions & Light Wood Walls	Diffuse reflectance of 50%
Floor	Diffuse reflectance of 35%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%
First Unitarian Church	
Walls, Floor & Ceiling	Diffuse reflectance of 35%
Light Wood Elements	Diffuse reflectance of 50%
Glazing	Clear glazing with visual transmittance of 80%; visual transmissivity of 87%

Table 8.2 – Rendering parameters for 28 semi-annual renderings

dt	dj	ds	dc	dr	dp	st	ab	aa	ar	ad	as	lr	lw	pj	ps	pt
.05	0	.15	.75	3	512	.15	3	.1	512	4096	2048	8	.005	0	2	.05

Table 8.3 – Rendering parameters for final renderings used in the experiment

dt	dj	ds	dc	dr	dp	st	ab	aa	ar	ad	as	lr	lw	ps	pt
.05	0	.02	.75	3	512	.15	4	.025	512	4096	1024	8	1e-9	1	-

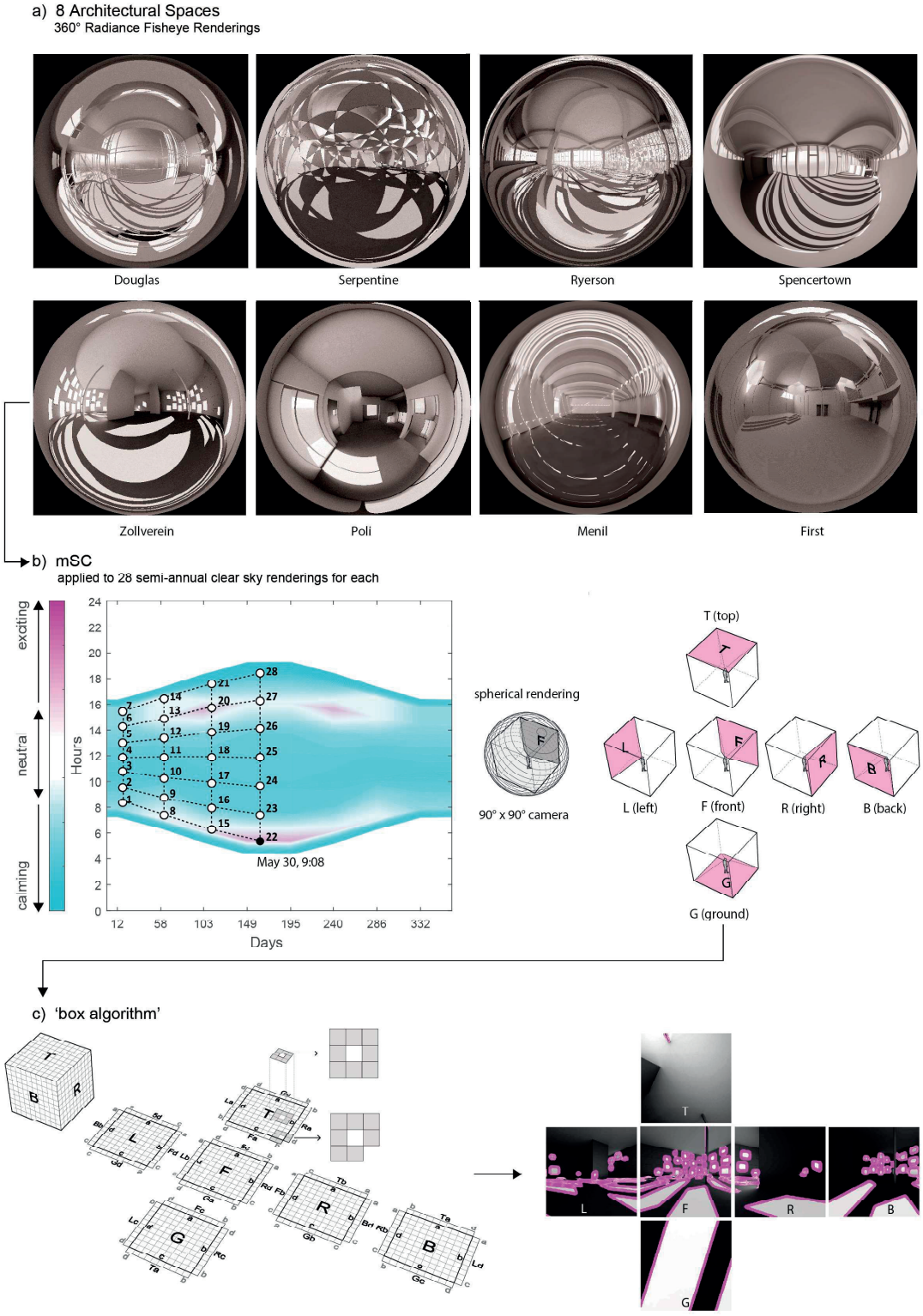


Figure 8.3 a) 360° angular fishese renderings of the 8 architectural spaces selected for this experiment, b) mSC_5 results across 28 semi-annual instances calculated from the average of six 90° x 90° box projections extracted from the full scene. The algorithm is implemented using a c) cube face adaptation with seamless edges and corners.

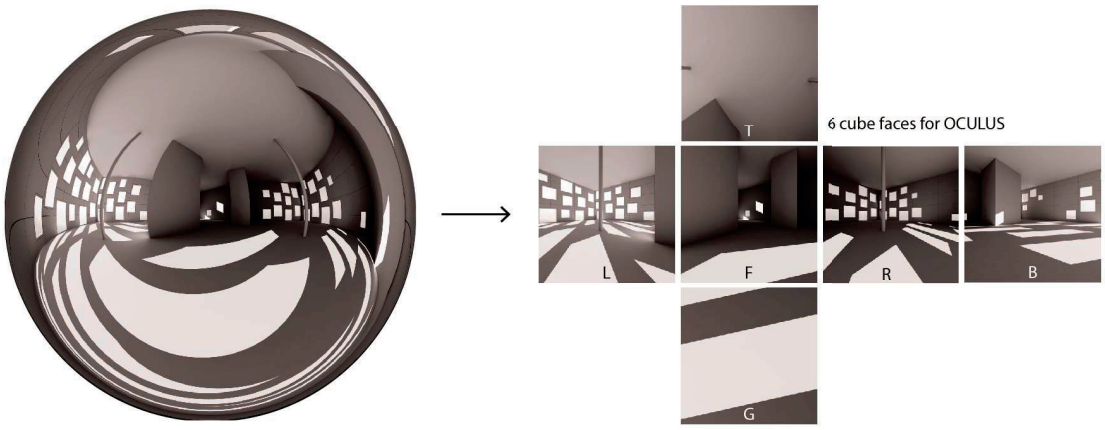


Figure 8.4 Showing one time step under clear sky conditions rendered from the central view position in the Zollverein School of Management in 360° angular fisheye (*-vta*) projection. On the right, this fisheye is extracted into six 90° x 90° cube faces, which make up the projection map in UNITY.

Each of the 28 angular fisheye renderings produced in this first rendering phase (Table 8.2) was tone-mapped using the *pcond* algorithm developed by Ward, et al. (1997) with the human contrast sensitivity function *-s.* While the literature suggests that other tone-mapping operators like Reinhard (2002) or Durand (2004) may be perceived as more realistic (see Section 1.3.3), we decided to use *pcond* as its native adaptation in Radiance allows for an accurate compression of luminance values using a fisheye image projection. Future work is needed to determine the impact of TMOs on perceived scenes in virtual reality as adaptive, view-dependent mapping provides a new challenge for existing algorithms which must transform with the users' view direction to smooth lighting between areas of the scene which vary in lighting composition. A discussion of this topic can be found in Section 9.2.2, where results for mSC_5 and RAMMG are compared using two tone-mapping functions as well as BMP vs. HDR compressions.

Relying on a method developed by Chamilothori et al. (2017), the tone-mapped 360° angular fisheye renderings were then transformed using *pinterp* to extract six 90° x 90° renderings (Figure 8.4), each corresponding to $1/6^{th}$ of the full scene. This set of 6 renderings was analyzed in Matlab using an adapted 'cube face' version of the mSC_5 algorithm (Figure 8.3c). Figure 8.5 shows how the author conceptually adapted the 2D mSC_5 algorithm to the 6 projected cube faces. Script A4.1

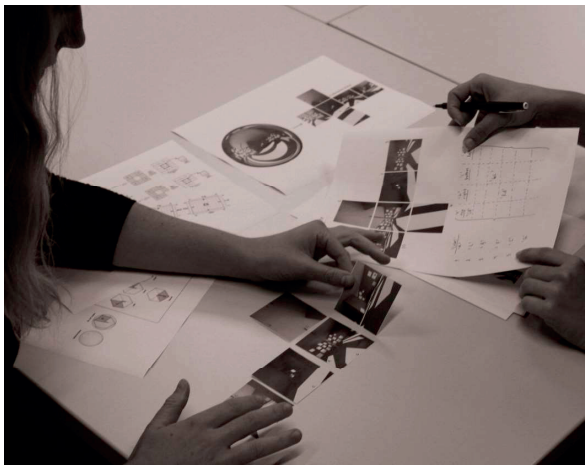


Figure 8.5 The process of developing our mSC_5 box algorithm from the original 2D algorithm described in Chapters 5-7.

(Appendix 4) shows this implementation in Matlab. As each face shares a virtual 'seam' with its neighbor (both top, bottom, left, and right), this algorithm was designed to address both edges and corners of the image. While conceptually straight forward, the implementation into a functional algorithm had to account for each face as a set of related faces, with the edge of each face sharing pixel neighborhoods with adjacent faces. A critical point in this procedure is that the 360° angular fisheye renderings were tone-mapped *before* the extraction of cube faces, as any compression of luminance values had to be consistent across the entire scene. If the faces are tone-mapped separately, seams between images become visible both in the virtual scene and in the application of mSC_5, creating contrast boundaries that do not exist within the scene.

To select instances to render under high quality radiance parameters and to show to subjects in this experiment, the box adaptation of mSC_5 was applied to the 28-step time series of cube-map projections for each space shown in Figure 8.3a under clear sky conditions. Instances of highest relative mSC_5 per time-series were identified (as shown in Figure 8.3b and for a selection of space in more depth in Appendix 4 - Figures A4.2 - A4.6) and re-rendered using high precision Radiance parameters for clear and overcast sky conditions for each of the selected architectural case studies (Table 8.3). The final 'cube face 'renderings (used to build the scenes in UNITY) can be found in Appendix 4 (Figures A4.9 - A4.16).

8.2.3 Selection of Hemispherical View Directions

As described briefly in the introduction of this chapter, subjects in this VR experiment were randomly placed into one of two possible groups; one that saw 180° scenes and one that saw 360° scenes. These distinct groups allowed us to collect subjective data from the whole scene as well as from specific view directions within that scene to evaluate how ratings of visual interest vary *across* the architectural scenes. To select the view directions for each of the 180° images presented to the semi-immersive group, an additional step in the simulation workflow was required. For each instance of clear and overcast sky selected using the procedure described in 8.2.2, a series of eighteen 180° angular fisheye projections were extracted using *pinterp* in 20° radial increments. The resulting 180° angular fisheye renderings were then analysed separately using the mSC_5 algorithm, this time using the adaptation for hemispherical image formats, to select the highest and lowest view directions based on the average mSC_5 in each view direction. Figure 8.6 shows an example of the four variations for one space: a) clear sky, high mSC_5 view direction, b) clear sky, low mSC_5 view direction, c) overcast sky, high mSC_5 view direction and d) overcast, low mSC_5 view direction. Figures A4.7-A4.8 in Appendix 4 show the high and low mSC_5 view directions selected for every space and sky condition.

8.2.4 Projection of Final Scenes

In order to project the final rendered scenes in Oculus, the author used a workflow developed by Chamilothori, et al. (2016), for the generation of immersive scenes. As briefly mentioned before,

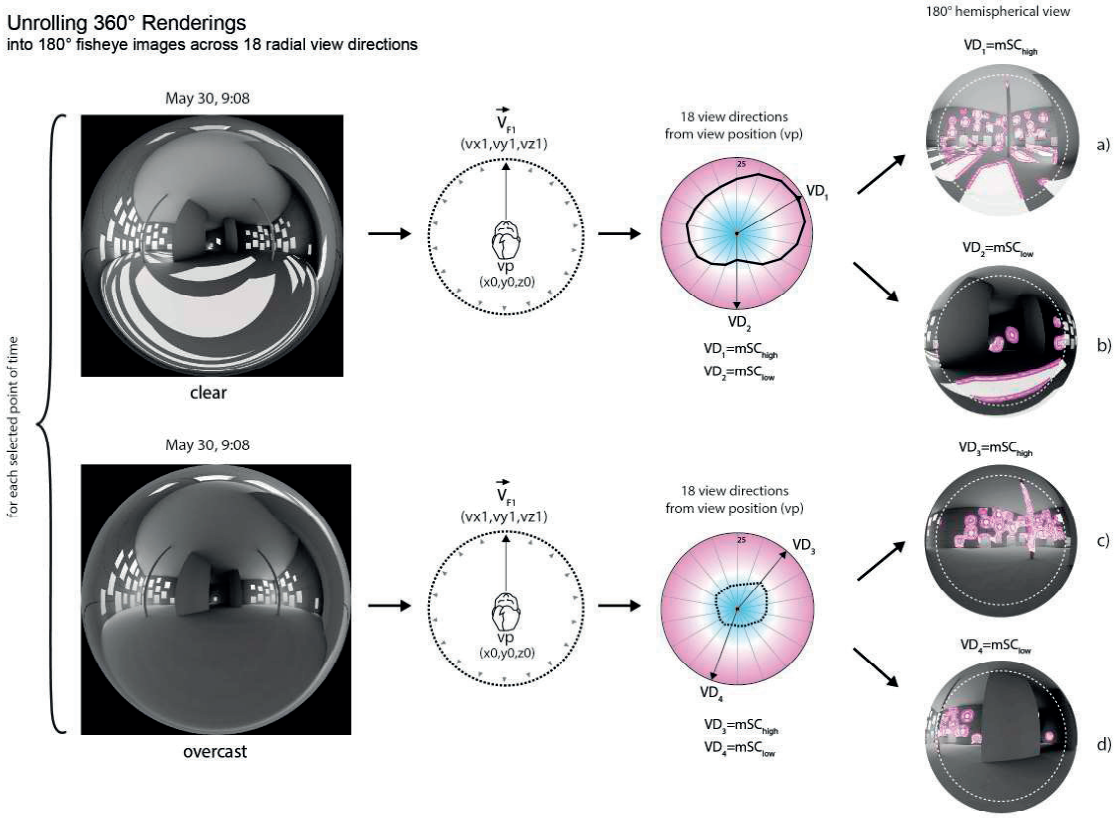


Figure 8.6 a) Semi immersive 3D scene, with 2 view directions selected per space and sky condition. Views were chosen based on the highest and lowest mSC_5 achieved across a ray of 20° radial view vectors, from the view position (vp). b) Fully immersive 3D scene, unrolled into 90° cube faces for projection in Oculus.

the scenes were created in the game engine Unity using the principle of cube map projection, which gives a seamless impression of 3D immersion to the scene observer. Although the projected images were not stereoscopic, as in this experiment the same image was projected to both eyes, the loss of 3D object perception was minimal due to the scale of each scene. Stereoscopic projection is most critical in scenes with objects close to the foreground of the observer. Future work may seek to implement a stereoscopic scene projection, although Banos, et al. (2008) found no significant effect of stereoscopy on ratings of presence or emotional response in the virtual environment. In addition enhanced visual-immersion, the virtual scenes projected in Oculus Rift CV1 allow for the collection of head-tracking data during each experimental session. This data allows for the analysis of recorded view behaviour within the different scenes and will be explained in more depth in Section 8.7.

Research conducted by Chamilothori, et al. (2017) compared subjective ratings of a test room with those of the same space rendered in Radiance and presented to subjects using the Oculus DK2, an older version of the headset used in our experiment. The objective of this study was

to determine the effect of an immersive virtual display on ratings of pleasantness, interest, complexity, excitement, and satisfaction with the amount of view to see if the virtual scenes could be used as a surrogate for daylit spaces to collect subjective impressions of daylight. However preliminary in scope and population size, there was no significant difference in ratings between the real and rendered scene for both clear and overcast sky conditions, suggesting that the images projected in VR could be used as a surrogate for real world experiments.

While some studies have suggested that local TMOs such as Reinhard, et al. (2002) may be more accurate for compressing HDR renderings to LDR screen luminances (Yoshida, et al., 2005), others have recommended global operators like Drago (2004) and *pcond* (Cadik, et al., 2008). In either case, we were unable to apply these algorithms to our 360° angular fisheye image format. Due to the image specification used in our workflow, non-native TMOs such as Reinhard, et al. (2002) could not mask the area beyond the hemispherical image (the black corners) and resulted in muddled grey compressions which appeared either too light or too dark. The *pcond* operator was ultimately selected because its native radiance adaptation for the fisheye image specification created more visually acceptable results. Future work by Chamilothori, et al. (forthcoming) seeks to determine the relative impact of different TMOs on subject impressions in the virtual space and their effect on the perceptual accuracy of the display device and method. Some of the architectural scenes rendered in this experiment showed large shifts in luminance strength and distribution across eighteen view directions, and a static TMO may have difficulty compressing the entire scene when it varies so strongly in light distribution. This topic will be discussed further Section 9.2.2.

8.3 Experimental Design & Procedure

This section will describe the experimental design used in our study, followed by the collection of subjective responses to qualitative daylight characteristics from participants immersed in the projected scenes.

8.3.1 Design of Experiment

The online experiment introduced in Chapter 5 used a 3 x 3 semi latin square to test the effect of space (high, medium and low groups) and clear sky conditions (sky 1, sky 2, and sky 3) on subject impressions of visual interest. This design was selected due to the use of a web-based survey, which necessitated a compact set of questions which could be answered quickly to encourage subject participation. Using this experimental design, each subject was only required to rate 9 of the 27 possible images produced for the study.

For the experiment introduced in this chapter, we used a fully randomized presentation of sky conditions and view directions (in the 180° group) for each space. While each subject saw all eight architectural spaces (Douglas, Serpentine, Ryerson, Spencertown, Zollverein, Poli, Menil, and First), the spaces was presented to subjects at random, using a randomized set of sky/view

conditions. In the 180° group, subjects saw one of four possible sky/view combinations per space (as determined in subsection 8.2.3 and shown in Figure 8.7), while subjects in the 360° group saw one of two possible sky conditions as shown in Figure 8.8 (entering the scene in the high mSC_5 view direction and freely explore the rest of the scene). Because the participants in the 180° group were only able to explore half the overall scene, we expected their impressions to vary depending on view direction and that collecting data between view directions would allow us to understand the effect of view on subject impressions throughout a space.

To collect subjective responses, we used two forms of verbal questioning: open response and semantic differential ratings. While scenes and sky conditions were randomly presented, we asked an open response question in the first scene for each subject to collect descriptive adjectives before subjects were influenced by the selected rating scales used in the rest of the experiment. For the purposes of understanding which words people use to describe lighting ambiance, we asked subject to begin by exploring the scene at their leisure then asked them to describe the lighting ambiance using 3 or more adjectives of their choice. These adjectives will be used in future work by the research team to refine the selection of semantic rating scales, but will not be presented in this thesis as it exceeds the scope of work.

After answering this first question, subjects were then asked to rate each scene using a 10-point unipolar semantic scale with verbal anchors on each end (1 - not at all to 10 – very). Subjects were asked to rate the space for how pleasant, interesting, exciting and calming it was and how diffuse and contrasted the light in the space was. These words were selected from two previous studies conducted by the research group (Chamilothori, et al. 2017; Rockcastle, et al. 2016).

While the online experiment presented in chapter 5 used bi-polar rating scales, we decided to switch to uni-polar scales in this experiment. This decision was made in part to understand the relationship between ratings of 'calming' and those of 'exciting' which were used as a bi-polar scale in the previous study and selected due to goodness of fit with mSC_5 predictions (Section 5.7.4). After the fact, we wanted to understand whether spaces could be both exciting and calming and to study the relationship between these adjectives in more depth. We also switched from a 7 to a 10-point rating scale to makes it easier for subjects to provide verbal responses, as people are generally more familiar with providing assessment values from 1 to 10 and previous work identified particular difficulty in the participant's ease of responding to a verbal questionnaire in VR with no visual reference to the scale (Chamilothori et al., 2017).

8.3.2 Subjects & Experiment Procedure

This virtual reality experiment was conducted at EPFL in October, 2016 over the course of three weeks and was approved by application to the EPFL Ethical Review Board. Subjects were unpaid volunteers who were recruited via email, social media, and printed posters (Appendix 4, Figure A4.19). The study took place in different seminar rooms around the EPFL campus as the semi-

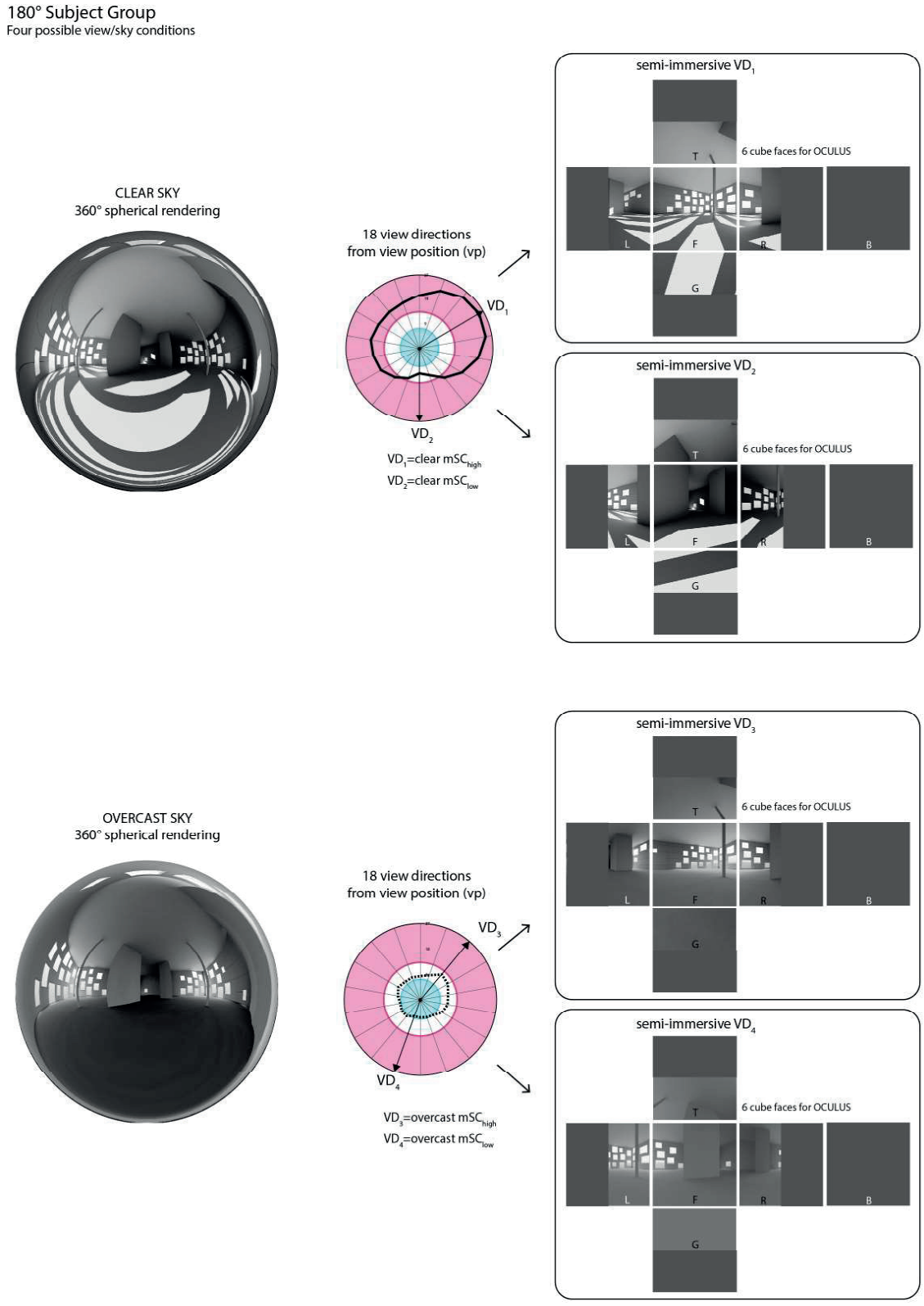


Figure 8.7 Showing the four possible scene parameters for the 180° group: $VD_{1 =}$ clear sky, high mSC_5 view direction, VD_2 = clear sky, low mSC_5 view direction, VD_3 = overcast sky, high mSC_5 view direction and VD_4 = overcast, low mSC_5 view direction

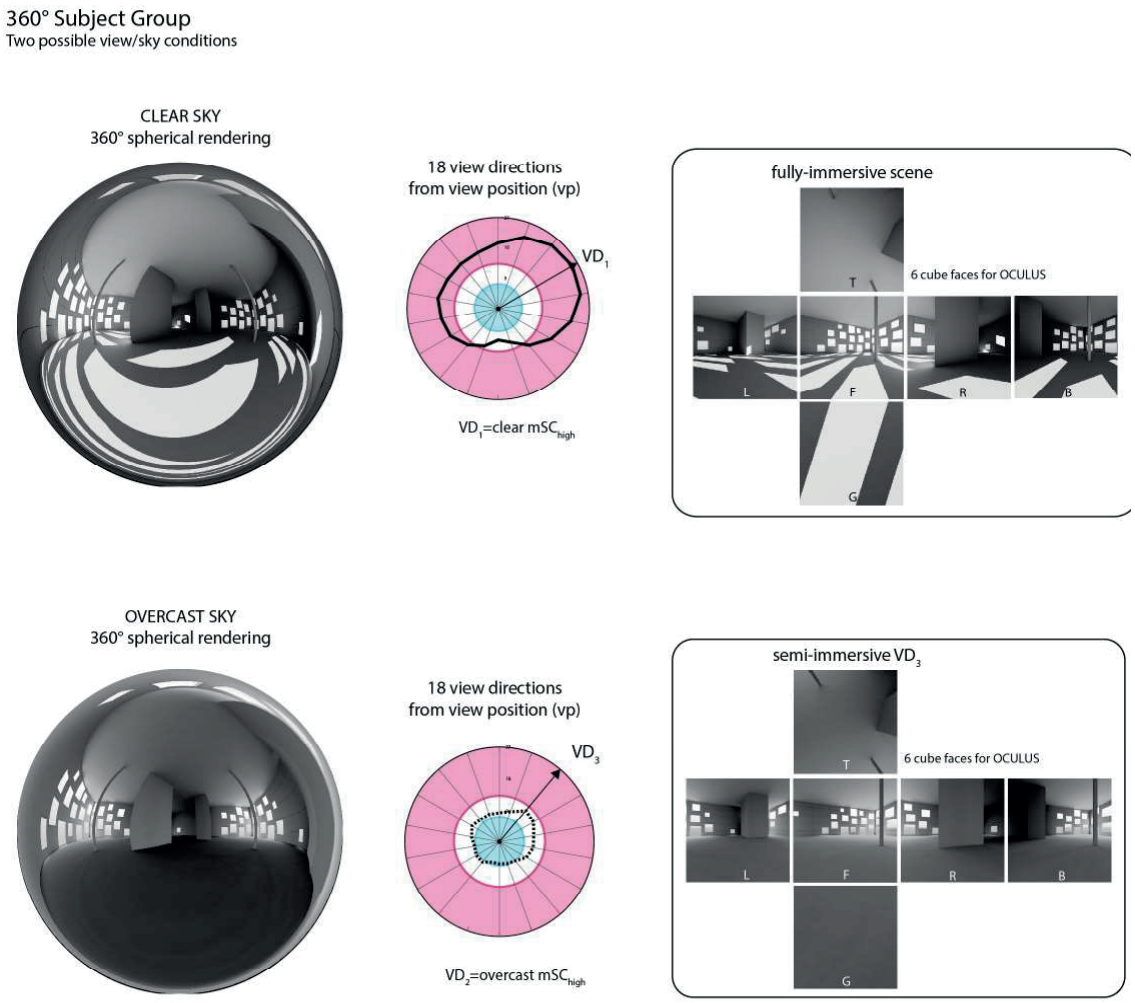


Figure 8.8 Showing the two possible scene parameters for the 360° group: VD_1 = clear sky, high mSC_5 view direction, and VD_2 = overcast sky, high mSC_5 view direction.

portable nature of this experiment allowed for easy set-up and access a larger population than a fixed laboratory experiment. The experimental equipment included the Oculus Rift CV1 and an Acer Predator 17-X laptop, capable of supporting the VR headset. The 109 subjects we recruited were between 18 and 50 years of age with a mean age of 29 (std=5.7 years, 30% female and 70% male) and were screened for English language capacity; eligible participants had a self-rated English proficiency of C1 or higher. They were asked to wear contact lenses or glasses, if needed, to ensure visual acuity. Each experimental session lasted roughly 20 minutes. Upon arriving for their scheduled appointment in one of the seminar rooms, subjects were asked to read an information sheet (shown in Appendix 4, Figure A4.7) about the experiment and sign a consent form regarding their voluntary participation. After this step, they were asked to respond to a series of demographic questions (shown in Appendix 4, Figure A4.18), the data from which will be explored in future work.

Upon arriving for their scheduled appointment in one of the seminar rooms, subjects were randomly assigned to a group, using a branch logic that placed 60% in the 180° group and 40% in the 360° group - determined based on the sample size needed for specific analyses. From there, subjects were asked to wear the virtual reality headset and adjust its fit in a training scene with the help of the researcher. They were told that the scenes they would see correspond to a field-of-view of either 360° or a 180° (depending on group assignment) and that they could turn their head, standing in a fixed position, to explore the space within these boundaries. This ensured that the scenes were perceived as immersive, as long as the participants rotated within these boundaries (Figure 8.9).

When they were ready, the participants were presented with each of the eight rendered spaces in randomized order under one of the possible scene parameters, two possible scenes for each space in the 360° group (Figure 8.8) or four possible scenes in the 180° group (Figure 8.7). After freely exploring the immersive scenes, subjects were asked to verbally respond to a series of ten-point unipolar scales on perceived characteristics in each scene. The presentation order for spaces and rating scales was random, being automatically dictated from the questionnaire and controlled by the researcher with the laptop's keyboard. The generated head tracking data log from Oculus Rift CV1 was saved after each session so that data could later be extracted and analysed for each participant and scene (presented in Section 8.7). At least 15 subjects rated each of the 4 conditions per scene in the 180° group and at least 20 subjects rated each of the 2 conditions per scene in the 360° group.

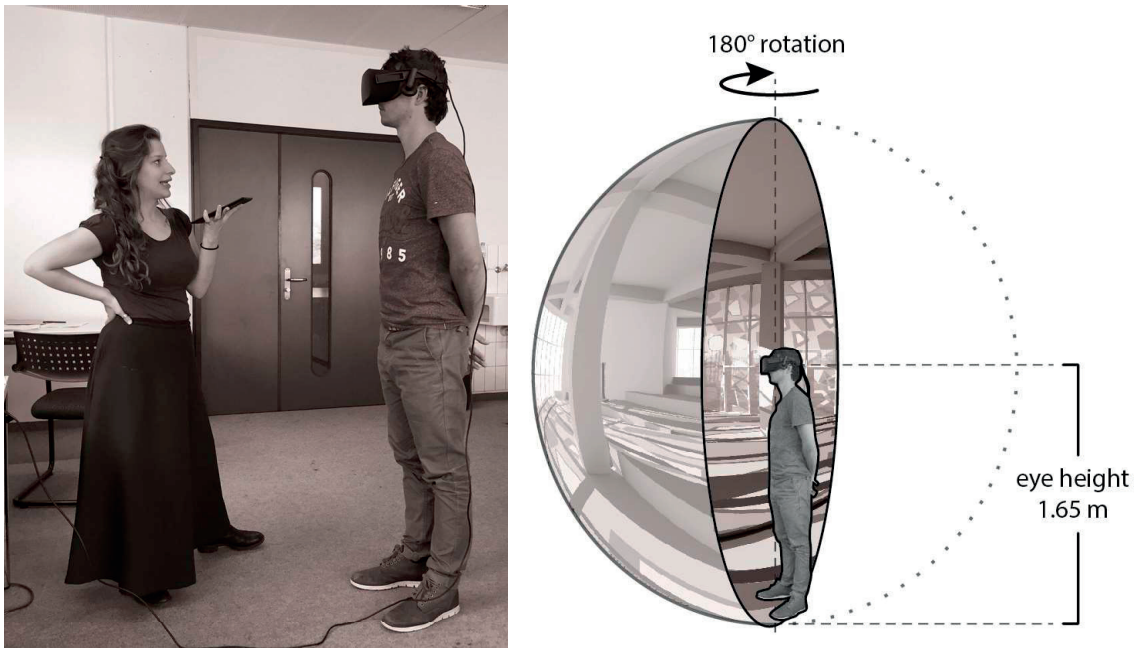


Figure 8.9 a) Image showing one of our researchers asking and recording subject responses from an anonymous test subject and b) a diagram showing the field-of-view displayed to the test subject within the Oculus Rift CV1.

8.4 Results for 180° Group

The results of this experiment will be presented in three parts. Section 8.4 will present results for the 180° group, where subjects were asked to rate semi visually-immersive scenes. Section 8.5, will introduce a comparative analysis between the 180° and 360° groups, and finally, Section 8.6 will present results from the 360° group, where subjects were asked to rate fully-immersive scenes.

To begin with our analysis of results from the 180° group, Section 8.4.1 will introduce the overall distribution of subject responses from each scene and rating scale followed by a more specific look at a selection of rating scales - pleasant, interesting, and exciting. As we were interested in creating a composite rating for visual interest, we also took the mean rating of pleasant, interesting, and exciting for each subject, hereafter referred to as 'PIE,' and considered it alongside the other uni-polar scales. Section 8.4.2 will then introduce the results from a non-parametric pair-wise comparison to present the effects of space and sky/view parameters on each individual rating scale. Section 8.4.3 will explore the relationship between subjective responses and image-based algorithms such as mSC_5 (alongside other related algorithms previously presented in Chapters 4 & 5) to evaluate their predictive capability in semi and fully visually-immersive scenes. To this end, we will use Pearson Correlation Coefficient values (r) to look for instances of high correlation between predictions from the selected metrics (SC, mSC_5, RAMMG, mean brightness and RMS contrast) and the median rating per scene (space, sky and view).

8.4.1 180° Group: Distribution and Dependence of Subject Ratings

To evaluate the relationship between subject responses for all rating scales in the 180° group, Table 8.4 shows Pearson Correlation Coefficient values between responses for each of the six rating scales. Correlation values reveal that ratings of pleasant, exciting, and interesting all show *relatively* high interdependence at ($r \geq 0.60$) with the highest correlation occurring between ratings of interesting and of exciting ($r = 0.79$). This analysis also revealed that ratings of exciting and calming are not bi-polar, as previously thought, but show a weak positive correlation ($r = 0.245$). Furthermore, the weak correlation values between ratings of diffuse and/or contrasted and those of pleasant, interesting, exciting, and calming show that these characteristics in and of themselves, do not seem to be related to more emotional indicators. This finding is contrary to the what we saw in the online experiment, where bi-polar ratings of contrast were highly correlated to ratings of excitement and stimulation. This could be explained by the design of the online experiment, where bi-polar scales were presented in sequence (Appendix 2, Figure A2.3) and may have had an influence on adjacent rating scales. Interestingly enough, we will see later that those contrast-based algorithms selected for comparison in this chapter (Section 8.4.3) show a higher dependence to factors of pleasant, interesting, and exciting than they do to contrast. This finding *is* consistent with the results from our online experiment, which found the highest correlations between the mSC_5 algorithm and subject ratings of direct, stimulating, and exciting.

Table 8.4 – Pearson Correlation Coefficient values between subject ratings for each rating scale. Bold italicized values indicate $r \geq 0.60$.

	pleasant	interesting	exciting	calming	diffuse
interesting	***0.681***				
exciting	***0.611***	***0.786***			
calming	***0.568***	0.318	0.245		
diffuse	0.124	0.085	0.042	0.231	
contrasted	0.141	0.247	0.297	-0.023	-0.373

Figure 8.10 shows the distribution of responses for each rating scale, space, and possible sky/ view parameters in the 180° group. As was mentioned in Section 8.2.3, 'clear high vd,' 'clear low vd,' 'overcast high vd,' and 'overcast low vd' refer to the four scene parameters selected per space using the mSC_5 algorithm (see Figures 8.6-8.7). Responses are marked in light grey for ratings 1-4 and dark grey for ratings 7-10. Ratings 5-6, which were considered neutral, are left blank to better visualize the distributions on either end of the scale. One of the most obvious trends we can visualize from this figure is the seemingly inverse relationship between ratings of contrast and ratings of diffuse and how this relationship varies between sky conditions. Clear sky parameters have generally lower diffuse ratings and higher contrast ratings, while the reverse can be seen under overcast sky conditions. Pleasant ratings also appear to vary between sky conditions and view directions as 'clear high' are generally rated more pleasant than 'clear low' and 'overcast high' are generally rated more pleasant than 'overcast low.' The same can be said about ratings of calming, which appear to decrease between high and low mSC_5 view directions, but not necessarily between sky conditions.

To look more specifically at some of the rating scales, Figure 8.11 shows the distribution of subject responses for 'pleasant,' 'interesting,' 'exciting,' and the composite 'PIE' for each of the 180° scenes, grouped by space, sky and view. Responses for ratings 1-5 are shown in a grey gradient while ratings 6-10 are shown in purple or pink. Ratings 1-3 and 8-10 are outlined in black to show the distribution of responses toward the extreme ends of the rating scale. When we look at the distribution of 'PIE' for Douglas, Ryerson, and Spencertown, the more asymmetrical spaces, we can see a shift in the distribution between high and low mSC_5 view directions for both sky conditions. This suggests that low view directions for both clear and overcast skies, as determined using mSC_5, were generally rated lower in one or more of the 'pleasant' 'interesting' or 'exciting' scales.

Overall trends in distribution tell us about the impacts of sky condition and view on visual impressions within each scene. As subjects were not aware which parameters we were testing, a noticeable shift in responses between sky conditions tells us that daylight does indeed have an impact on perception. Shifts in responses between view directions, from a fixed view position, also tell us that interior view direction can affect our perception and appraisal of space, a somewhat intuitive finding, but one that could have impacts on spatial planning and design.

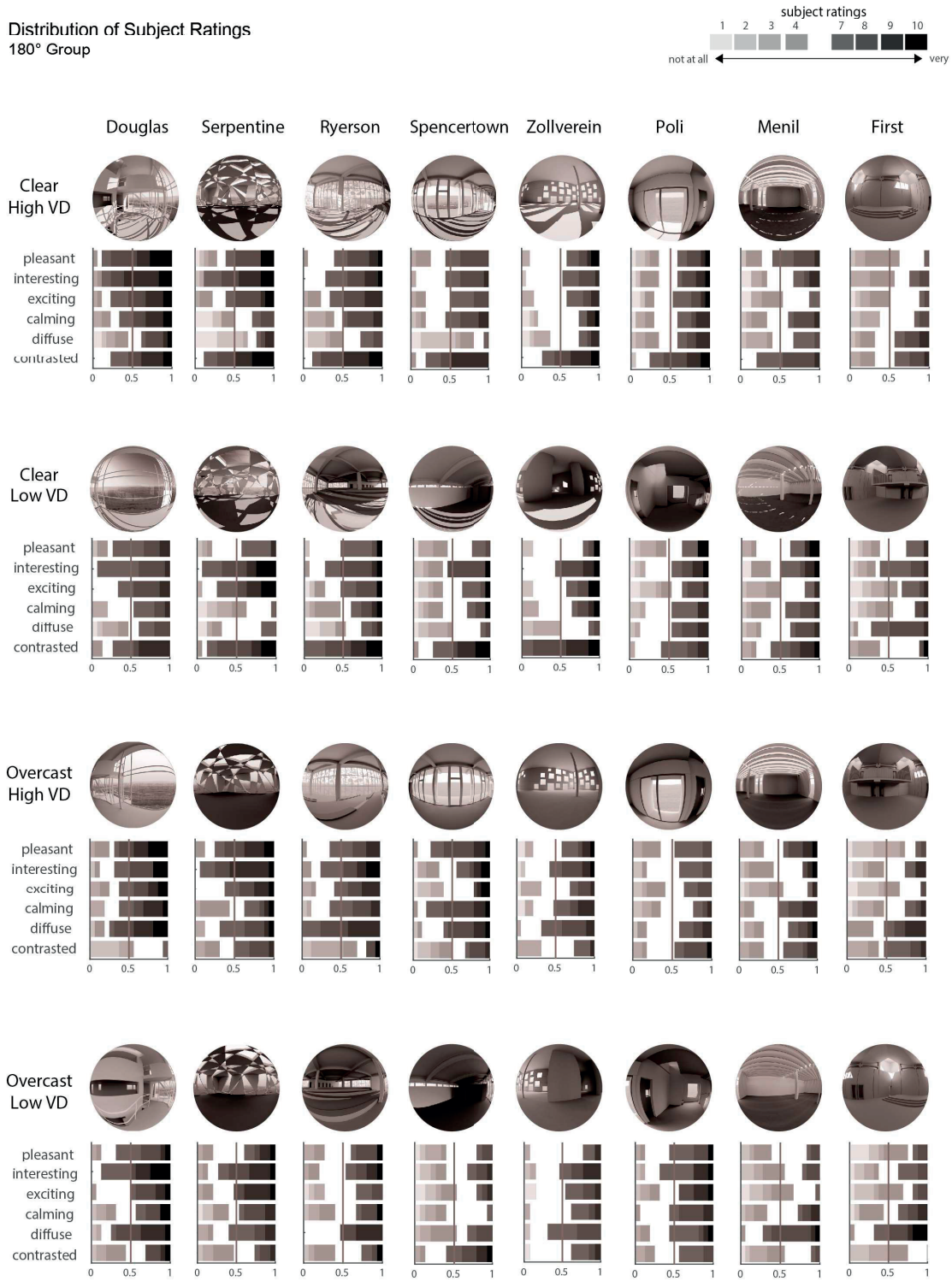


Figure 8.10 Distribution of subject reponses from each scene (clear high mSC_5, clear low mSC_5, overcast high mSC_5, overcast low mSC_5) for each of the 6 rating scales (pleasant, interesting, exciting, contrasted, calming, and diffuse). Responses are marked in light grey for ratings 1-4 and dark grey for ratings 7-10. Ratings 5-6, which were considered neutral and are left blank.

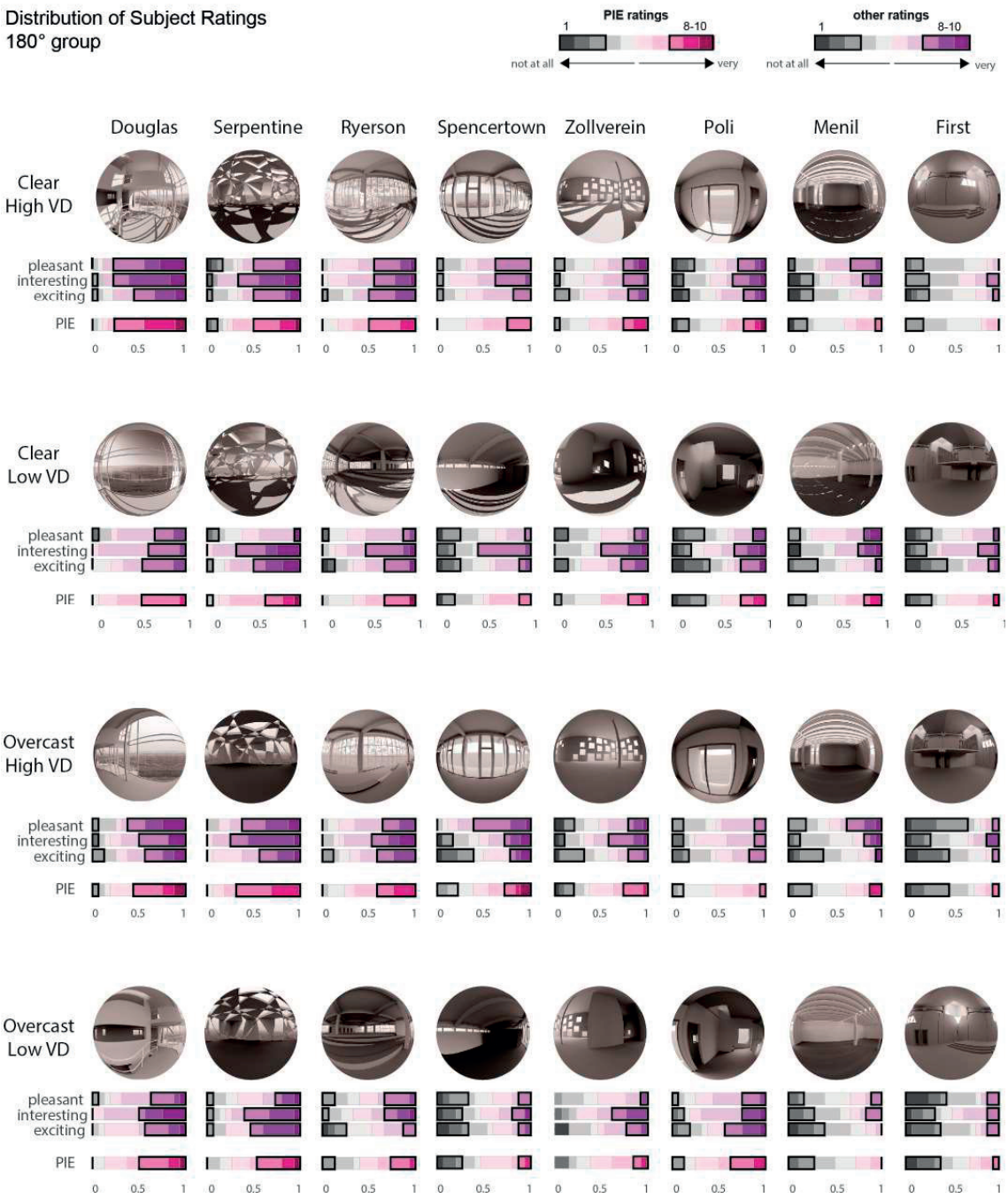


Figure 8.11 Distribution of subject reponses from each scene (clear high mSC_5, clear low mSC_5, overcast high mSC_5, overcast low mSC_5) for each of 3 selected rating scales (pleasant, interesting, exciting) plus a composite of the median responses for those 3 scales, called 'PIE.' Responses 1-3 and 8-10 are highlighted to identify extremes of each scale.

8.4.2 180° Group: Effects of Space and Sky/View Direction

To explore the effects of space and view/sky parameters on the distribution of responses for each rating scale (pleasant, interesting, exciting, calming, diffuse, contrasted, and 'PIE'), a non-parametric Kruskal-Wallis post-hoc pair-wise comparison was used. Figure 8.12 shows box-plots for each of the rating scales and spaces, merging all four sky and view parameters to isolate the impact of space). Kruskal-Wallis tests, equivalent to a one-way ANOVA for ordinal scales, are shown to the right of each rating scale, identifying space comparisons whose response distribution were significantly different at $p \leq 0.05$ with a check mark. Using this test, we can see a significant difference between responses for at least one space-to-space comparison in all ratings scales, but the most significant differences (between combinations of spaces) were observed in ratings of pleasant, interesting, exciting, and PIE.

Figure 8.13 shows box plots and Kruskal-Wallis results for each rating scale and sky/view parameter (collapsing all 8 spaces to isolate the effect of sky/view). In these KW tests, we can see the most significant differences between sky and view conditions for ratings of pleasant, diffuse, and contrasted. It is interesting to see that the significance ($p \leq 0.05$) of space as an independent effect is strongest on ratings of pleasant, exciting, interesting, and PIE, while the effect of sky and view is strongest on ratings of pleasant, diffuse and contrasted.

While we can see a slight shift in the median responses for PIE (Figure 8.13) when grouped by sky and view, the difference is not statistically significant unless we also account for space. This is expected considering the fact that ratings for clear low mSC_5 view directions were often higher than the overcast high mSC_5 view directions and that this varied depending on architectural space. As such, clear skies and high mSC_5 view directions alone did not always result in the highest ratings of PIE. The influence of space as an independent factor did, however, reveal a significant effect (<0.05) on ratings of 'PIE' for 15 of the 28 possible space-to-space comparisons.

Considering the effects of sky/view and space as independent factors is useful in understanding the relative effect of those parameters on subject ratings, but it is also interesting to see the effect of sky and view on each space *independently*. It makes intuitive sense that not all architectural scenes selected for this experiment would show the same influence of view direction and sky on subjective impressions, because the scenes vary in composition across the field-of-view. Both the degree of asymmetry and the location of direct openings create variation depending on architectural design. To test the effect of view direction and sky on each space independently, we conducted a Kolmogorov-Smirnov two-sample test (Siegel, 1956) between the 'clear high' mSC_5 and 'overcast low' mSC_5 scenes for each space and rating scale.

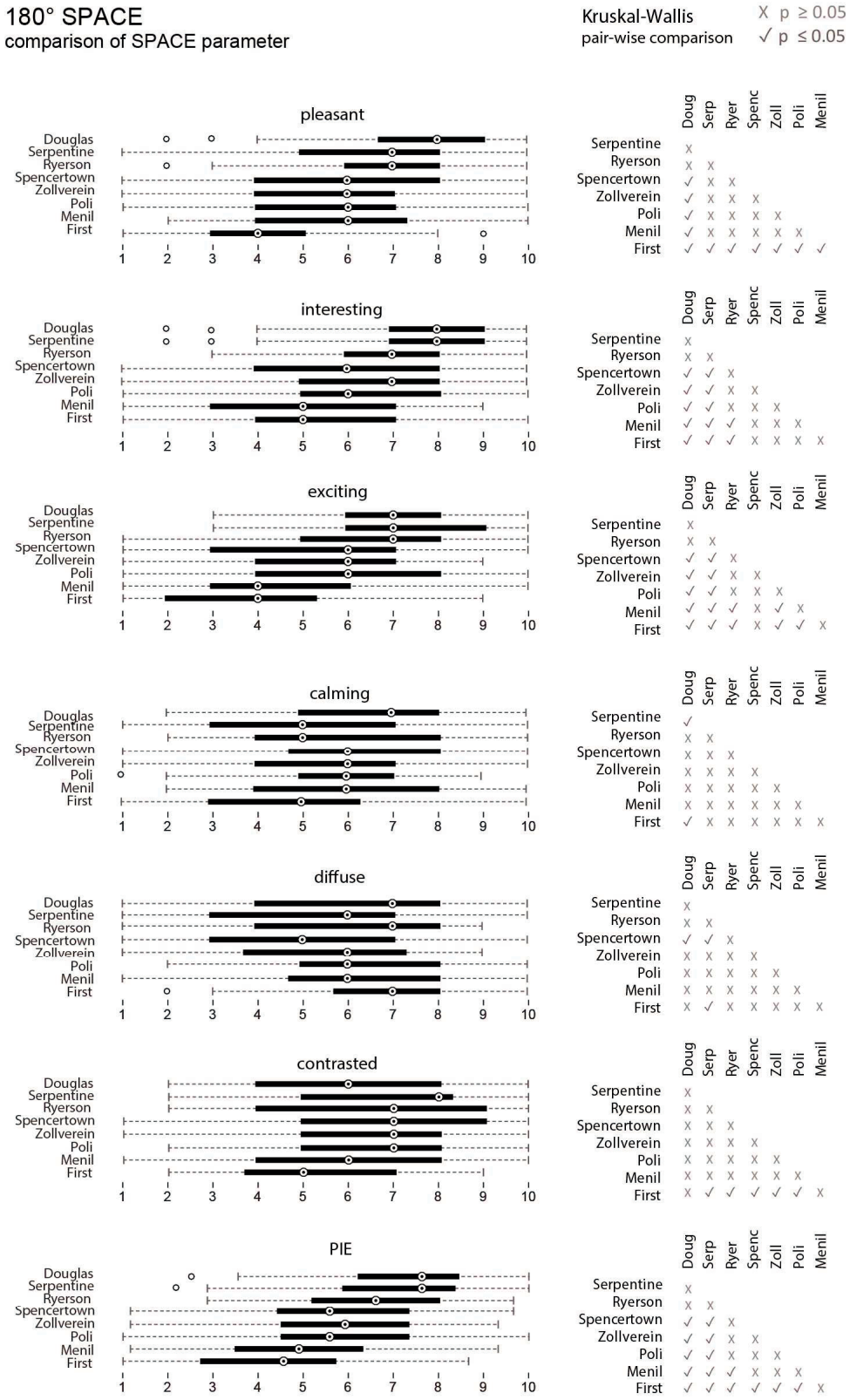


Figure 8.12 Kruskal Wallis Pair-Wise comparisons between spaces on the left (collapsing all four sky and view parameters into one) and KW tests on the right. Significant KW tests between pair-wise comparisons are shown with a check mark.

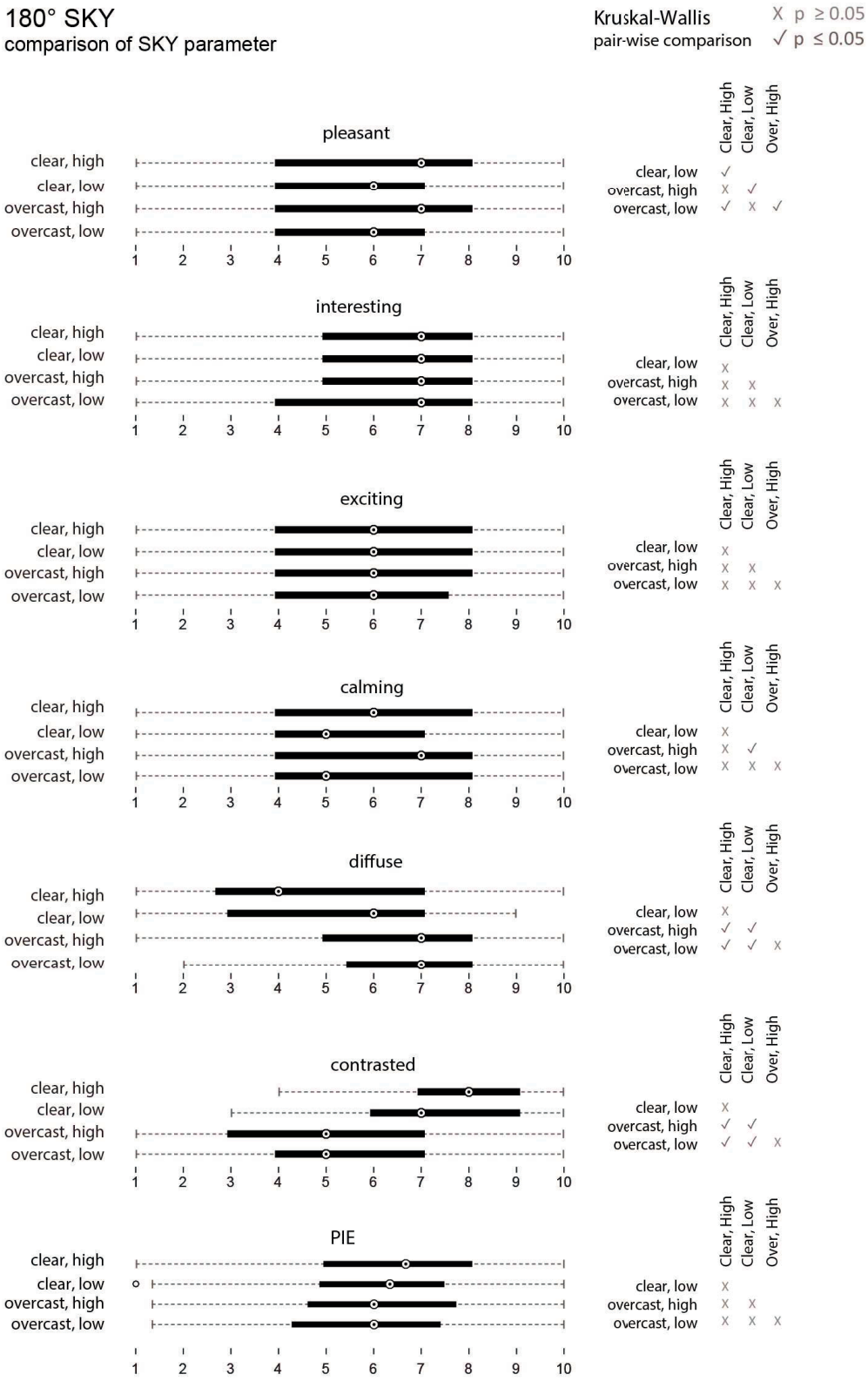


Figure 8.13 Kruskal Wallis Pair-Wise comparisons between sky and view parameters on the left (collapsing all spaces into one) and KW tests on the right. Significant KW tests between pair-wise comparisons are shown with a check mark.

Table 8.5 shows the comparison between 'clear high' mSC_5 and 'overcast low' mSC_5 scenes with significantly different distributions highlighted in bold* for (p<0.10) and bold** for (p<0.05). Those spaces with asymmetrical architecture (like Ryerson and Spencertown) show significantly different rating distributions between clear high and overcast low scenes in 4 out of the 7 rating scales. While all spaces showed at least one rating scale with a significantly different response distribution (p<0.10) between scenes, ratings of pleasant, exciting, diffuse, contrasted, and PIE show significant distributions in two or more spaces each.

Table 8.5 – Kolmogorov-Smirnov two-sample test between clear high and overcast low view directions for each space and rating scale in the 180°group.

	Douglas	First	Menil	Poli	Ryerson	Serp	Spenc	Zoll
pleasant	**0.02****	0.21	**0.06***	0.38	0.12	0.65	**0.05****	**0.08***
interesting	0.55	0.91	0.41	0.35	0.17	0.45	**0.08***	0.8
exciting	0.41	0.46	0.33	0.54	**0.04***	0.6	**0.05****	0.39
calming	0.22	0.54	0.74	0.27	0.67	0.14	**0.01****	0.98
diffuse	**0.03****	**0.07***	**0.00****	0.47	**0.00****	**0.00****	**0.01****	0.37
contrasted	**0.00****	**0.00****	**0.00****	**0.01****	**0.00****	**0.00****	0.12	**0.06***
PIE	**0.10***	0.62	0.12	0.3	**0.05****	0.8	**0.06***	0.29

This is an important finding in that both daylight distribution and view direction appear to have a significant effect on our perceptual evaluation of architectural space, especially when that space is compositionally varied across our field-of-view.

8.4.3 180° Group: Regression Analysis

To compare subject responses for each rating scale to quantitative algorithms that can be used to predict elements of contrast, visual interest, and brightness, a Pearson Correlation Coefficient analysis was done between the median response for each scene and a selection of algorithms. Figure 8.14 shows the Pearson Correlation Coefficient values (r) between each rating scale and contrast algorithm (SC, mSC_5, RAMMG, Brightness, and RMS).

The RAMMG predictor (with a seven level average N=7) was the most highly correlated to median ratings of 'pleasant' (r = 0.65) and the composite 'PIE' rating (r = 0.66). The mSC_5 predictor was also highly correlated to ratings of 'pleasant' (r = 0.64) and the composite 'PIE' rating (r = 0.63). Positive correlations were also relatively strong between ratings of excitement and interest, while negative correlations were relatively strong between diffuse responses for both mSC_5 and RAMMG. It is interesting to note that while mean brightness showed an inferior fit to most rating scales (relative to the other predictors), it showed the highest fit to ratings of 'calm' (r = 0.38). Logistic regression results are shown in Appendix 4, Figure A4.21.

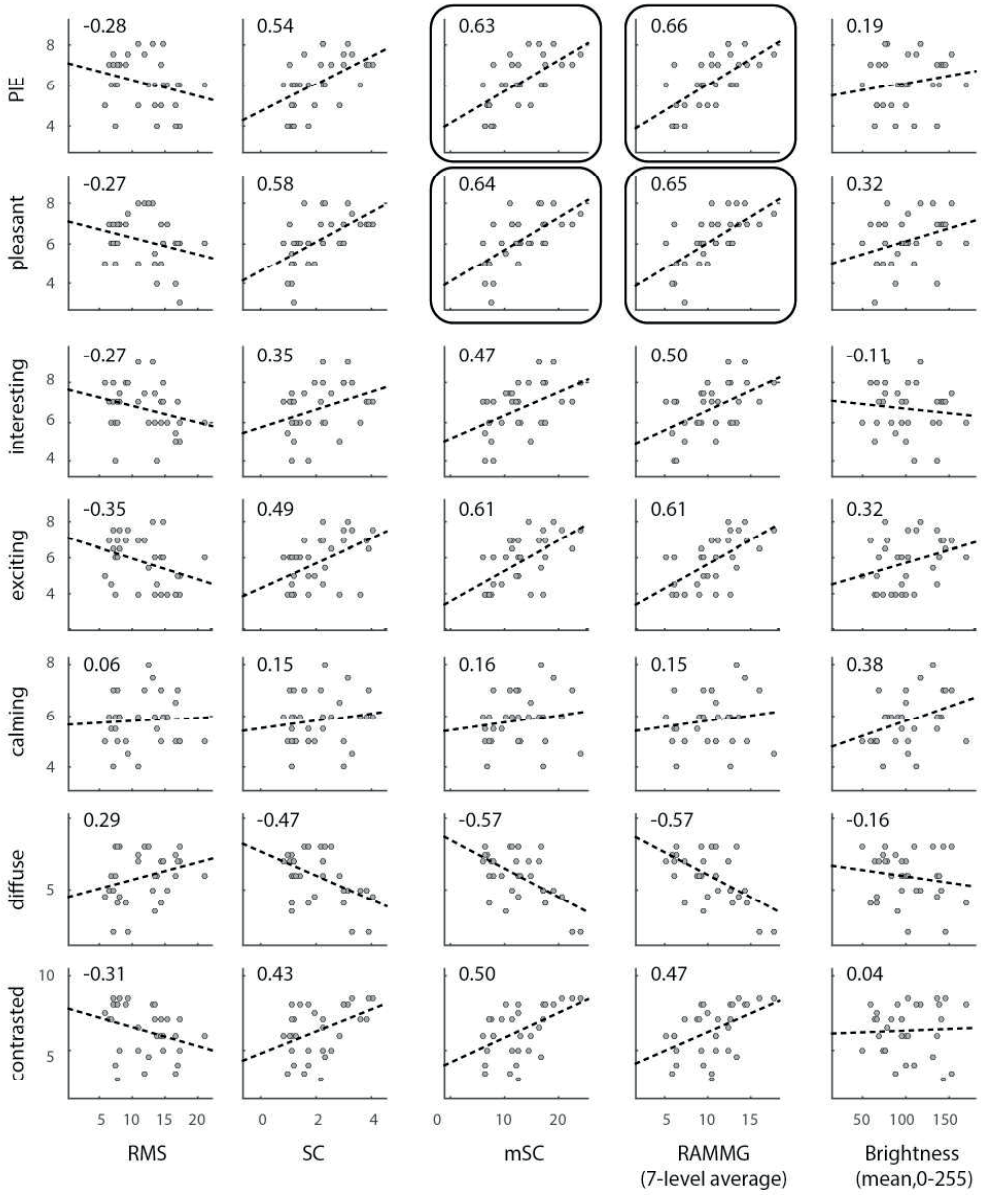


Figure 8.14 Showing Pearson Correlation Coefficient values (*r*) between median subject ratings and quantitative predictors RMS, SC, mSC_5, RAMMG, mean brightness.

8.5 *Comparative Analysis of Results: 180° Group vs. 360° Group*

As described in Section 8.3.2, subjects in this experiment were placed in one of two groups depending on a randomization algorithm which assigned 2/3 of subjects into the 180° and 1/3 of subjects into the 360° group. Section 8.4 introduced data analysis from the 180° group, describing the effects of both view direction and sky conditions on subjective ratings. As previously stated, subjects in the 180° group saw one of four possible sky and view combinations per space. Unlike the 180° group, subjects in the 360° group saw one of only two possible conditions per space, (either clear or overcast) entering the scene in the *high* mSC_5 view direction and then freely

exploring the remainder of the scene. This section will compare the subjective data from *between* the two groups. In order to compare the ratings from the 180° to the 360° group, we will only assess the high mSC_5 view direction for each of the 180° scenes as these correspond to the same starting view direction in the 360° scenes. To understand the difference in responses between groups that saw the180° scenes and those that saw the 360°, we conducted a Kolmogorov-Smirnov two-sample statistical test (Siegel, 1956). The KS two-sample test can be used to determine whether two probability distributions, with one-dimensional independent samples vary, in this case – whether subjective responses from the 180° scene are significantly different than those reported from the 360° scene.

Table 8.6 shows the p values for this test, applied to subjective responses for each rating scale (pleasant, interesting, exciting, calming, diffuse, contrasted, and PIE) between each space and sky (16 scenes in either 180° or 360°). The null-hypothesis (that the data from one sample is draw from the same distribution as the other) can be rejected for $p \leq 0.05$, however we have also highlighted p values that are ≤ 0.10 as these may also show considerable difference in the response distributions between scene conditions. From the highlighted values in Table 8.6, we can see that subjective responses between the 180° and 360° group are significantly different ($p \leq 0.05$) for one or more scales in 6 of the 18 scenes and moderately different ($p \leq 0.10$) in an additional 4 scenes. This test suggests that subjective responses to daylight conditions within a given scene are impacted by the degree of visual immersion experienced in that space.

Appendix 4, Table A4.20 shows the median subject responses for each space, sky, and high view direction in both the 180° and the 360° groups. Significant differences in median response due to scene immersion, as determined by the KS test in Table A4.20 are highlighted in bold* for ($p<0.10$) and bold** for ($p<0.05$). It does appear, from both Table 8.6 and Table A4.20 that overcast skies had a significant effect on more ratings than clear skies between the 180° and 360° groups.

To compare ratings for each scenes and rating scale between the 180° and 360° groups, Figure 8.15 shows Pearson Correlation Coefficient values (*r)* between median subject responses for each group. From this correlation analysis, we can see that ratings for interesting, exciting, diffuse, and PIE are most highly correlated ($r \geq 0.80$), while ratings for pleasant, calming, and contrasted have Pearson values $r \geq 0.75$. Given the added range of visual information and composition in the 360° scenes, it is somewhat surprising to see such a high correlation between these two groups. To speculate on this, it is useful to consider a limitation in our experimental design. As subjects in the 360° group were always introduced into the scene facing in the high view mSC_5 direction, perhaps it is not surprising that their responses were so closely correlated to the high view direction mSC_5 in the 180° group. We did observe, during the experiment, that after a brief period of exploration, subjects in the 360° group remained facing forward during the verbal delivery of subjective ratings, often directing their bodies and heads toward the location of the researcher's voice. This rather harmless human tendency to face the direction of the person you

Table 8.6 – Kolmogorov-Smirnov two-sample test

	Douglas 180° vs 360°		**First** 180° vs 360°		**Menil** 180° vs 360°		**Poli** 180° vs 360°	
	clear	overcast	clear	overcast	clear	overcast	clear	overcast
pleasant	0.15	0.86	0.74	0.18	0.43	1.00	0.36	0.29
interesting	0.15	0.34	0.92	0.12	0.85	0.61	0.54	0.95
exciting	0.83	0.43	0.86	0.15	0.31	0.60	0.56	0.69
calming	0.62	0.96	0.74	0.97	0.61	0.74	0.80	***0.06****
diffuse	***0.09****	0.29	0.93	0.37	0.91	***0.09****	0.77	0.24
contrasted	0.90	***0.02*****	0.58	0.30	0.37	0.18	0.23	0.32
PIE	0.17	0.17	0.84	0.53	0.90	0.53	0.39	0.61

	Ryerson 180° vs 360°		**Serpentine** 180° vs 360°		**Spencertown** 180° vs 360°		**Zollverein** 180° vs 360°	
	clear	overcast	clear	overcast	clear	overcast	clear	overcast
pleasant	0.37	0.70	0.49	***0.07****	0.88	***0.001*****	***0.04*****	0.52
interesting	0.54	0.68	0.68	0.49	0.74	0.68	0.34	0.64
exciting	0.63	0.43	0.82	0.90	0.41	0.66	0.58	***0.03*****
calming	0.11	0.17	0.21	0.74	0.49	***0.04*****	0.50	0.34
diffuse	0.83	0.17	0.32	***0.07****	***0.04*****	0.66	0.18	0.74
contrasted	0.34	***0.06****	0.93	***0.02*****	0.94	0.45	0.66	***0.10****
PIE	0.38	0.68	0.62	0.32	0.65	0.38	0.18	0.19

*$p \leq 0.10$, **$p \leq 0.05$

are speaking to may pose a significant experimental bias when testing the effect of view immersion as an independent factor. Subjects inadvertently spent more time looking in the high direction (as it was the direction they entered and the direction of the researcher) even if they could freely explore the entire scene. To address this further, we would have to conduct a new study using randomized entry view directions for both groups and find a way to deliver the questions through a speaker in the ear phones, to ensure that people are evaluating the entire scene, and not just the view direction facing the researcher's voice.

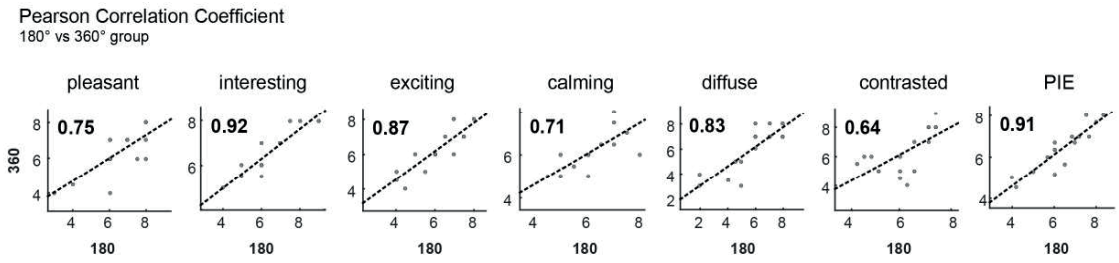


Figure 8.15 Pearson Correlation Coefficient (*r*) values between 180° and 360° groups for ratings of pleasant, interesting, exciting, calming, diffuse, contrasted, and PIE.

8.6 Results for the 360° Group

In Section 8.5, a comparative analysis of subjective data was presented from the two subject groups in this experiment to find whether responses varied depending on scene immersion. In this section, we will assess the 360° group as an independent data set, as was done for the 180° group in Section 8.4. Through this analysis, we will also compare the 'cube face' mSC_5 algorithm prediction - used to select instances of time in this experiment - to subjective assessments from the fully-immersive scenes. As was described in Section 8.2.2, the cube face mSC_5 algorithm takes an average of mSC_5 from across the 6 faces of the scene. While we believe it's valuable to understand how visual interest varies across view directions (as can be computed using the hemispherical adaptation of mSC_5 across a range of view directions), it is also useful to have a single value across the entire scene, for use is temporal evaluations where single view directions provide many data points and can be overwhelming to comprehend.

8.6.1 360° Group: Distribution of Subject Ratings

To test the relationship between subject ratings for each rating scale in the 360° group, Table 8.7 shows Pearson Correlation Coefficient values (r). Similar to the results from the 180° group shown in Table 8.4, we can see a relatively high correlation between ratings of pleasant, interesting, and exciting ($r \geq 0.65$). In this case, ratings of pleasant and calming also have a relatively high correlated $r = 0.686$. As was found in the 180° group, ratings of diffuse and contrasted have a low correlation to ratings of pleasant, interesting, exciting, and calming, but show some low negative correlation to one another.

Table 8.7 – Pearson Correlation Coefficient values (r) between subject ratings for each rating scale in the 360° group. Bold values indicate $r > 0.65$.

	pleasant	interesting	exciting	calming	diffuse
interesting	***0.676***				
exciting	***0.668***	***0.775***			
calming	***0.686***	0.402	0.40		
diffuse	0.2364	0.157	0.08	0.31	
contrasted	0.122	0.239	0.274	0.012	-0.34

Figure 8.16 shows the distribution of all subject responses, grouped by space and then sky condition. Ratings 1-3 and 8-10 are highlighted in black to emphasize the distribution of responses on either extreme of the scale. While subjects in the 180° group saw one of four possible scenes, subjects in the 360° group saw one of only two possible scenes per space; clear or overcast. An initial observation of this data shows more spread than when compared to the 180° group for most rating scales. While the distribution of responses for ratings of pleasant, interesting, exciting, composite PIE, and contrast generally diminish between spaces like Douglas, Serpentine, Ryerson, and Spencertown and spaces like Menil and First (as expected based on the mSC_5 predictions), there is a less obvious and sometimes counter-intuitive influence of sky condition. The most obvious and

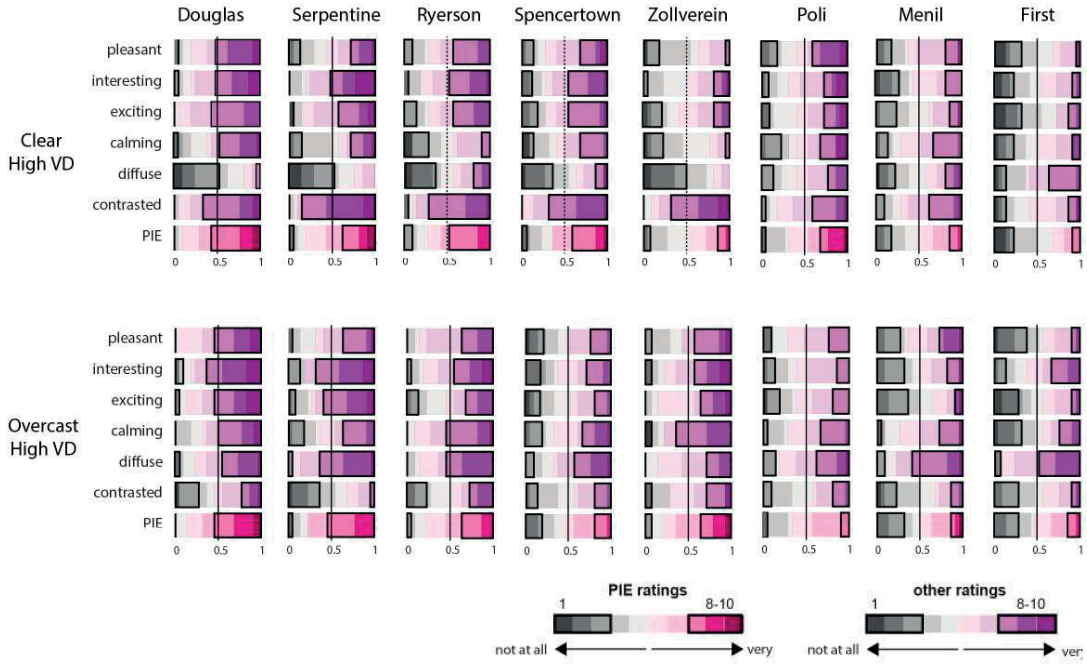


Figure 8.16 Shows the distribution of responses for each space, sky, and rating scale in the 360° subject group. Ratings 1-3 and 8-10 are highlighted with a black border, while composite PIE ratings are shown in magenta.

consistent impact of sky can be observed in ratings of contrast and diffuse. While high contrast ratings under clear skies generally relate to lower contrast ratings under overcast skies for the same space, high contrast ratings also generally relate to lower diffuse ratings. In other words, a low median response on the diffuse rating scale under clear sky parameters is always met with a higher median response under overcast skies. This observation may be intuitive, based on the assumption that the indirect skylight experienced under overcast sky conditions is usually more diffuse than the direct sunlight under clear sky conditions.

When we observe PIE ratings between spaces, we see a clear shift in the distribution of ratings between high mSC_5 spaces like Douglas, Serpentine, and Ryerson and low mSC_5 spaces like Menil or First, supporting the idea that architectural composition has an influence on subjective ratings of pleasantness, interest, and/or excitement even if sky conditions are less impactful. The distribution of high PIE ratings (8-10) is equal or higher for clear sky conditions than overcast sky conditions in Douglas, Ryerson, Spencertown, Poli, and Menil, but lower for Serpentine, Zollverein and First. A more precise look at the isolated impacts of space and sky on each rating scale in the following section will provide a more detailed assessment of this observation.

8.6.2 360° Group: Effects of Space and Sky

To measure the effect of space and sky as independent factors, we will first present the results of a Kruskal-Wallis test using a post-hoc pair-wise comparison. This test will show us which

combination of independent samples are likely to originate from the same distribution. In this case, a significant KW test determines when an individual sample, such as a space or sky condition, shows a stochastic dominance over another space or sky condition. To compare the effect of sky on each space independently, we will then conduct a Kolmogorov-Smirnov two-sample test to determine whether subject responses to each sky condition come from the same distribution.

Figure 8.17 shows box plots for each space in the 360° group - merging the two sky directions to isolate the effect of space - and Kruskal-Wallis tests using a post-hoc comparison of responses for each rating scale. Significant KW tests ($p \leq 0.05$) are shown with a check mark. In this analysis, the highest number of significant differences in KW results in space-to-space pair-wise comparisons were found for the following rating scales: pleasant, interesting, exciting and PIE. Responses for calming and diffuse had the fewest significant differences in the KW results in space-to-space comparisons, similar to the KW tests run for the 180° group. That being said, there were fewer significant comparisons found in the 360° group than the 180° group, perhaps due to the broader spread in responses for the 360° group.

Figure 8.18 shows box plots for clear and overcast sky conditions - merging all spaces to isolate the effect of sky - for each ratings scale. Using a Kruskal-Wallis test, significant differences between the distribution of responses for each sky condition are shown to the right, with $p \leq 0.05$ highlighted using a check mark. It is interesting to note that those rating scales with a significant KW result in sky-to-sky comparisons (calming, diffuse, and contrasted in Figure 8.18), had the fewest significant KW tests in space-to-space comparisons (Figure 8.17). Given this finding, it appears that the impact of architectural composition may have a stronger effect on ratings of pleasant, interest, and excitement while sky has stronger effect on ratings of calm, diffuse, and contrast. This is consistent with the findings from the 180° group.

When we look at the effect of sky condition on subject ratings in each of the architectural spaces separately using a Kolmogorov-Smirnov two-sample test (Table 8.8), we can see that the distribution of ratings for diffuse and/or contrasted scales show a moderately significant different in most spaces ($p \leq 0.10$). We can also see that some spaces have more significant differences in subject ratings than others. Ryerson, Spencertown, and Zollverein all show significant differences

Table 8.8 – Kolmogorov-Smirnov two-sample results between Clear and Overcast sky conditions for each space and rating scale in the 360° group.

	Douglas	First	Menil	Poli	Ryerson	Serp	Spenc	Zoll
pleasant	0.64	0.86	1	0.74	0.46	0.51	0.24	**0.06***
interesting	0.54	**0.10***	0.85	0.33	0.86	1	0.16	**0.09***
exciting	0.9	0.69	0.37	0.48	0.29	0.52	**0.08***	0.13
calming	0.75	0.8	0.96	0.67	**0.00****	0.62	0.65	**0.00****
diffuse	**0.00****	**0.10***	**0.00****	0.24	**0.00****	**0.00****	**0.02****	**0.00****
contrasted	**0.00****	0.29	**0.01****	**0.03****	**0.00****	**0.00****	**0.00****	**0.01****
PIE	0.63	0.47	0.85	0.53	0.71	0.61	**0.10***	**0.05****

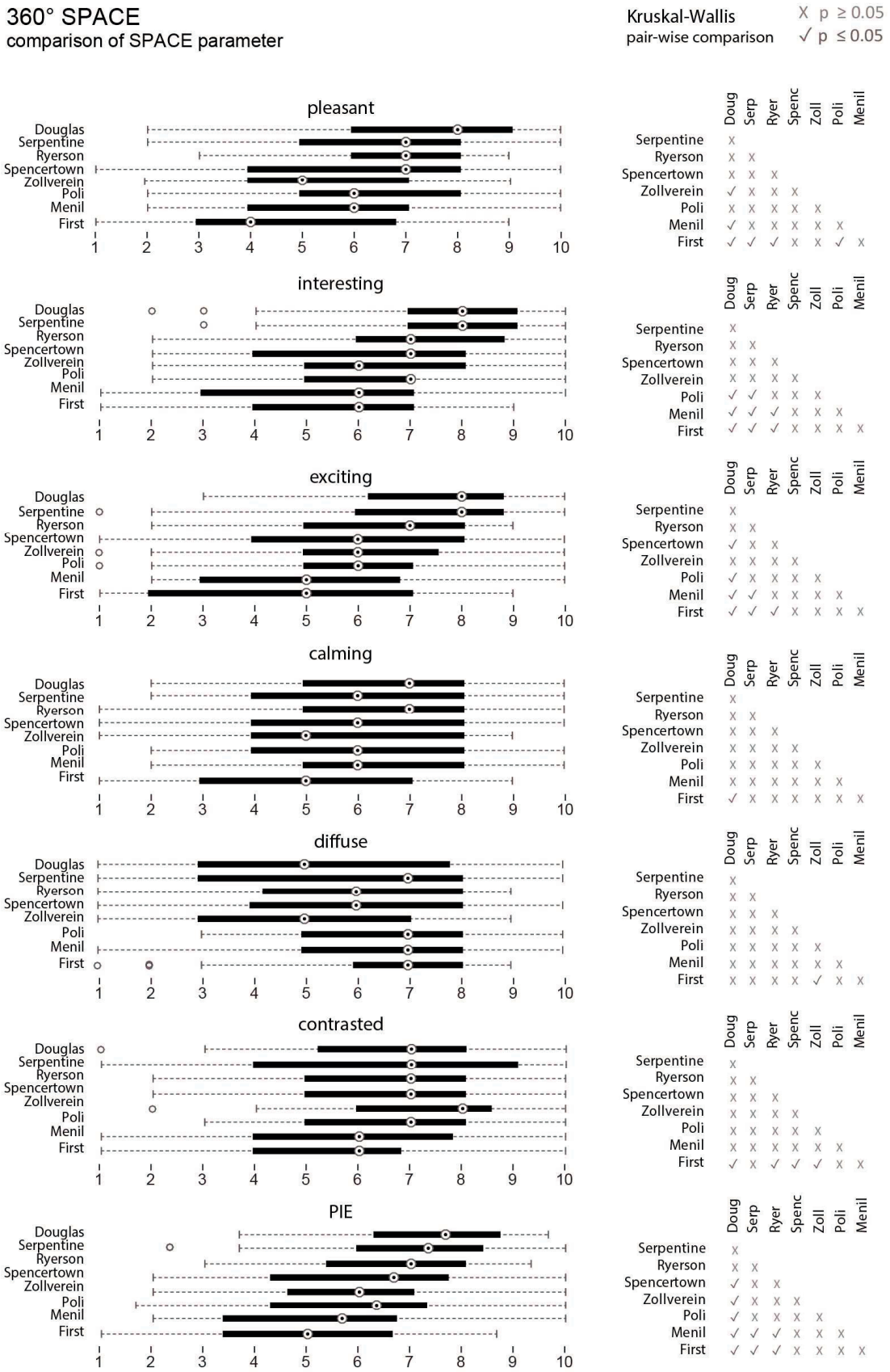


Figure 8.17 Kruskal-Wallis pair-wise comparisons for space and sky parameters

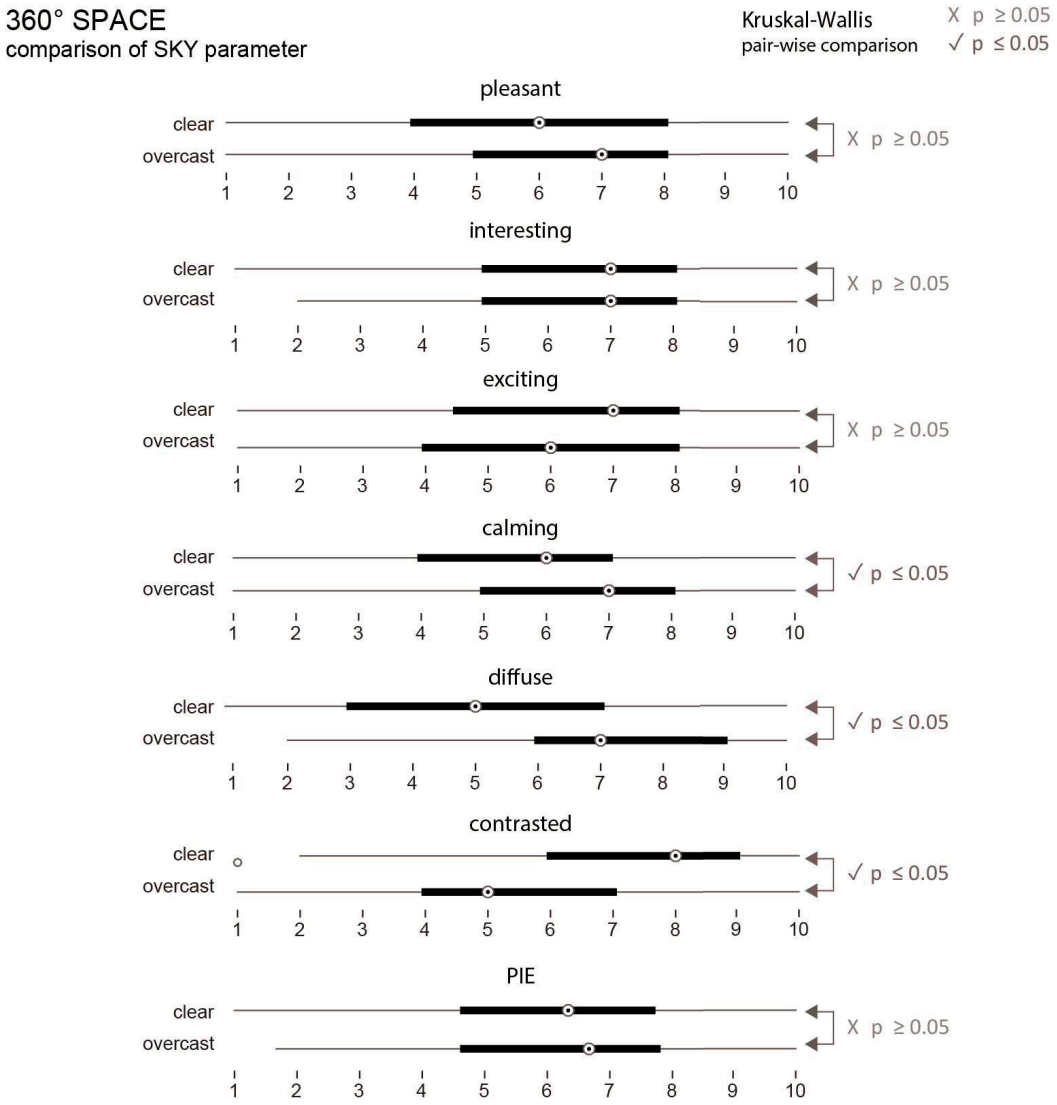


Figure 8.18 Kruskal-Wallis pair-wise comparisons for space and sky parameters

in 3 or more rating scales, including calming, exciting, pleasant, and interesting. Given the asymmetrical distribution of architectural elements that emit daylight in these spaces, it makes sense that sky condition would have a greater impact on subject impressions in these spaces. First Unitarian Church, with its diffusing roof monitors, shows the fewest number of significantly different rating scales between the two sky conditions.

8.6.3 360° Group: Regression Analysis

To determine whether subject responses for any of the rating scales in the 360° group show dependence to the 'cube face' mSC_5 algorithm used to predict instances of visual interest in the 360° scenes (as described in Section 8.2.2), a regression analysis was conducted. Figure 8.19 shows the Pearson Correlation Coefficient values between 'cube face' mSC_5 and RAMMG algorithms and median responses for each rating scale. While Pearson Correlation values for

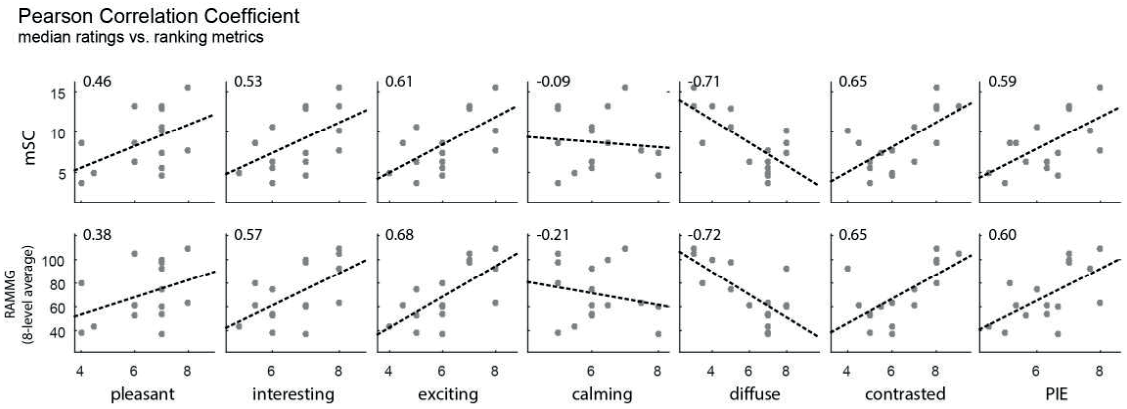


Figure 8.19 Shows the Pearson Correlation Coefficient values (*r*) and linear regression fits between mSC_5 and RAMMG box algorithms and subject ratings in the 360° group.

RAMMG are greater than or equal to 0.60 for ratings of exciting, contrasted, and PIE, there is also a negative correlation between RAMMG and ratings of diffuse (r = -0.72). Correlation values between mSC_5 and exciting, contrasted, and PIE are also relatively strong, with the biggest difference between the two algorithms appearing for the exciting rating scale (where r for mSC_5 equals r = 0.61 and r for RAMMG, r = 0.68) and the pleasant scale (where r for mSC_5 equals r = 0.46 and for RAMMG, r = 0.38). The correlation between cube face mSC_5 and RAMMG predictions and ratings of calming is low, as was also seen in the 180° group. A logistic regression analysis through subject ratings and mSC_5/RAMMG was also conducted for each rating scale and is presented in Appendix A4.22. The lowest deviance in distribution and fit was found between ratings of interesting and mSC_5 (7.09) and composite ratings of PIE and mSC_5 (7.22).

Table 8.9 shows the side-by-side comparison of correlation values for both mSC_5 and RAMMG algorithms (180° and 360° implementations) for median ratings of pleasantness, interest, excitement, and PIE. Pearson Correlation Coefficient values between ratings in the 180° group and the hemispherical mSC_5 prediction in Table 8.9 are higher for ratings of pleasant and PIE than those between ratings in the 360° group and the cube face mSC_5 prediction. We can also see, however that ratings of interesting and exciting are more highly correlated to predictions made using the cube face mSC_5 prediction in the 360° group than in the 180° group using the hemispherical mSC_5 algorithm.

Table 8.9 - Pearson Correlation Coefficient (*r*)values between median subject ratings and mSC_5/ RAMMG predictions for both 180° and 360° groups.

	mSC			RAMMG		
	180°		360°	180°		360°
pleasant	***0.64***	>	0.46	***0.65***	>	0.38
interesting	0.47	<	***0.53***	0.5	<	***0.57***
exciting	0.61	=	0.61	0.61	<	***0.68***
PIE	***0.63***	>	0.6	***0.66***	>	0.6

8.7 Head Tracking

One of the main motivations for using VR in this experiment was the ability to extract head tracking information from the data log in Unity, allowing the research team to see where subjects looked within each scene. While a more detailed analysis of this data is ongoing and will not be presented in this thesis, this section will present a first overview of results for the 180° group which offer a possible explanation for variations, where applicable, between quantitative image algorithms and subject ratings.

The collected head tracking data consists of a series of normalized vectors, generated every 11 milliseconds, from the center of the headset in that instance. By using a custom script in Unity, a series of head tracking view directions were extracted to corresponded to each space and sky/view combination for every experimental session, as shown in Figure 8.20a. These vectors were then separated into three groups based on their absolute vertical distance from the horizontal, expressed as a fraction of the vertical field of view: 0-25%, 25-50% and 50-100%, as illustrated in Figure 8.20b. By merging all the experimental sessions and conditions for each space, we calculated the normalized frequency distribution of the participants' vertical head movement in Figure 8.20c. This analysis confirms an observation that was made by the research team during the experimental sessions: most of the time, the participants' vertical head movement was within the 0-25% band of absolute distance from the horizontal, relative to the subjects' view direction. For all the spaces on average, the head tracking vectors stay within the 0-25% band for 74.77% of the time, within the 26-50% band for 19.48% of the time and within the outer band for 5.74% of the time. The region between the horizontal and ±25% of the vertical field of view corresponds to 45°, which is in line with the suggestion of a 40° horizontal band as the main region of influence on perceptual impressions of space by Loe et al. (1994). This behavior could explain some discrepancies between the mSC_5 predicted excitement and subjective evaluations for a given space, if the main interest-inducing source is outside of the focus of the users, as is the case with the roof of Menil gallery, as seen in Appendix 4, Figure A4.15.

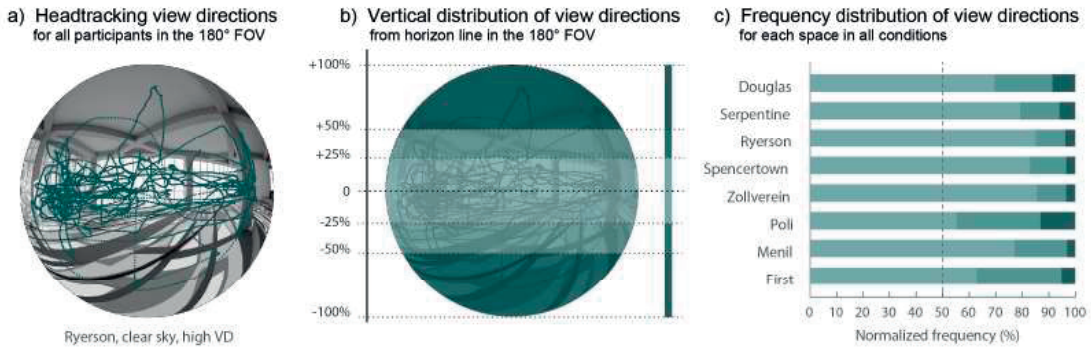


Figure 8.20 a) Headtracking view directions for Ryerson (clear sky, high VD) for all participants, b) grouping of vertical distribution of view directions and c) frequency distribution of view directions in the vertical axis for each space.

This finding indicates that prediction algorithms could potentially be improved with the integration of a view-dependent weight. Further analysis of the head-tracking data set collected in this experiment may provide future insights into a more robust algorithms and will be discussed in the Conclusion chapter.

8.8 Chapter Summary

This chapter introduced an experimental study using visually-immersive rendered scenes from a selection of 8 daylit architectural spaces presented to subjects in the Oculus Rift CV1 virtual reality headset. The authors collected subjective and objective data, through verbal questionnaires and head tracking respectively, introducing a novel experimental approach for use in qualitative lighting research. By varying the sky conditions and view direction of rendered scenes, the authors were able to compare subjective ratings of those scenes from a population of subjects to quantitative algorithms designed to predict perceptual responses across the subject's field-of-view. To test the influence of view immersion, two subject groups were created; one that saw semi-visually immersive 180° scenes and one that saw fully-immersive 360° scenes.

While the first experimental study presented in Chapter 5 of this thesis used 2D rectangular images (from a single view direction) to collect subjective impressions of calming - exciting, this chapter introduces the first study of its kind to use a visually-immersive virtual approach, allowing for the collection of data from a fixed position in space across a range of view directions and sky conditions. The relationship between subjective ratings of pleasant, interesting, exciting, and contrasted and image-based algorithms mSC_5 and RAMMG that have been adapted to immersive image formats, provides convincing evidence that impressions of visual interest can be modelled and predicted in immersive rendered scenes.

From the 180° subject group in this experimental study, we found significant effects of architectural space on ratings of pleasant, interesting, exciting, and contrasted. When we isolated the effect of sky and view direction across all 8 spaces, it was significant on ratings of pleasant, diffuse, and contrasted. When we considered each space independently, the effect of sky and view direction was significant on at least one (but often 3 or more) rating scales for all spaces included in this study. Pearson Correlation Coefficient values were highest between the hemispherical mSC_5 and RAMMG algorithms and median ratings of pleasant, interesting, exciting, and contrast for any of the image-based algorithms we considered. This result is consistent with the findings from our online experiment in Chapter 5.

From the 360° group, our findings show a similar trend, with significant effects of architectural space on subject impressions of pleasant, interesting, exciting, and contrast. Effects of sky varied slightly, showing significant effects on ratings of calm, diffuse, and contrast. When we compared ratings between clear and overcast sky conditions in each space individually, we saw significant effects on ratings of contrast and/or diffuse in all spaces, with significant effects on pleasant,

interesting, exciting or calming in those architectural spaces with varied and asymmetrical facade composition (like Ryerson, Spencertown and Zollverein).

Linear regression analyses between both the mSC_5 and RAMMG algorithms and median ratings of pleasant and PIE revealed higher Pearson Correlation Coefficient for the 180° scenes than those in the 360° scenes (see Table 8.9). As was discussed in Chapter 5 with regards to data analysis methods in the online survey, a linear fit through *median* ratings does not always represent a robust fit with ordinal data as the distribution of subjective ratings is important. A linear fit through *all* ratings would be deceiving as well, because we expect a spread in subjective responses, especially when using a 10-point scale. For these reasons, a logistic regression analysis was considered most appropriate for the ordinal data collected in this experiment. Similarly to the approach used in Chapter 5, we ran a logistic regression analysis (proportional odds model) between subject ratings for each rating scale and results from the mSC_5 and RAMMG algorithms. The full set of results for these analyses can be found in Figures A4.21 and A4.22 (for the 180° and 360° groups, respectively).

As the focus of this thesis through Chapters 6 & 7 was to develop a prediction model that could analyze a range of specific view-directions from a fixed view position or an array of view positions, we decided to delve further into the results from the 180° scenes. The lowest deviance between ratings and hemispherical mSC_5 and RAMMG algorithms for any of the rating scale was found between ratings of PIE and RAMMG (19.13) and between ratings of PIE and mSC_5 (19.29). The high correlation between mSC_5 and RAMMG predictions ($r = 0.96$) leads the author to conclude that either of these algorithms could be used as an indicator of visual interest. Further work with an expanded set of spaces may be needed to determine which of these algorithms is more appropriate for immersive scenes.

To illustrate how uni-polar PIE responses could be predicted using one of the algorithms mentioned above, the Figure 8.21 shows the logistic regression between PIE responses for each rendered scene and the RAMMG algorithm (which revealed a *sightly* lower deviance than mSC_5). As can be seen in this analysis, an increase RAMMG results in a higher percentage of subjects who would likely rate the scenes as higher in 'PIE.' In other words, as RAMMG increases, so too does the percentage of subjects who rated those images as more pleasant, interesting, and/or exciting.

Figure 8.22 shows RAMMG values for each view direction plotted across a series of eighteen view directions from a central view position in the Ryerson Student Learning Center on January 13, 9:47am. Using the 50% population lines, shown in the graph on the right, three thresholds are determined: one over which 50% of the population rates the scenes as 5 or higher (RAMMG=7.5), one over which 50% of the population rates the scene as 6 or higher (RAMMG=11.75), or one over which 50% rates the scenes as 7 or higher (RAMMG=16.10). While this chapter does not intend to propose a new visually-immersive model for predicting subject responses, in part due to

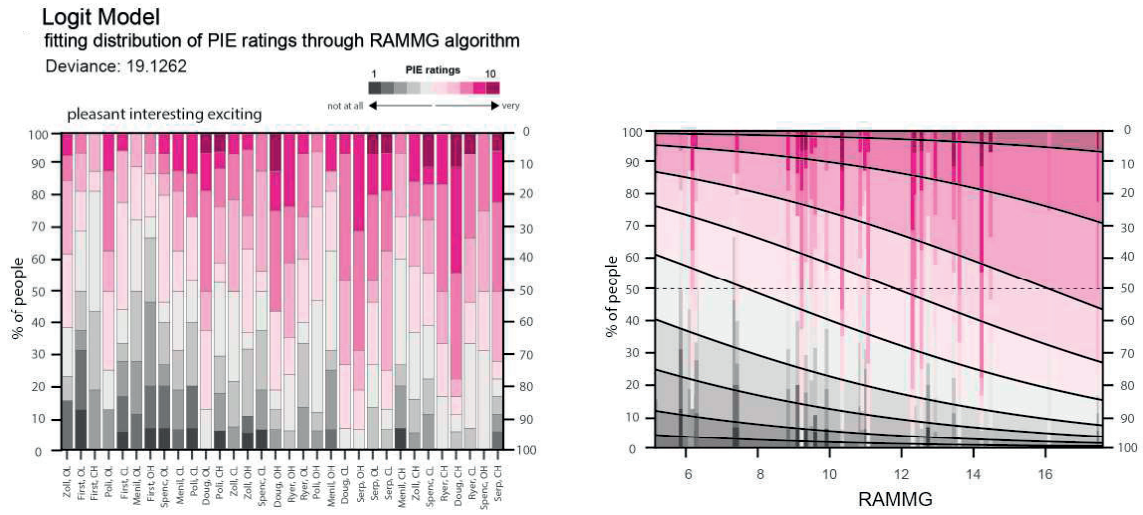


Figure 8.21 Shows the logistic regression and resulting logit model between RAMMG and subject ratings of PIE in the 180° group.

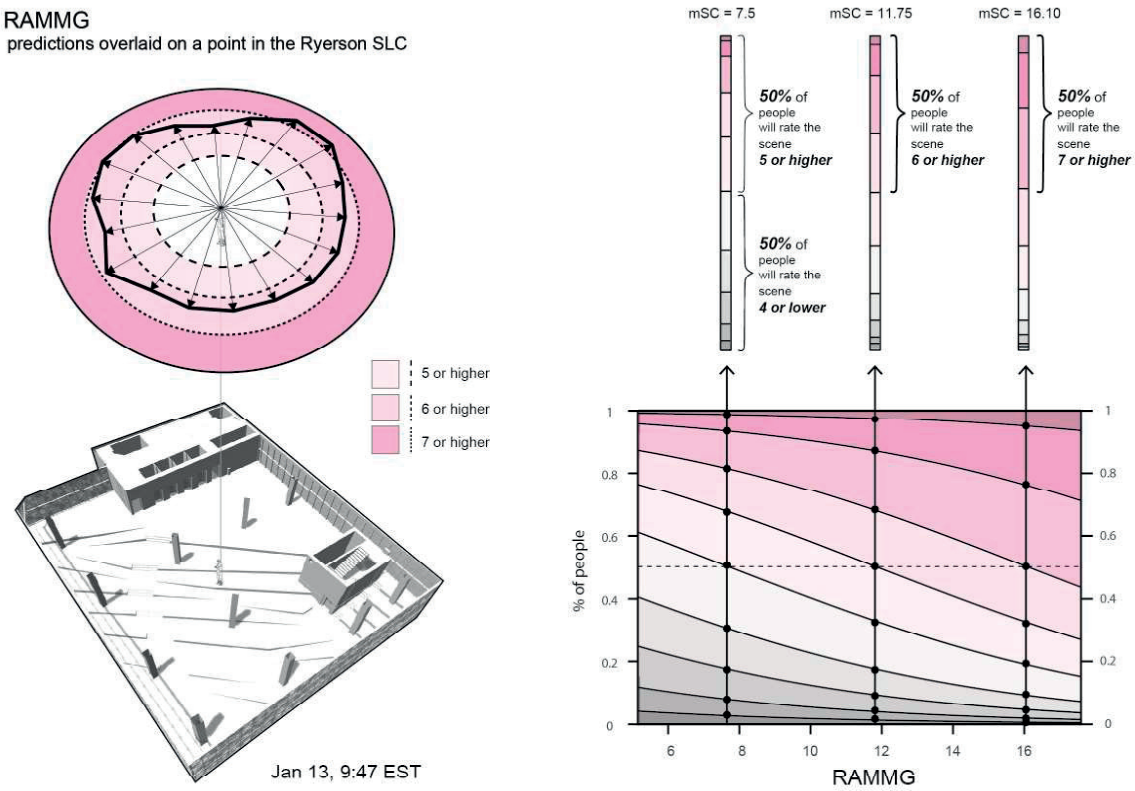


Figure 8.22 Shows the application of the logit model from Figure 8.19 to subject ratings of PIE across a range of view directions in the Ryersen SLC. A 50% population threshold was used to determine 3 hypothetical thresholds (7.5, 11.75, and 16.10) which correspond to higher subject ratings of PIE.

the relatively small number of subjects per scene condition in this experiment, Figure 8.22 does provide an example of how a composite indicator like PIE could be fit to subjective data from an immersive scene after further validation.

The insights gained from our preliminary assessment of the head-tracking data also suggest that a subjects' view behaviour should be accounted for in the development of future image-based prediction algorithms. The finding that subjects explored the 180° scenes primarily within a 45° wide horizontal band, centred in the field-of-view, is enlightening when we consider where the impacts of daylight-driven visual interest may have the most impact in architectural design from an occupant perspective. Future development of this immersive occupant-centric approach to qualitative lighting predictions can help designers understand the dynamic impacts of daylight on subjective appraisals of space across space and over time. A discussion on this issue will be elaborated on in the conclusion.

While the methods of computation are different, we can see that both mSC_5 and RAMMG predictions show relatively high correlation to subject ratings of PIE for both the 180° and 360° groups. So then, which prediction algorithm do we choose to provide immersive predictions of visual interest - the hemispherical, view-specific algorithm or the cube face algorithm which returns and average across the entire 360° field-of-view? On the one hand, an average value across the 360° view (using the cube face algorithm) provides for a compact prediction. On the other hand, this does not account for the variations in daylight and interest *across* our field-of-view. Due to the number of significant differences observed in subjective responses between 180° and 360° groups for the more asymmetrical architectural spaces in our study (described in Table 8.5), the selection of algorithms and method may depend on the goal of the analysis. If the aim is to evaluate a view position for a general (or average) impression of interest across the field-of-view and to see how that changes over time, then the box algorithm may be sufficient. If the aim is rather to evaluate a view position across many view directions to see how specific directions vary in visual interest, then the hemispherical approach is obviously preferred.

To end, it is also important to emphasize that the use of a composite rating like PIE does not represent the most robust measure of 'visual interest.' In Chapter 5, we used the bi-polar scale calming - exciting (which was also highly correlated to subdued - stimulating) to predict instances of visual interest. In this chapter, we averaged ratings of pleasant, interesting, and exciting, but is also possible to imagine a more nuanced approach, where a series of *different* algorithms can be used to predict a series of *different* subjective characteristics. Rather than try and reduce 'visual interest' to a single rating scale or a composite (such as 'PIE'), the author could see several emotional indicators overlaid to create a composite set of responses. This will be discussed further in the conclusion chapter (Section 9.2.3).

9 CONCLUSION & OUTLOOK

This conclusion will present a general overview of findings and contributions in each research phase (0-4 as proposed in Section 1.2.3), followed by a discussion of various topics impacting this work, and ending with an overview of next steps motivated by this thesis.

9.1 Findings & Contributions in Each Research Phase

The following section will present findings from the simulation-based study in Phase 0 to the immersive user-based study in Phase 4.

9.1.1 Phase 0: Evaluating Dynamic Contrast: A Proof-of-Concept Approach

In Phase 0, a proof-of-concept study compared the compositional contrast of daylight over time as a result of dynamic clear sky conditions in a series of simple geometry models. The implementation of tan algorithm called Spatial Contrast SC was first proposed by the author during her master thesis (Rockcastle, 2011). After conducting a survey of contemporary architecture to categorize compositional formalities of daylight and its variability over time, the author created 10 simple geometry models that represented a gradient of contrast and temporal variability from high to low. Using these simple geometry models, Phase 0 applied the algorithms mentioned above to an annual time-series of 56 rendered moments. From these simulations, a relative comparison was made using the results from each algorithm to discuss the dynamics of 'spatial contrast' for each geometry model over time. This compositional approach to computing contrast between neighboring areas within the rendered scenes and the relationship between those algorithms and the author's own intuition of contrast and temporal variability provided a promising new direction for daylight performance assessment. While the algorithm SC, first used to quantify local contrast in Rockcastle (2011) and Rockcastle, et al. (2014), was borrowed from computational methods in the field of computer graphics, the implementation of these methods into a simulation time-series to compare dynamic compositional characteristics of daylight in architecture was a novel contribution. At the time of this proof-of-concept study, a broader range of computational methods for measuring contrast-based characteristics was not explored, nor was a validation of these methods using subjective data beyond the author's own intuition due to time-constraints. This proof-of-concept approach offered a relative comparison of contrast-based characteristics to show the *potential* for evaluating daylit architectural models and provided a foundation for subsequent phases of research in this dissertation.

9.1.2 Phase 1: Comparative Study Using Contrast Algorithms

Building upon initial findings from the proof-of-concept study in Phase 0, Phase 1 presented a comparative study of algorithms developed to predict contrast perception in digital images. Using a selection of 'global' and 'local' algorithms introduced in Section 2.4, Chapter 4 applied a series of these algorithms to the 10 abstract geometry models presented in Phase 0. Values from each algorithm were normalized and then compared to find dependencies and or distinct differences between computational methods. Spatial Contrast (SC), proposed in Phase 0, was found to be highly correlated to an algorithm called RAMM1 and allowed us to simplify our computational approach in future steps, using the RAMM1 measure instead of SC. This comparative study also showed that several algorithms (and variations of those algorithms) were redundant and could be eliminated in future studies. Finally, the values from each algorithm were plotted against the author's intuitive ranking from low to high. Linear regression found weak correlations between the author's intuitive rank and global algorithms (RMS and Michelson), but found strong correlations to a selection of local algorithms (SC, RAMM1, RAMMG, DOG (Rc=1, Rs=2), DOG (Rc=1, Rs=4), and DOG (Rc=2, Rs=4)), identifying a group of neighborhood methods to be further evaluated. From this analysis, the author selected RAMMG, a multi-level neighborhood algorithm, for use in Phase 2, where it was applied to a rendered time-series. Using instances selected by this algorithm, a web-based experiment was used to gather subjective impressions of daylight composition.

9.1.3 Phase 2: Web-Based Experiment Using 2D Renderings

Phase 2 introduced the first of two experimental approaches used to collect subjective impressions of daylight in this thesis. Using an online survey, an experiment was designed to collect subjective ratings related to composition, contrast, and excitement using a 7-point bi-polar scale. Based on the literature presented in Chapter 2, several bi-polar scales were borrowed and others adapted to create a series of 7 rating scales: low contrast – high contrast, uniform – non-uniform, unvaried – varied, diffuse – direct, simple – complex, calming – exciting, and subdued – stimulating. Nine existing architectural spaces, which vary in composition and daylight distribution, were modeled and rendered to create a clear sky semi-annual time-series of tone-mapped renderings. Using the RAMMG algorithm from Phase 1, contrast was computed across each of the 28 semi-annual renderings to select high, medium, and low relative instances of RAMMG in each space. Groups of spaces were then organized so that each subject in our experiment would see a range of high, medium, and low contrast conditions across each of the nine architectural spaces. The survey was distributed using Survey Gizmo and collected 200 complete responses.

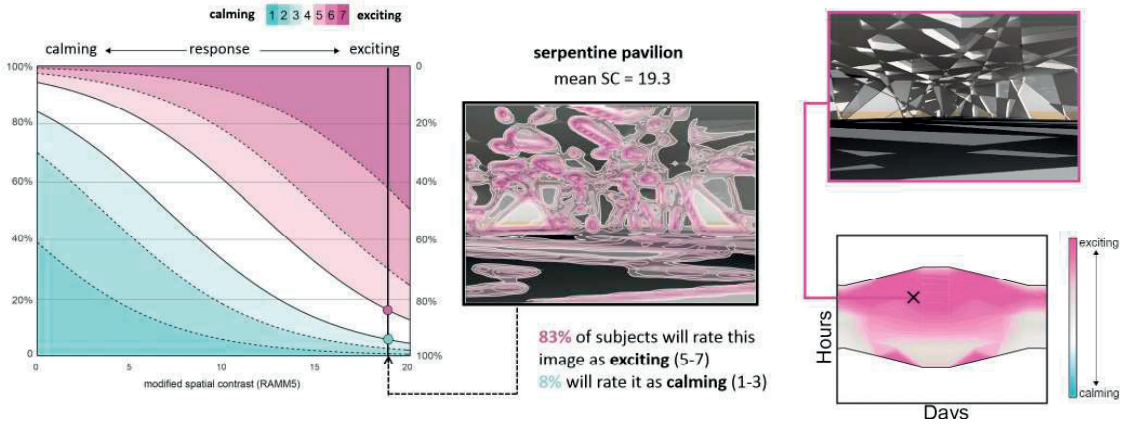
This Web-Based experiment resulted in the following findings & contributions:

1) Both space and sky conditions were found to have a significant effect on subject responses for ratings of low contrast - high contrast, diffuse - direct, simple - complex, calming - exciting, and subdued - stimulating.

2) Local neighborhood contrast measures such as RAMMG and specific levels within that metric (RAMM5, i.e. mSC_5) were found to be the best predictors of contrast-based visual effects, especially ratings of diffuse – direct, calming – exciting and subdued – stimulating, which supports the authors' initial findings in Phase 1.

3) Using ordered logistic regression between ratings of calming - exciting and results from the mSC_5 algorithm, this chapter introduced a method to predict the proportional odds of subject responses to attributes of visual interest using daylight renderings.

The mSC_5 algorithm and proportional odds model developed in this research phase allows architects and building performance engineers to predict attributes of visual interest for a single view across a series of annual instances. While at each instance, a single point-in-time quantitative analysis may be less useful to designers who can evaluate this performance qualitatively (Figure 9.1a), Modified Spatial Contrast becomes very useful in its ability to evaluate a time-series and predict changes in attributes which may be unanticipated due to dynamic sky and sun positions (Figure 9.1b). This annual method of predicting excitement can then be integrated alongside dynamic illumination and glare-risk metrics. By predicting how visually exciting and stimulating a space may be and how this changes over time, this research offers a new dimension in daylight performance assessment. Rather than be satisfied with the knowledge that a space achieves *enough* or *too much* daylight, the model proposed by this phase evaluates characteristics of daylight composition that impact emotional attributes in architecture. Building upon this work, Phase 3 will present an application method for simulating mSC_5 across an immersive *human-perspective*. Phase 4 will then evaluate the generalizability of this model to a broader range of architectural spaces and immersive view parameters.



a. b.

Figure 9.1 Shows a) the proportional odds model developed using subjective data in the Online experiment and the mSC_5 algorithm and b) the application of mSC_5 across an annual time-series to predict instances of calming, neutral, and exciting predictions based on dynamic sunlight.

9.1.4 Phase 3: Application of Visual Interest to a Simulation Workflow

Using the mSC_5 algorithm and proportional odds model developed from the online experiment in Phase 2, this phase presented a two-part approach for simulating attributes of visual interest across a range of view-directions in an architectural geometry model. In Chapter 6, the algorithm developed in Chapter 5 was adapted to a hemispherical image format and a predictive probability of 50% was used to compute threshold values for calming, neutral, and exciting using mSC_5 and the data provided by the online experiment shown in Figure 9.2a. Using a tone-mapped 360° HDR fisheye rendering from Radiance, *pinterp* was then used to extract a series of hemispherical fisheye renderings across 18 view directions in 20° radial increments. The adapted mSC_5 algorithm was then applied to each of the 18 resulting hemispherical renderings under both clear and overcast sky conditions to predict the degree of calming – exciting in each view direction (Figure 9.2b).

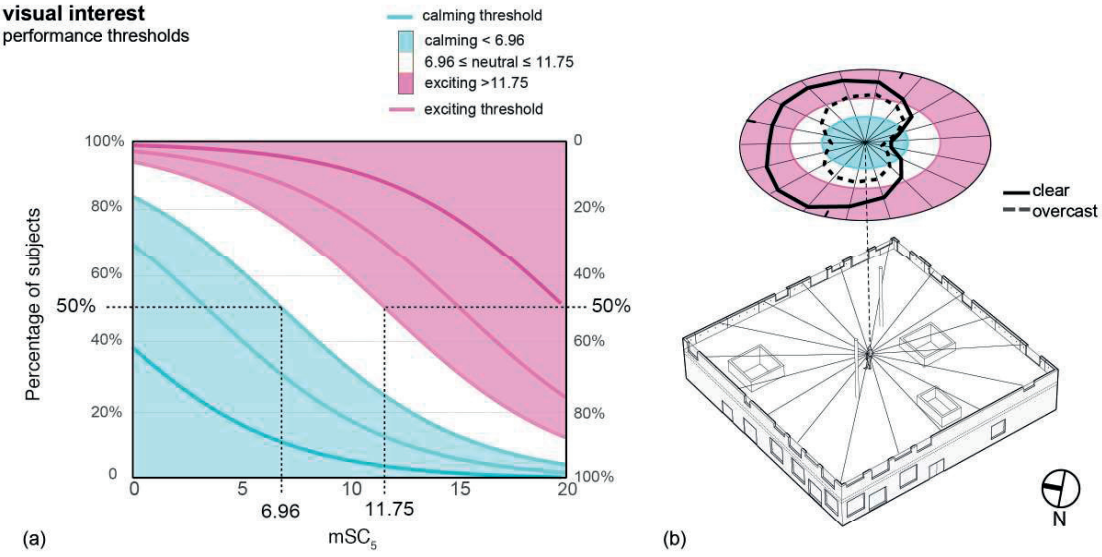
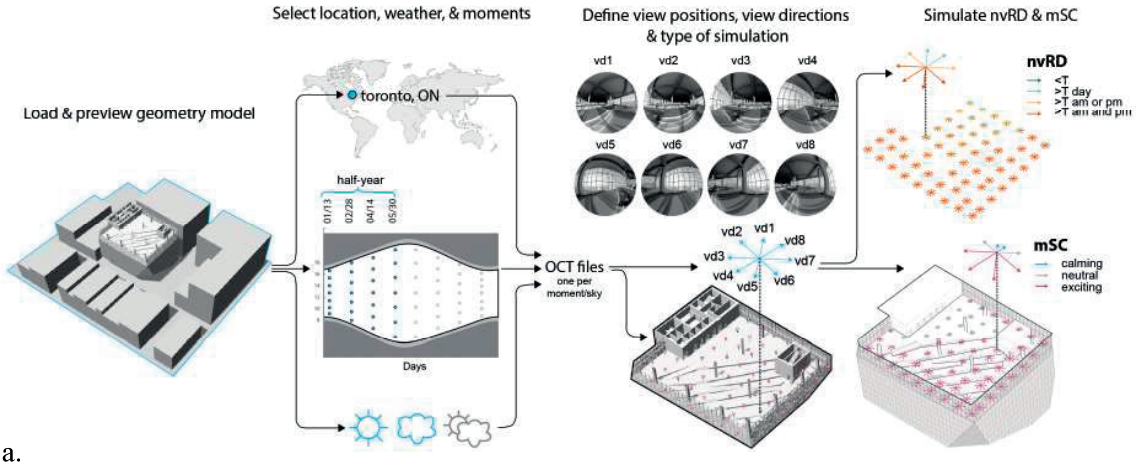


Figure 9.2 Shows a) the proportional odds model developed with threshold mSC_5 values for calming, neutral, and exciting based on a population percentage of 50% and b) the application of mSC_5 across 18 radial view directions in an architectural case study.

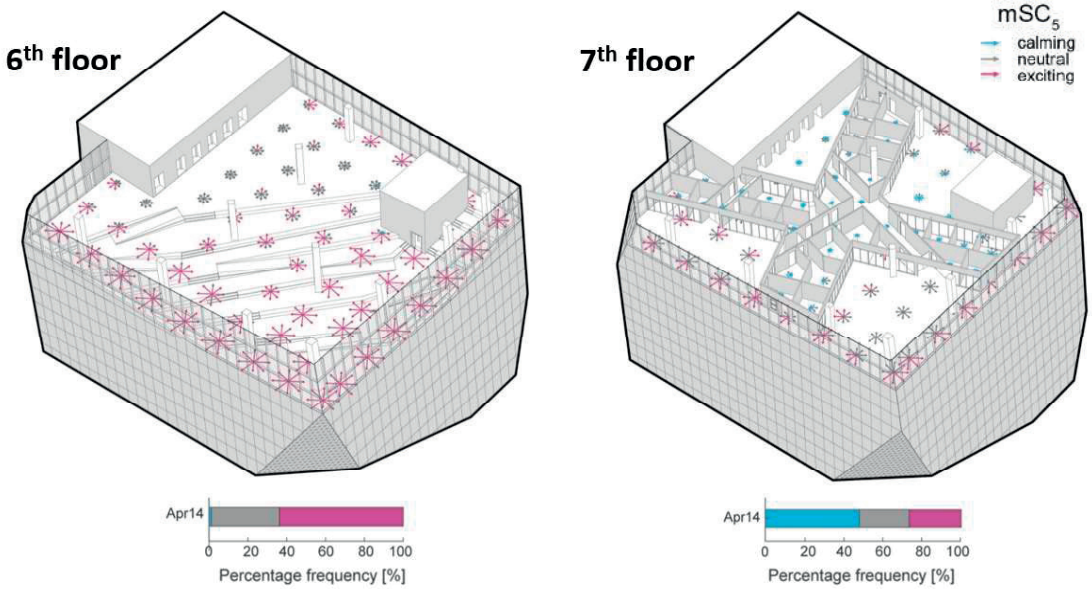
This method was then applied to a central view position in each of two case studies: the Zollverein School of Management and a side-lit test room in Fribourg over a time-series under overcast and clear sky conditions. The results show the effect of sky condition, view direction, and time-of-day on predicted instances of calming – exciting. The highly variable results illustrate the importance of temporally-dynamic and view-specific evaluations as the impacts of daylight vary greatly under these two inputs.

In the second part of this application phase, the method described in Chapter 6 was integrated into a new simulation-based workflow and platform, developed to evaluate mSC_5 alongside a non-visual health-based model called nvR_D. Where the method in Chapter 6 was applied to a single central view position, Chapter 7 introduced a workflow to evaluate an array of view positions. In

this platform, inputs for geometry, site, weather, location, view position, and view direction were streamlined into a simple user-interface whereby batch scripts for Radiance were generated and the simulation of many points was optimized using shared ambient files between multiple view positions in a single geometry file (Figure 9.3a). Using the threshold values for calming, neutral, and exciting established in Chapter 6, the results for visual interest (in the case of calming – exciting) were then computed across a time-series of 360° renderings under clear and overcast sky conditions. In this chapter, results for mSC_5 were presented as both daily averages (Figure 9.3b) and instantaneous predictions alongside daily cumulative responses for the nvR_D non-visual health model, introduced in Chapter 7. This multi-point, multi-view-direction, multi-sky condition approach introduced a new paradigm in human-centric daylight evaluation and visualization,



a.



b.

Figure 9.3 Illustrating a) simulation workflow for applying mSC_5 and nvR_D to an immersive set of view positions and b) the average daily mSC_5 under clear skies on the 6th (left) and 7th floors (right).

providing dynamic performance evaluations in a relatively user-friendly simulation-based workflow. The implications of this human-centric analysis method and simulation workflow are widespread, with the potential to shift the paradigm of daylight performance from an orthographic plan or section to the visual experience of an occupant in a building.

9.1.5 Phase 4: Visually-Immersive Experiment in Virtual Reality

The final phase of this thesis presented an experiment using visually-immersive rendered scenes in a virtual reality headset. To test the effect of visual immersion and sky condition factors on subjective ratings for a selection of attribute ratings, two groups were established. Group one was asked to rate semi-immersive 180° scenes and group two was asked to rate fully-immersive 360° scenes. Eight architectural case studies were selected for this experiment. A semi-annual time-series of 28 clear-sky instances was rendered using a 360° hemispherical fisheye format, and a 3D version of the mSC_5 algorithm called the 'cube face' adaptation was used to select instances of high mSC_5 from each time series. This same instance was then rendered under overcast skies. From these instances of clear and overcast sky, a series of view directions were extracted from each rendered scene and evaluated using the hemispherical mSC_5 algorithm presented in Phase 3. High and low mSC_5 view directions were selected for each scene and sky condition to make up the scene variations presented to subjects in each group.

In our design of experiment, subjects were randomly assigned a group, either semi-immersive 180° or fully-immersive 360° and then asked to rate a selection of scenes using a 10-point unipolar rating scale, delivered verbally while subjects explored the scenes in VR. In the 180° group, subjects saw each scene under one of four random view direction and sky condition parameters. Subjects in the 360° group saw each scene under one of two random conditions. Using the Oculus Rift CV1 headset, semi and fully immersive scenes were then shown to subjects who were asked to rate the scenes.

Results of this experiment found that architectural space had the most significant effect on ratings of pleasant, interesting, and exciting for both the fully and the semi-immersive groups. Sky and view parameters, on the other hand showed the most significant effects on ratings of pleasant, calming, diffuse, and contrasted. When comparing ratings between the view and sky parameters in the semi-immersive scenes, the most significant differences occurring in asymmetrical spaces like Douglas, Ryerson, and Spencertown. A comparison between median subject responses and algorithms used to rank them (RAMM1, mSC_5, RAMMG, mean brightness, and RMS), mSC_5 and RAMMG showed the highest positive correlation to ratings of pleasant, interesting, exciting, and contrasted for both groups. Using linear and logistic regression, we also find high correlations between mSC_5 and RAMMG algorithms and the responses for a composite rating called PIE (the average between ratings of pleasant, interesting, and exciting, which were found to be the most highly correlated).

While it is tempting to compare the results from this VR experiment to those from the online experiment in Phase 2, the shift from a bi-polar to a uni-polar scale and the disparity in subjects

per experiment make it difficult. In the Online experiment, a minimum of 50 subjects rated each image, but only 15-20 subjects rated each of the immersive scene conditions in this VR experiment. Still, we can see that both mSC_5 and RAMMG provide relatively good fits between subject responses for ratings of excitement and related factors in both experiments, out-performing all other algorithms under consideration.

Between the 180° and 360° groups, median ratings for the 360° scenes were all moderately to highly correlated to those of the 180° scenes. In a comparison between the 180° and 360° groups, and rankings from the mSC_5 and RAMMG algorithms, we found that ratings of pleasant and PIE were more highly correlated in the 180° scenes than the 360° scenes. Ratings of interesting and exciting were equally correlated in both groups or more highly correlated in the 360° scenes.

Head-tracking data revealed some interesting findings regarding the frequency of view activity in this experiment. From an initial overview of view direction vectors across all subjects and scenes in the 180° group, it appears that the majority of activity was spent viewing the middle 50° band (+/- 25° from the horizon of Oculus Rift CV1). Despite higher localized mSC_5 regions in some scenes, like the roof of the Menil Gallery, the highest frequency of view activity was still around the middle 50° horizontal band and not on the ceiling as would be expected. This may help explain why the mSC_5 algorithm was less effective in predicting responses in the Menil Gallery, where a view-dependent weight centered on a specific area of the image may better predict view-specific tendencies.

The intention of this research phase was to evaluate whether mSC_5 and RAMMG algorithms (adapted to the specific scene parameters) may serve as a reasonable predictor of visual interest in visually-immersive scenes. While the findings of this experiment support findings from the online experiment in Chapter 5 and confirm that adapted algorithms can serve as a reasonable predictor of visual interest in 180° and 360° scenes, an equally important contribution of this research phase may be the experimental method itself. Using the protocol first proposed by Chamilothori, et al. (2016) and further developed in collaboration with the author for this experiment, one of the main contributions of this research is the method of collecting subjective *and* objective data in daylit architecture from visually-immersive virtual scenes. The ability to extract head-tracking data alongside subjective ratings of daylight in rendered architectural scenes provides researchers with a new virtual laboratory, ripe with possibilities for experimental work. Experiments in real world scenes provide a broader range of luminance levels, but rendered scenes allow the researcher to manipulate an infinitely greater number of parameters while virtual reality headsets like Oculus allow the subjects to feel more realistically immersed in the environment. The contributions of this research and the potential for future interdisciplinary work that builds upon it's foundation have immediate and valuable impact on the fields of engineering, psychology, computer graphics, and architecture.

9.2 Discussion

9.2.1 Comparing Online vs. Immersive Experimental Studies

This thesis introduced a pair of user-based experiments with the aim of collecting subjective responses to ratings of visual interest. Using algorithms developed to predict contrast perception in digital images using a number of different computational techniques, quantitative contrast results were then compared to subjective responses from these experiments to search for dependencies. The aim of this comparison was to find and/or develop an algorithm capable of predicting subject responses to characteristics of visual interest in daylit scenes. Depending on the degree of visual immersion in the digital scene (Chapter 5 used 60° x 80° rectangular images, while Chapter 8 used semi-immersive 180° fisheye and fully-immersive 360° fisheye images) two algorithms - mSC_5 and RAMMG - emerged as possible contenders based on their relationship to subjective ratings of pleasant, interest, excitement, and contrast. To collect these responses, the 2D online experiment in Chapter 5 used semantic bipolar scales while the immersive VR experiment in Chapter 8 switched to unipolar scales. We made the decision to switch from bi-polar to uni-polar scales in the immersive experiment to explore whether ratings of calming and exciting were, indeed, bi-polar as assumed in our online study. We also shifted from a 7-point to a 10-point scale as a result of the virtual reality headset and the need for a verbal response scale based on this display medium. The bi-polar 7-point scales used in the online experiment had each rating (i.e. calming and exciting) displayed on either end with 3 bubbles to the left of center and 3 bubbles to the right. When we tested a 7-point rating scale in an informal pilot using the VR headset, subjects commented on the difficulty associated with visualizing a 7-point scale. When asked how we may improve the experiment, a 10-point scale was recommended.

While the decision to switch from 7-point bi-polar to 10-point uni-polar ratings was logical based on specific experimental parameters, this made it impossible to compare subject responses and algorithm predictions between experiments. In hindsight, it would have been advantageous to use the same rating scales between experiments so that a full comparison between images and contrast algorithms would have been possible. Furthermore, some architectural spaces were redundant between the online and virtual experiments (Serpentine, Zollverein, Poli, Menil, and First), but the algorithm used to select moments under clear sky conditions for each experiment shifted from RAMMG to mSC_5. In addition to using the same rating scale between experiments, the author would consider selecting the same view directions and time-series instances so that view parameters (2D rectangular, 180° fisheye, and 360° fisheye) and display medium (online uncontrolled vs. Oculus Rift CV1) could be directly compared.

9.2.2 Display of HDR Images & Tone-Mapping Considerations

In both the online and immersive user-based experiments, HDR renderings were tone-mapped down to display screen luminance ranges (0.5 to 200 cd/m^2 in the online experiment and 0 to 80 cd/m2 in the VR experiment) using the *pcond* tone-mapping operator with the contrast sensitivity

function -*s* (Ward, 1997). The *pcond* operator was selected based on default settings in the online experiment, but was used again in the VR experiment. While studies by Yoshida et al. (2005) and Kuang et al. (2004) have recommended the use of local tone-mapping operators like Reinhard (2002), studies by Ćadík et al. (2008) found global tone-mapping operators such as *pcond* by Ward (1997) and Drago (2003) were rated higher in overall image quality. While we wanted to test the effect of various tone-mapping operators on subject ratings in the immersive VR experiment, it was not possible to adapt Reinhard or Durand to the specific 360° angular fisheye format used in our study. *pcond* was the only algorithm with native implementation in Radiance, allowing for a visually acceptable compression of luminance levels, where the use of Reinhard and Durand resulted in grey muddled scenes due to the inclusion of black masking pixels in the corners of our fisheye images. Future work may seek to implement a range of alternate TMOs like Reinhard to a 360° image format. Ongoing work by Chamilothori, et al. (forthcoming) seeks to compare the effects of both local and global tone-mapping operators on ratings of excitement, interest, and contrast in 180° photographs, visualized in the Oculus Rift CV1.

To evaluate the impacts of tone-mapping on predictions of visual interest using mSC_5 or RAMMG, a quick comparative analysis was done using the *pcond* mapping algorithm *with* contrast sensitivity function (*s*+) and *without* contrast sensitivity function (*s*-) on both BMP and HDR image formats. As was briefly discussed above, the potential impact of tone-mapping operators TMOs on the perceptual accuracy of *immersive images* in VR is relatively unknown. In the VR experiment presented in Chapter 8, the *pcond* algorithm was used due to the ease of implementation on our 360° fisheye image format. Rather than enable the human visual response output -*h* function in *pcond,* which packages -*a* -*v,* -*s,* and -*c* functions, we manually enabled the contrast sensitivity -*s* and color -*c* functions. This decision was made to ensure the most accurate human contrast and color response, without implementing those functions that would defocus or alter local regions of the image based on the view direction of the 360° fisheye. As the 360° fisheye had to be tone-mapped before using *pinterp* to unroll new view-directions, we couldn't use any mapping function that would alter localized regions of the image dependent on view direction (i.e. -*a* or

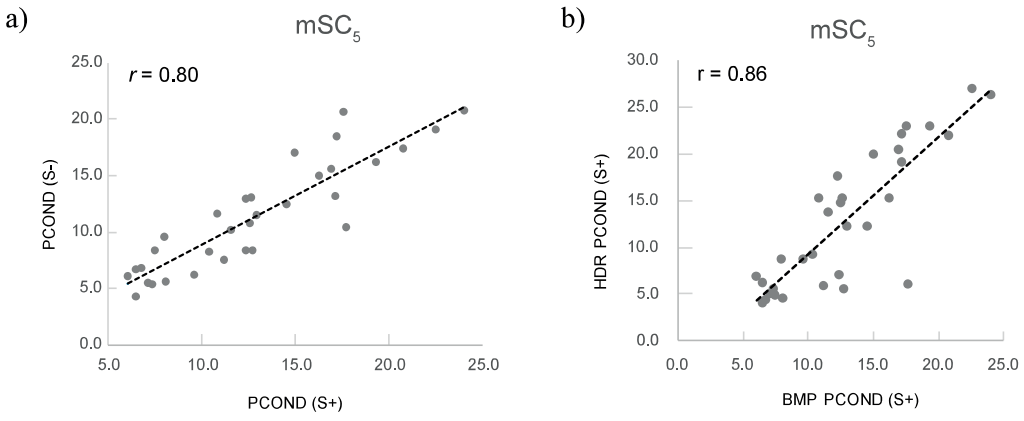


Figure 9.4 Pearson Correlation Coefficient values (*r*) for a) PCOND (s+) vs. PCOND (s-)for mSC_5 results both using a BMP compression and b) HDR vs. BMP for mSC_5 results both using PCOND (s+).

-*v* functions). While we were unable to implement and compare TMOs that were non-native in Radiance (i.e. Reinhard, Durand, etc.) while using this specific image format, a comparison between images mapped using *pcond* with and without the contrast sensitivity function (*s*+/-) allowed us to compare the impact of that function on results for mSC_5 and RAMMG.

Figure 9.4 shows Pearson Correlation Coefficient (*r*) values between results for mSC_5 on BMP images using *pcond* with and without the -*s* function enabled ($r = 0.80$) and for mSC_5 using *pcond* with the -*s* function enabled on HDR vs. BMP image compressions ($r = 0.72$). The full table of results can be found in Appendix A5.1 for BMP images and Appendix A5.2 for HDR images (with results multiplied by 255 to equal the same range as BMP images). One-way ANOVA results can be found in Appendix A5.3 for the comparison of image compression and contrast sensitivity function where no significant effects were observed between groups. While we cannot compare subjective data from our experiment to the results for *pcond* without -*s* enabled or for HDR formats (as subjects only saw the BMP images with -*s* enabled), the Pearson values in Figure 9.4 reveal a potential source of error for future implementation. If the mSC_5 measure is applied to a rendering without -s enabled, the thresholds established in Section 6.3 might not be precisely accurate for that computed data set. While we continue to develop the metrics for predicting elements of visual interest in daylight renderings, future work should also explore the impacts of tone-mapping algorithms and specific functions there within on subjective responses to see if they scale in a predictable manner and to determine the most appropriate image format.

Another topic of discussion raised by this thesis is the use of static TMOs on immersive HDR renderings, where contrast and brightness across the field-of-view often vary significantly. Studies that have tested the perceptual accuracy of available tone-mapping operators have used 2D photographs and renderings, but we noticed that our tone-mapped 360° scenes (especially in some view directions) often appeared too dark or too light. We speculate that a dynamic, view-dependent tone-mapping algorithm would need to be developed for use in 360° virtual scenes to achieve the same perceptual accuracy as those used on 2D rectangular images.

9.2.3 Defining 'Visual Interest'

A final topic of discussion related to this research is how we define 'visual interest' in architecture and the challenges associated with trying to relate it to a series of independent, albeit related, attributes like excitement, stimulation, complexity, and contrast. The author of this thesis has struggled to establish consistent terminology in this thesis, due to the relative fluidity of semantics surrounding the perceptual attributes we are trying to measure and simultaneously relate. In other words, we have associated attributes such as excitement and stimulation with visual interest, but find ourselves questioning this approach. Instead of combining attributes that appear to be correlated and isolating one algorithm to describe the group, it would be interesting to also consider a discreet approach. If, for example, we could find four different algorithms that are more specifically related to four different attributes such as excitement, calming, dullness, and

pleasantness, then we could offer a toolbox of emotional indicators that may be overlaid.
Related to this issue is the question of lower and upper limits for visual interest. In our online study, bi-polar ratings were between calming - exciting, neither of which offers a negative association. Does that mean that 1 out of 7 is the most calming? Perhaps in the context of our experiment, where there were arguably no 'bad' architectural spaces, this would be valid, but it would be interesting to know when a space goes from calming to boring, or from exciting to disturbing.

9.3 Next Steps

As mentioned briefly in Section 8.7, future work will analyze the full head-tracking data set collected in Phase 4 to reveal specific relationships between characteristics of contrast and composition in the virtual scenes and view direction behavior from our test subjects. While this head-tracking data does not account for the same level of detail as eye-tracking methods would allow, the potential for collecting view behavior during subjective experiments in immersive virtual scenes opens the door for an exciting set of opportunities for both architects and experimental researchers. Based on the findings from this data set, future work can seek to develop more sophisticated algorithms for predicting perceptual responses to daylight on architectural composition using view-dependent weights or focusing on specific areas within the field-of-view.

Related to our initial observations in the VR experiment and based on comments recorded during each session, we will work toward a new VR-specific protocol for delivering verbal questions that improve future head and behavioural tracking results. Ensuring that subjects feel realistically immersed in the virtual world will also allow for the future study of other environmental factors like air temperature, radiant heat, relative humidity, air speed, and acoustical stimuli.

Other areas of future experimental work include studying the effects of program use (space function) on daylight-driven impressions. As all furniture and non-structural elements were removed from the rendered scenes used in this thesis, we are unable to speculate how subjective impressions of excitement, calm, interest, or pleasantness may be impacted by program use (i.e. office vs. museum vs. entry hall). Using a similar methodology to that introduced in Phase 4, it would be possible to repeat a similar experiment with scenes that vary in furniture layout to imply a specific space use and measure the presence of any effects on subjective impressions.

Building upon the discussion in Section 9.2.3, it would also be interesting to relate algorithms like mSC_5 and RAMMG to negative lower (i.e. boring) and upper (i.e. overwhelming or uncomfortable) limits. Future work should explore the relationship between daylight glare metrics such as DGP and higher level visual interest predictions. Could a glare metric like DGP help to define an upper limit for visual interest? Do these predictions even relate to one another? Collecting subjective data to validate this hypothesis would require assessments in real space or a display medium capable of inducing high-level luminances. Field studies on glare could seek to capture HDR photographs and assess the scenes using mSC_5 and RAMMG algorithms to search for thresholds

over which daylight conditions are likely to move from exciting to visually disturbing. As our current algorithm works with tone-mapped grey-scale images, this would also require a weighting method for high-level luminance ranges in HDR format.

Phase 3 presented two application-based studies to implement mSC_5 predictions across a range of view directions and view positions in an architectural case study. Future developments in this work will seek to integrate illumination (instantaneous illuminance and climate-based illuminance-based metrics) and comfort (DGP and or gaze-driven metrics) into a single simulation platform. While many existing software tools integrate daylight illumination and comfort metrics (DIVA-for-Rhino, Lightsolve, Sefeira, etc.), they have been developed to assess a horizontal array of sensors or single view positions for instantaneous and/or annual comfort assessments. Based on the novelty of predicting visual interest using an immersive mSC_5 algorithm alongside nvR_D for non-visual health potential (presented in Chapter 7), the value of a human-centric simulation approach can revolutionize the way we predict daylight performance. Where previous metrics have sought to measure the potential for daylight to provide useful task illumination while limiting discomfort, this thesis introduces a complimentary method for predicting the emotional impacts of daylight from a *human-perspective*. While this thesis does not suggest that all daylight performance goals can and should be achieved in parallel - as good performance in some criteria may contradict others - it does aim to balance the available tools to provide a more holistic approach. This approach transforms the dialogue of daylight performance from the building to the *experience* of an occupant *in that building*. The author is currently working on improving the simulation-based platform in Chapter 7 to provide immersive visualizations of performance alongside temporal representations for annual evaluations.

A field-based study in architectural offices is currently under development for later this year. To evaluate the value and impact of prediction models proposed in this thesis, the author plans to provide assessments of visual interest alongside a range of other more traditional daylight performance measures, for architectural projects in the design development phase. By collecting feedback from architectural designers during the design process, we hope to transition this work from an academic environment to the professional design and consulting world. This leap from primary research to application will be facilitated by the simulation-based workflow and platform presented in Chapter 7. Working with architects and consultants, the author will implement dynamic visual interest analyses on their projects using a new simulation-based tool currently under development. Future work will strive to close the loop which began with the final objective of this thesis (presented in Section 1.2.2). That objective it to provide architects with the ability to evaluate the perceptual performance of their own designs alongside existing performance measures for a more holistic assessment of daylight in architecture.

A1 APPENDIX: CHAPTER 4

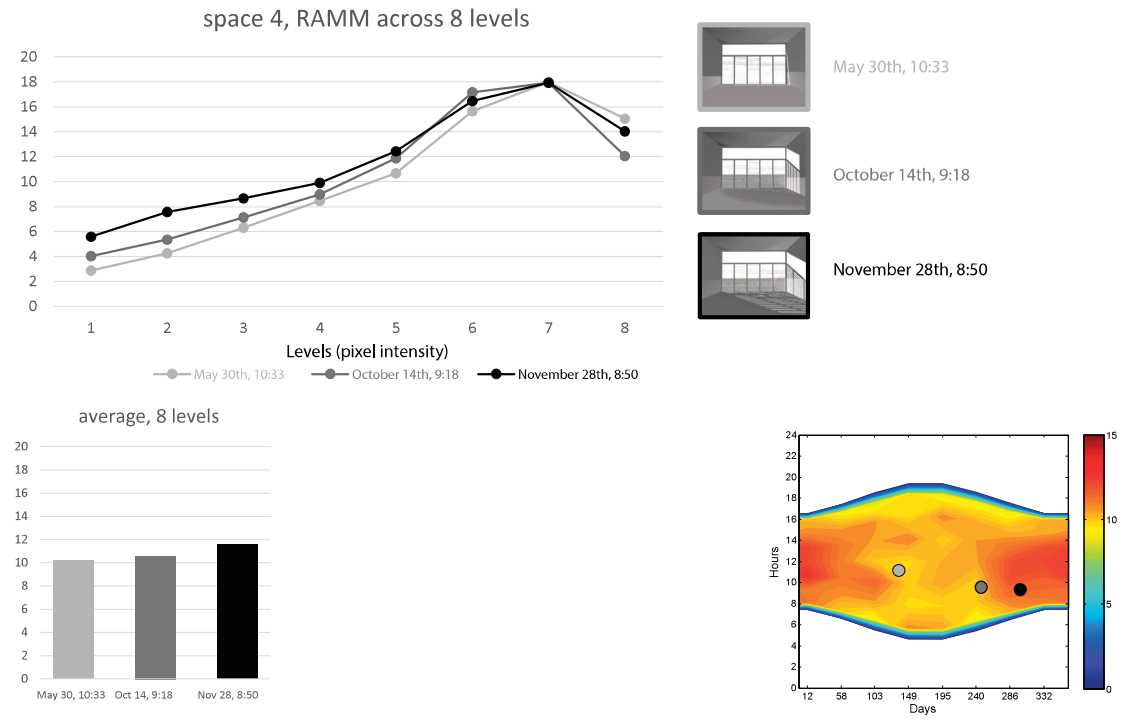
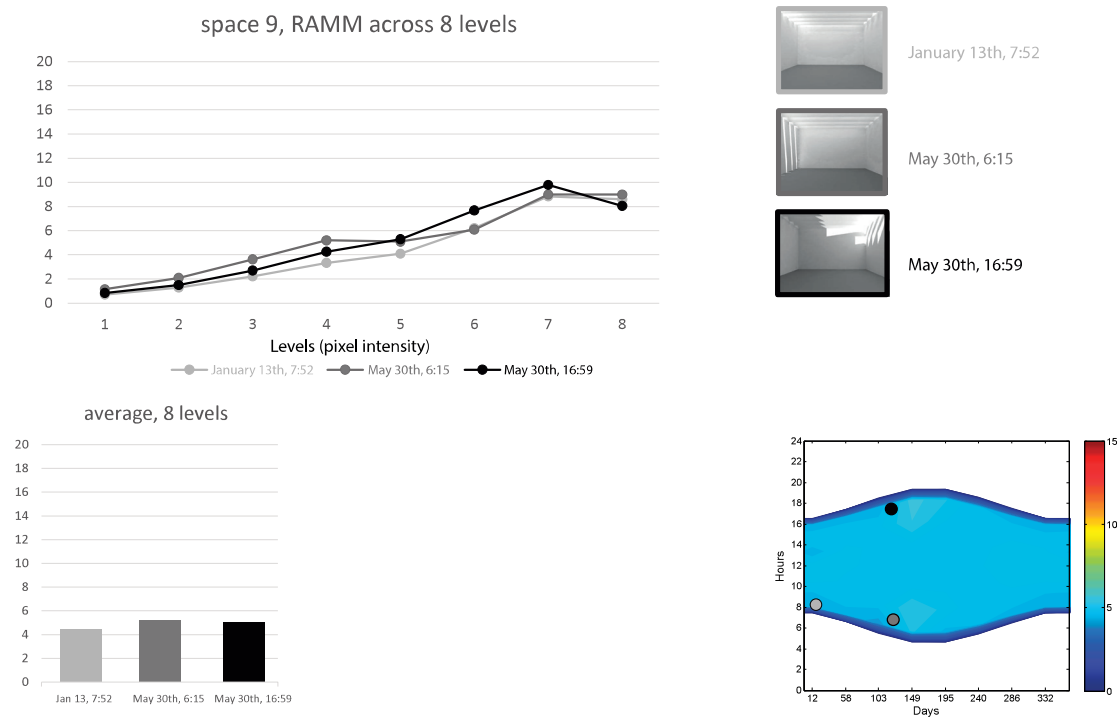


Figure A1.1 RAMM results from space 4



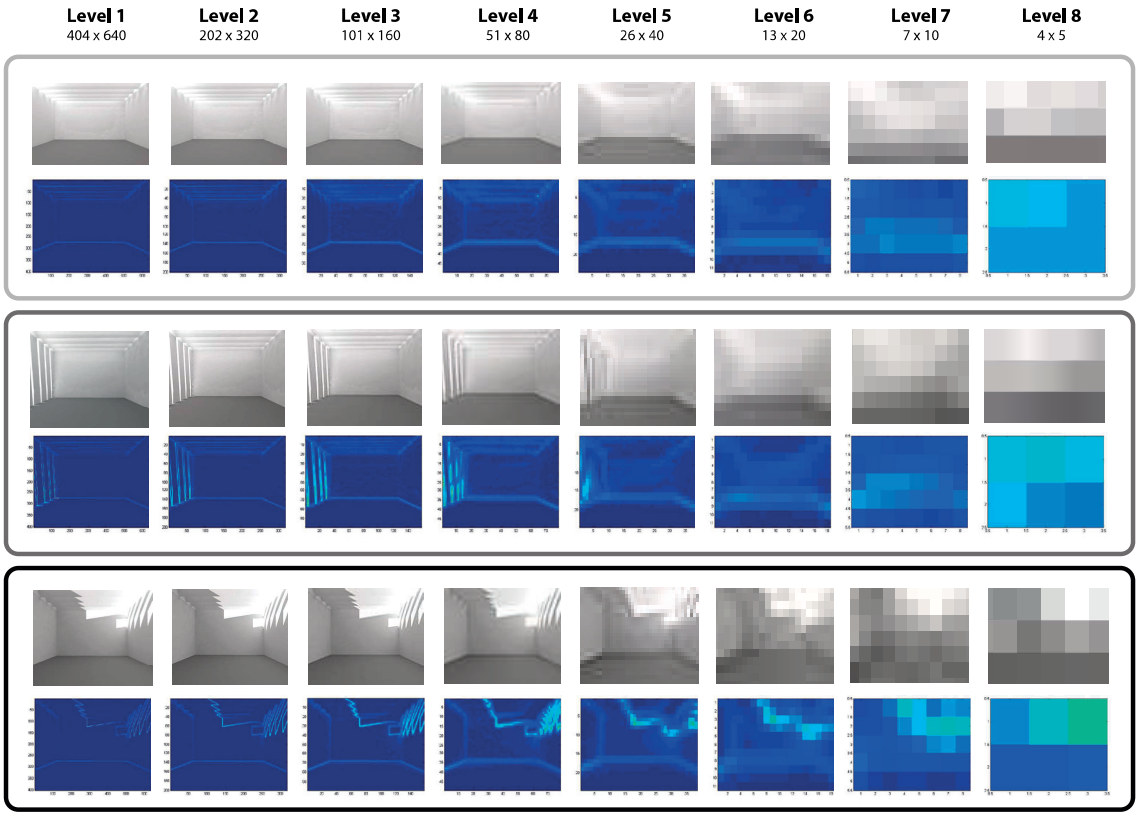


Figure A1.2 RAMM results from space 9

A2 APPENDIX: CHAPTER 5

low contrast – high contrast

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	24.765	8	3.0957	1.23	0.2995
sky	**37.062**	**2**	**18.5309**	**7.34**	**0.0014****
space	29.877	8	3.7346	1.48	0.1835
error	156.617	62	2.5261		
total	248.321	80			

uniform – non-uniform

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	20.444	8	2.55556	0.79	0.6156
sky	7.63	2	3.81481	1.18	0.3155
space	46.667	8	5.83333	1.8	0.0947
error	201.259	62	3.24612		
total	276	80			

unvaried - varied

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	23.556	8	2.9444	0.99	0.4517
sky	**24.963**	**2**	**12.4815**	**4.2**	**0.0194***
space	**55.556**	**8**	**6.9444**	**2.34**	**0.029***
error	184.148	62	2.9701		
total	288.222	80			

simple - complex

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	15.778	8	1.9722	0.96	0.4746
sky	**34.296**	**2**	**17.1481**	**8.35**	**0.0006****
space	**122.222**	**8**	**15.2778**	**7.44**	**<0.000****
error	127.259	62	2.0526		
total	299.556	80			

calming - exciting

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	27.432	8	3.429	1.41	0.2102
sky	**20.469**	**2**	**10.2346**	**4.21**	**0.0193***
space	**98.099**	**8**	**12.2623**	**5.04**	**0.0001****
error	150.765	62	2.4317		
total	296.765	80			

sedating - stimulating

Source	*Sum Sq.*	*d.f.*	*Mean Sq.*	*F*	*Prob>F*
subject	23.778	8	2.9722	1.66	0.1257
sky	**21.407**	**2**	**10.7037**	**5.99**	**0.0042****
space	**66**	**8**	**8.25**	**4.62**	**0.0002****
error	110.815	62	1.7873		
total	222	80			

*Prob>F is less than 5% ** Prob>F is less than 1%

Figure A2.1 ANOVA results from pilot study

ANOVAN

low contrast - high contrast

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	123.024285	2	61.5121424	26.2294704	**6.38E-12**
space	1340.63042	8	167.578803	71.4574891	**1.28E-99**
group	58.9666488	2	29.4833244	12.5720216	**3.85E-06**
error	3494.27918	1490	2.34515381		
total	5016.90053	1502			

non-uniform - uniform

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	57.497006	2	28.748503	10.5090604	**2.94E-05**
space	1266.33672	8	158.29209	57.8639221	**2.74E-82**
group	24.8376026	2	12.4188013	4.53971232	0.01082457
error	4076.03227	1490	2.73559213		
total	5424.70359	1502			

unvaried - varied

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	15.8087159	2	7.90435795	3.15883806	0.04275966
space	1331.44283	8	166.430353	66.5109724	**2.04E-93**
group	9.05351769	2	4.52675885	1.8090398	0.16417096
error	3728.42581	1490	2.5022992		
total	5084.73087	1502			

diffuse - direct

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	180.802728	2	90.4013639	33.2544577	**7.43E-15**
space	1411.82782	8	176.478477	64.9182245	**2.15E-91**
group	30.0737228	2	15.0368614	5.53136201	0.00404235
error	4050.52561	1490	2.71847356		
total	5673.22987	1502			

simple - complex

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	68.0701929	2	34.0350965	14.6321587	**5.09E-07**
space	1975.1442	8	246.893025	106.142726	**3.99E-140**
group	12.6158528	2	6.30792641	2.71186481	0.06674065
error	3465.81081	1490	2.32604752		
total	5521.64105	1502			

calming - exciting

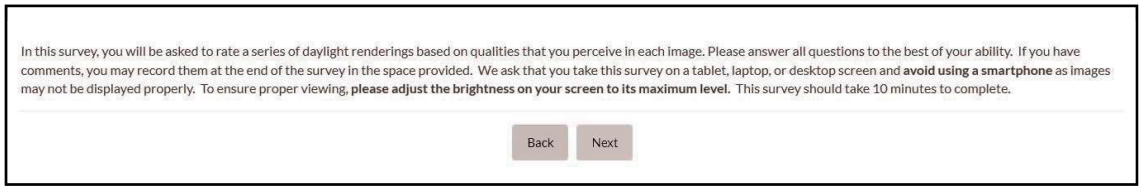
source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	82.2448436	2	41.1224218	19.8742172	**3.03E-09**
space	1664.33274	8	208.041593	100.545241	**6.46E-134**
group	4.32401933	2	2.16200967	1.04488617	0.35198944
error	3083.0099	1490	2.06913416		
total	4833.91151	1502			

subdued - stimulating

source	*sum of sq.*	*d.f*	*mean sq.*	*F*	*Prob>F*
sky	74.9554225	2	37.4777112	18.0927024	**1.42E-08**
space	1637.22434	8	204.653042	98.7980978	**5.92E-132**
group	4.51816886	2	2.25908443	1.09059334	0.33628513
error	3086.42615	1490	2.07142695		
total	4803.12409	1502			

Figure A2.2 ANOVA results from Online study

page 1



page 2

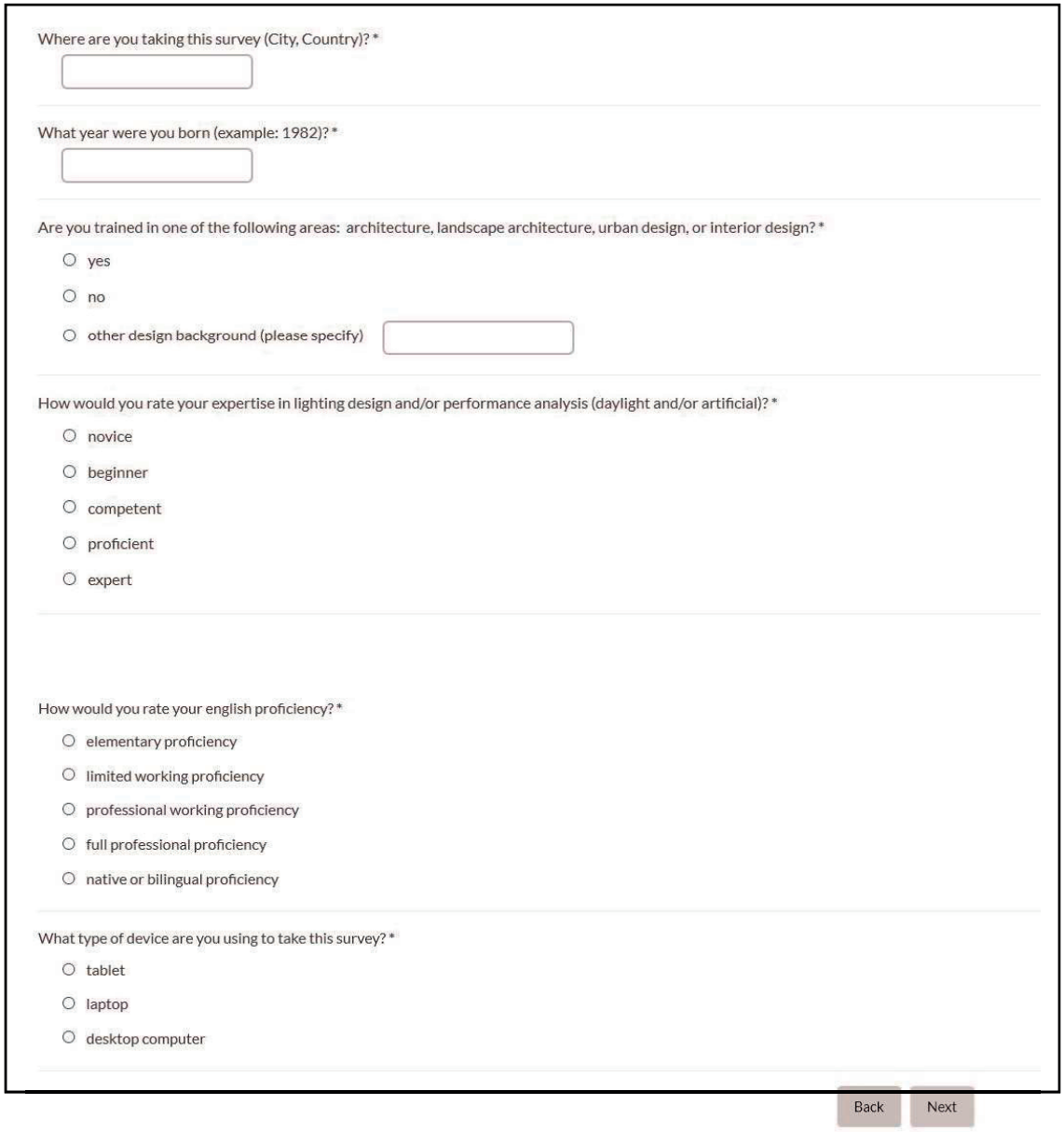
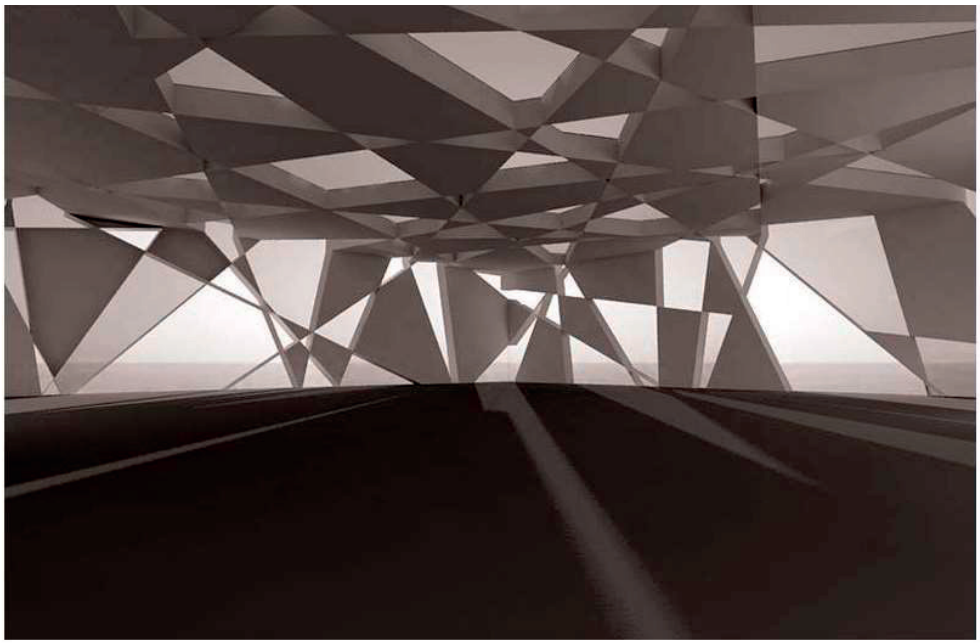


Figure A2.3 Online Survey

pages 3 - 12



How would you rate the daylight composition in this image based on the following: *

	1	2	3	4	5	6	7	
low contrast	○	○	○	○	○	○	○	high contrast
uniform	○	○	○	○	○	○	○	non-uniform
unvaried	○	○	○	○	○	○	○	varied
diffuse	○	○	○	○	○	○	○	direct
simple	○	○	○	○	○	○	○	complex
calming	○	○	○	○	○	○	○	exciting
subdued	○	○	○	○	○	○	○	stimulating

Back Next

Figure A2.3 continued...

pages 13

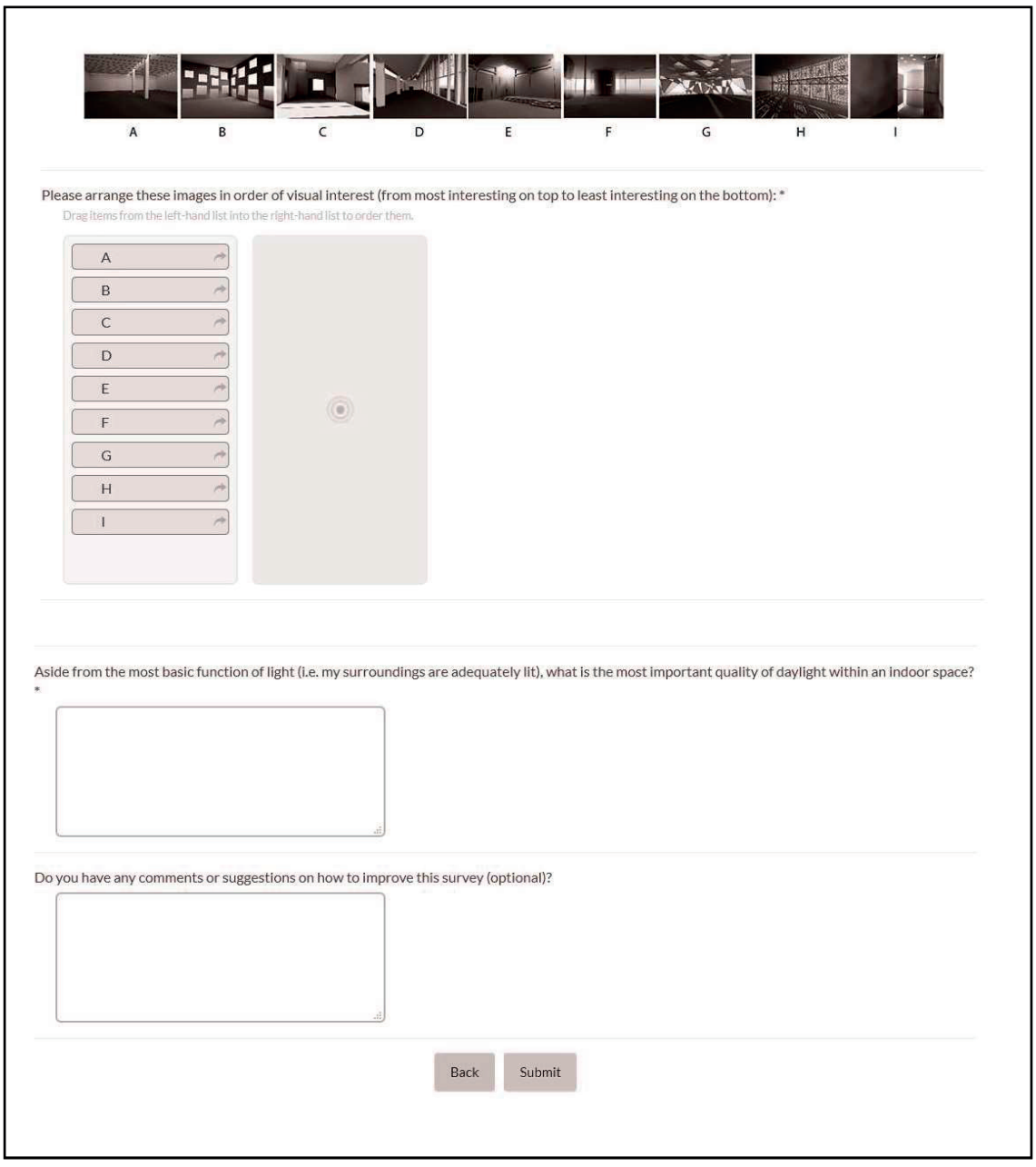


Figure A2.3 continued...

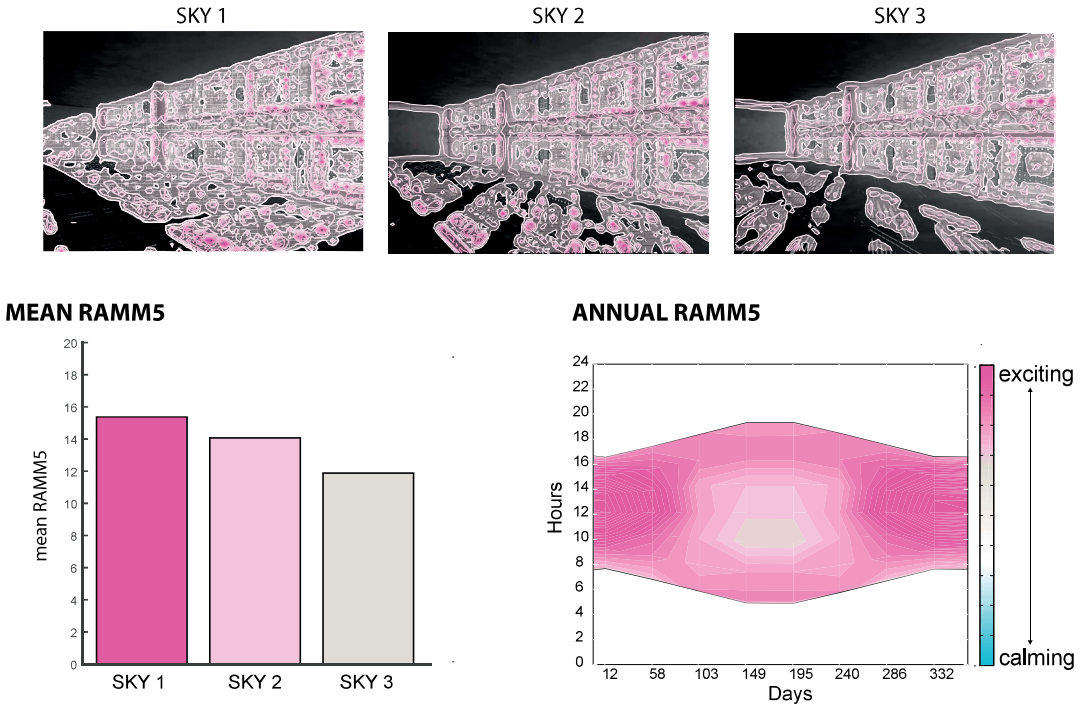


Figure A2.4 RAMM5 (mSC) results for the Arab World Institute by Jean Nouvel

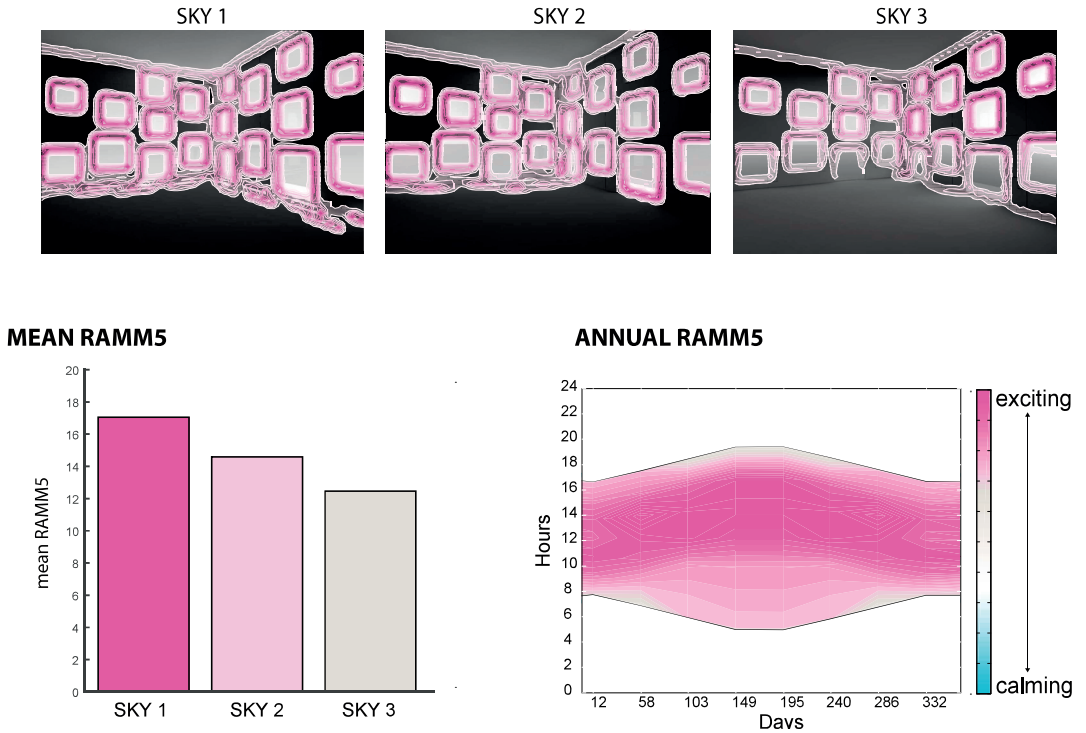


Figure A2.5 RAMM5 (mSC) results for the Zollverein School of Management by SANAA

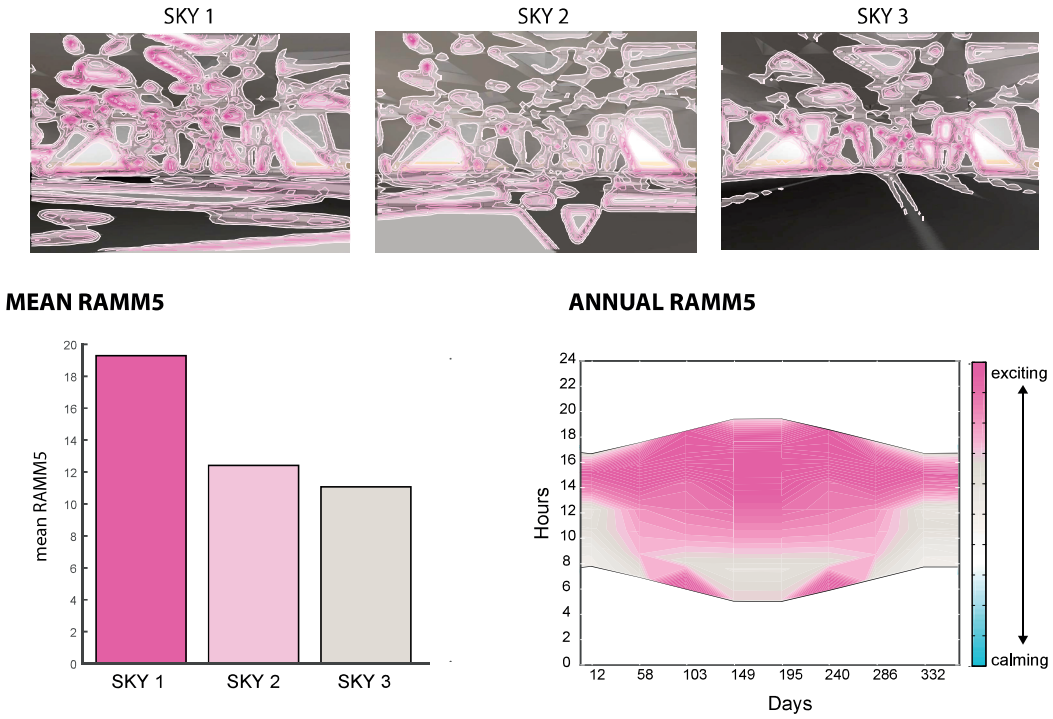


Figure A2.6 RAMM5 (mSC) results for the Serpentine Pavilion by Toyo Ito

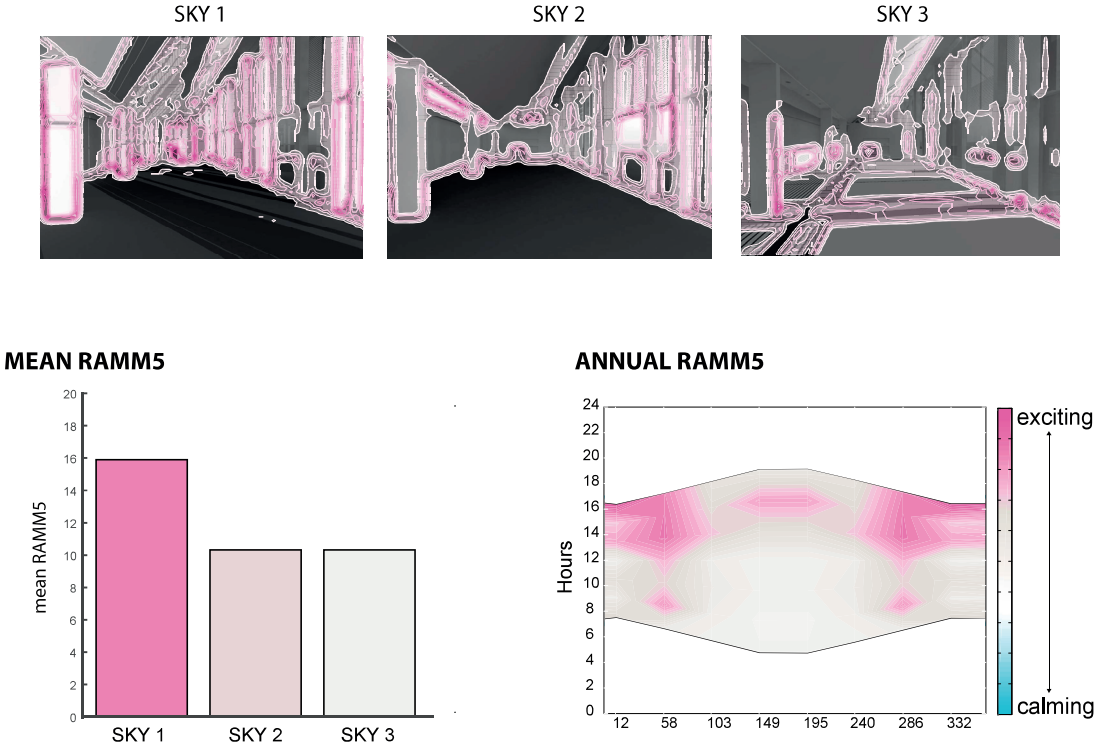


Figure A2.7 RAMM5 (mSC) results for the Neughebauer House by Richard Meier

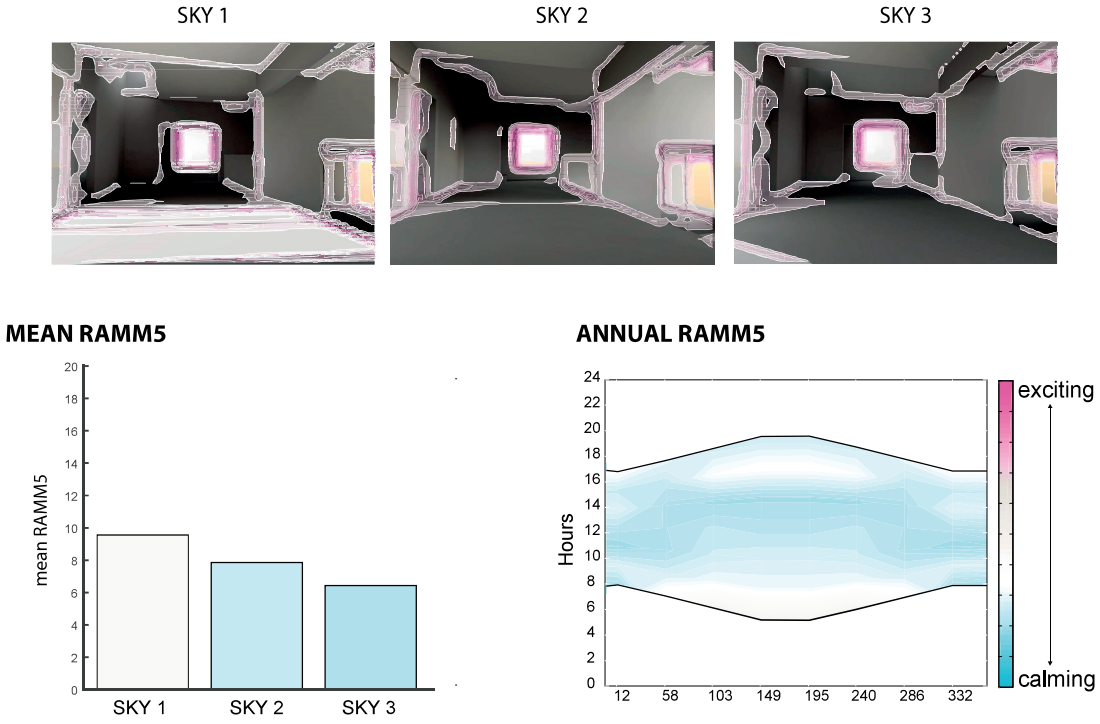


Figure A2.8 RAMM5 (mSC) results for the Poli House by Pezo von Ellrichshausen

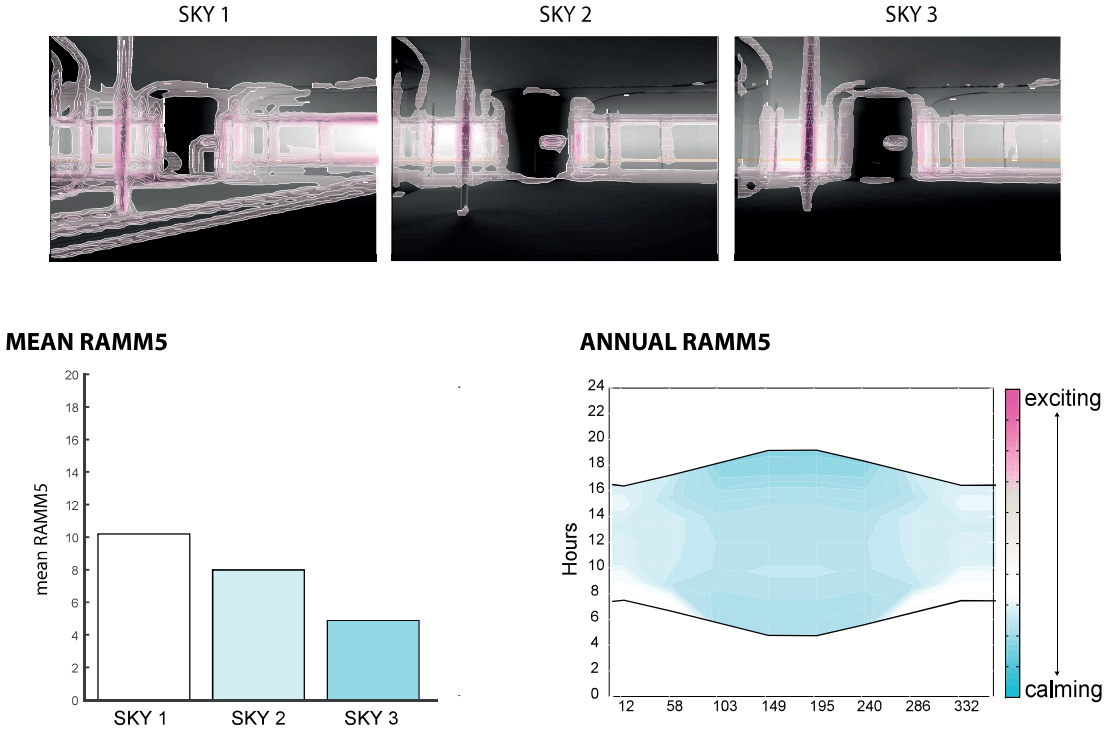


Figure A2.9 RAMM5 (mSC) results for the Toledo Glass Museum by SANAA

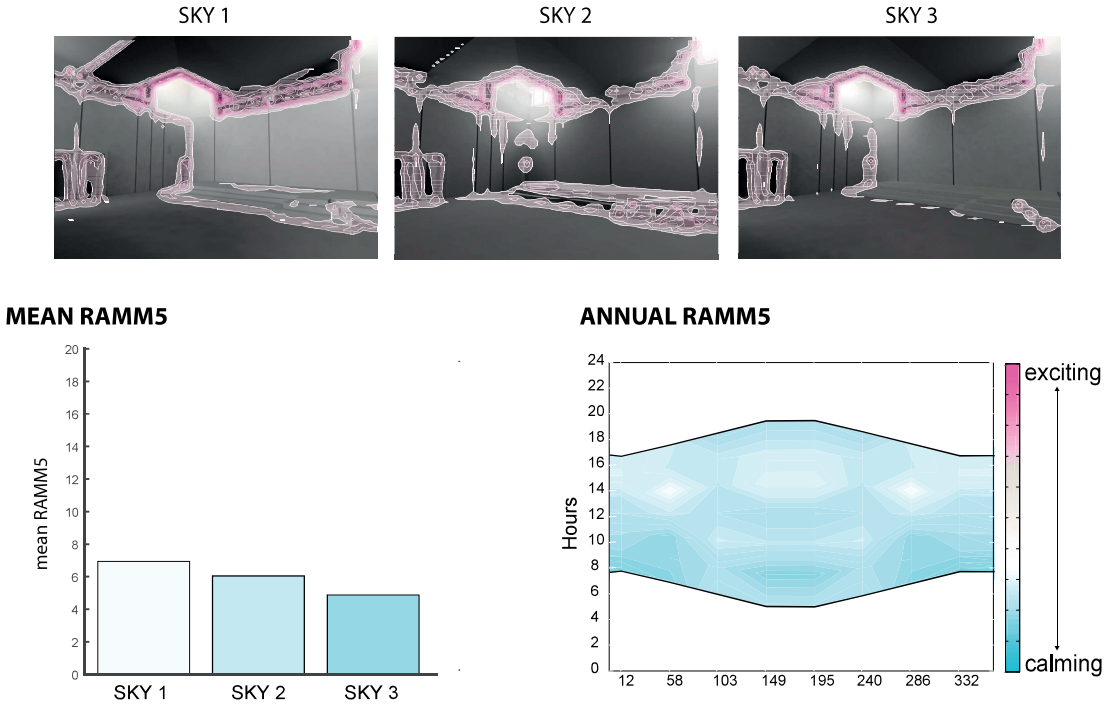


Figure A2.10 RAMM5 (mSC) results for the First Unitarian Church by Louis Kahn

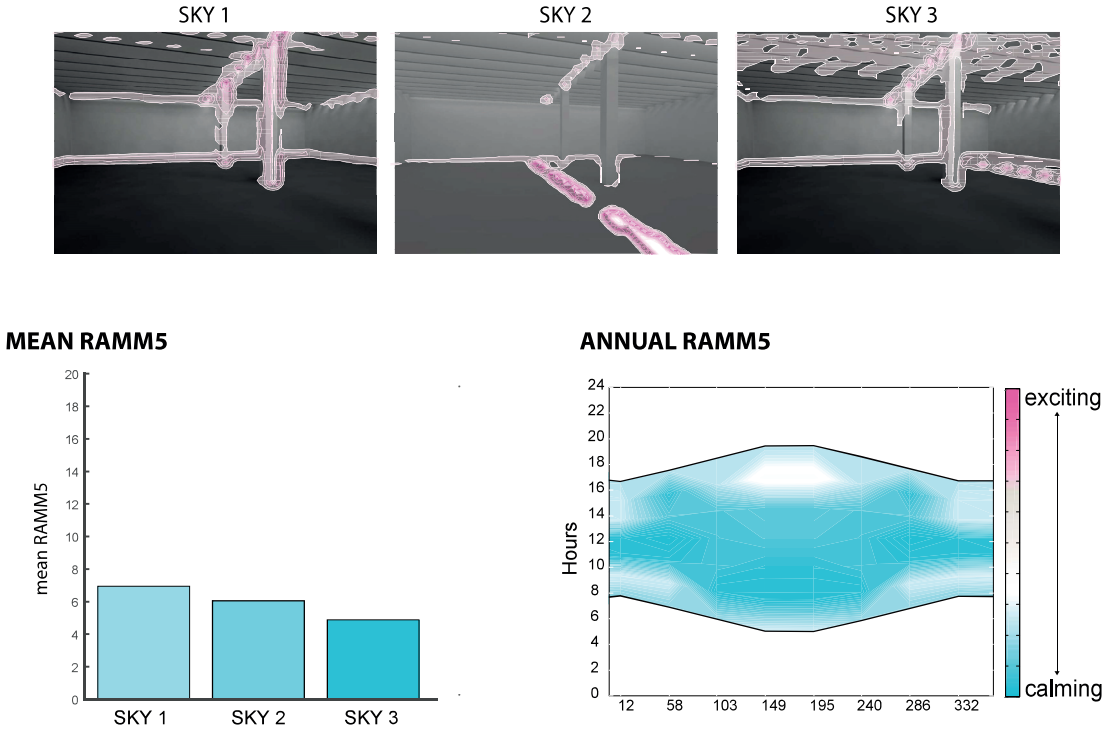


Figure A2.11 RAMM5 (mSC) results for the Menil Gallery by Renzo Piano

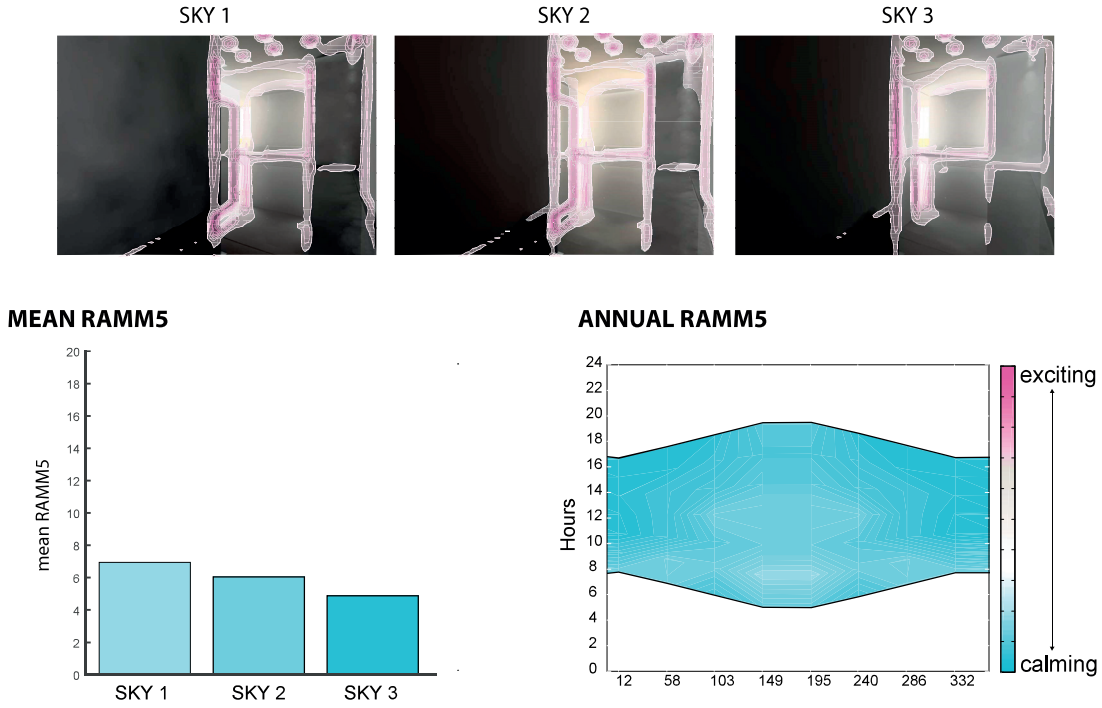


Figure A2.12 RAMM5 (mSC) results for the Vals Thermal Baths by Peter Zumthor

A3 APPENDIX: CHAPTER 6

Script A3.1 Matlab implementation of mSC hemisphere algorithm

run_multistep_SC_on_a_hemisphere.m

```
% Implemented 19.04.2016, MLA
% Developed by SR

% PYRAMID
% Compute Spatial Contrast on a hemisphere using a multistep method,
% where the image size is halved at each step.

% Needs: modified_SC_on_a_hemisphere.m

function [SC_step5_mean] = run_multistep_SC_on_a_hemisphere(dir_name)

% INPUT ---------------------------------------------------------------

%   Example:
%   dir_name = '~/Documents/BAE-cut/05_30/clear/cut-zollverein_8_1.hdr';

% ---------------------------------------------------------------------

% Minimum resolution:
min_res = 60;

% Read in original image:
IMG = hdrread(dir_name);

IMG_step1 = flipud(double(rgb2gray(IMG))).*255;
[N1,M1] = size(IMG_step1);
clear IMG;

% Set black pixels to nan:
IMG_step1(IMG_step1==0) = nan;

% Half the image size until min_res:
step_count = 1;
N_half = ceil(N1/2); %height
M_half = ceil(M1/2); %width
while (N_half > min_res)
   step_count = step_count + 1;
   eval(['IMG_step' num2str(step_count) ' = nan(N_half,M_half);'])
%     eval(['IMG_step' num2str(step_count) ' = imresize(IMG_step' num2str(step_count-1) ',
0.5,''bilinear'');'])
   eval(['IMG_step' num2str(step_count) ' = imresize(IMG_step' num2str(step_count-1) ', 0.5);'])

   % Reduce image size for next step:
   N_half = ceil(N_half/2);
   M_half = ceil(M_half/2);
end

% Compute spatial contrast:
[~,SC_step5_mean] = modified_SC_on_a_hemisphere(IMG_step5);
```

modified_SC_on_a_hemisphere.m

```
% Implemented 19.04.2016, MLA
% Works on round images.
% MODIFIED SPATIAL CONTRAST
% Developed by Siobhan Rockcastle

function [SC, mean_SC, sum_SC] = modified_SC_on_a_hemisphere(IMG)

[n,m] = size(IMG);

SC = nan(n,m);

SC(2:end-1,2:end-1) = ...
    ( (abs(IMG(2:end-1,2:end-1) - IMG(1:end-2,2:end-1)) + ...
      abs(IMG(2:end-1,2:end-1) - IMG(3:end,2:end-1)) + ...
      abs(IMG(2:end-1,2:end-1) - IMG(2:end-1,1:end-2)) + ...
      abs(IMG(2:end-1,2:end-1) - IMG(2:end-1,3:end))) + ...
      ( (1/sqrt(2)) .* (...
      abs(IMG(2:end-1,2:end-1)-IMG(1:end-2,1:end-2)) + ...
      abs(IMG(2:end-1,2:end-1)-IMG(1:end-2,3:end)) + ...
      abs(IMG(2:end-1,2:end-1)-IMG(3:end,3:end)) + ...
      abs(IMG(2:end-1,2:end-1)-IMG(3:end,1:end-2)) ...
      ) ) )./ (4+2*sqrt(2));

SC = round(SC);

% Count non-nan's:
Np = sum(sum(~isnan(SC)));

sum_SC = squeeze(sum(sum(SC(~isnan(SC)))));
mean_SC = sum_SC/Np;
```

A4 APPENDIX: CHAPTER 8

Script A4.1 Matlab implementation of mSC box algorithm

Multistep_spatial_contrast_on_a_box.m

```
% Implemented 15.04.2016, MLA
% Developed by SR

% Modified 19.04.2016 MLA
% flipud added for reading in .bmp images
% try hdrread and normalize

% PYRAMID
% Compute Spatial Contrast on a box using a multistep method,
% where the image size is halved at each step.

% Needs: modified_SC_on_a_box.m and run_modified_SC_on_a_box.m

% load custom colormap
load excitement2


% INPUTS
%-----------------------------------------------
% We assume that the images are numbered 1-6 and are of equal size.

% Directory:
% dir_name = 'Name/img000';
dir_name = 'Name/month_day/name_hour_';

% Minimum resolution:
min_res = 60;
%-----------------------------------------------

% Get original image size:
% eval(['IMG_step1 = imread(['' dir_name num2str(1) '.bmp'']);'])
eval(['IMG_step1 = hdrread(['' dir_name num2str(1) '.hdr'']);'])
IMG_step1 = double(rgb2gray(IMG_step1))+1;
[N1,M1] = size(IMG_step1);

% Read in original images:
Nimg = 6;
IMG_step1 = nan(N1,M1,Nimg);
for img=1:Nimg
% eval(['IMG = imread(['' dir_name num2str(img) '.bmp'']);'])
   eval(['IMG = hdrread(['' dir_name num2str(img) '.hdr'']);'])
   IMG_step1(:,:,img) = (flipud(double(rgb2gray(IMG))+1)-1).*255;
   clear IMG
end

% Half the image size until min_res:
step_count = 1;
N_half = N1/2;
while (N_half > min_res)
```

```
    step_count = step_count + 1;
    eval(['IMG_step' num2str(step_count) ' = nan(N_half,N_half,Nimg);'])
    for img=1:Nimg
        eval(['IMG_step' num2str(step_count) ...
            '(:,:,img) = imresize(IMG_step' num2str(step_count-1) '(:,:,img), 0.5);'])
    end
    N_half = N_half/2;
end

% Compute spatial contrast:
for istep=1:step_count
    eval(['SC_step' num2str(istep) ' = run_modified_SC_on_a_box(IMG_step' num2str(istep) ');'])
end

%%
SC_mean_step5= mean(mean(mean(SC_step5)));
```

run_modified_SC_on_a_box.m

```
% Implemented 15.04.2016, MLA
% Developed by SR
% Compute SC on images from an unrolled box.
% Needs: modified_SC_on_a_box.m

function [SC, IMG_ext_mtx] = run_modified_SC_on_a_box(IMG)

[N,M,Nimg] = size(IMG);

% From a viewpoint:
%             1      2      3      4      5      6
% img_label = {   'Right' 'Left'  'Back'  'Front' 'Top'   'Ground'};

% Clockwise from top:
%       a
%   d  IMG  b
%       c
edges_label = {'a' 'b' 'c' 'd'};        %IMG
sur_img = [     5   3   6   4; ...      % 1
           5   4   6   3; ...           % 2
           5   2   6   1; ...           % 3
           5   1   6   2; ...           % 4
           4   2   3   1; ...           % 5
           4   1   3   2];              % 6
%
%            a   b   c   d              %IMG
sur_edges = [   4   4   2   2; ...      % 1
           2   4   4   2; ...           % 2
           3   4   3   2; ...           % 3
           1   4   1   2; ...           % 4
           1   1   1   1; ...           % 5
           3   3   3   3];              % 6

% Memory:
IMG_ext_mtx = nan(N+2,M+2,Nimg);
SC = nan(N,M,Nimg);

% Extend images:
for img=1:Nimg

    % Increase size:
    IMG_ext = nan(N+2,M+2);
```

```
    IMG_ext(2:end-1,2:end-1) = IMG(:,:,img);

    % Find edges:
     for edges=1:4

        eval(['IMG' edges_label{edges} ' = IMG(:,:,' num2str(sur_img(img,edges)) ');'])

        if sur_edges(img,edges)==1
            eval([edges_label{edges} ' = IMG' edges_label{edges} '(end,:);'])
        elseif sur_edges(img,edges)==2
            eval([edges_label{edges} ' = IMG' edges_label{edges} '(:,end);'])
        elseif sur_edges(img,edges)==3
           eval([edges_label{edges} ' = IMG' edges_label{edges} '(1,:);'])
        elseif sur_edges(img,edges)==4
           eval([edges_label{edges} ' = IMG' edges_label{edges} '(:,1);'])
        end
        clear IMGa IMGb IMGc IMGd
    end

    if img == 1
        a = fliplr(a');
        c = fliplr(c');
    end

     if img == 2
        a = fliplr(a);
    end

    if img == 3
       c = fliplr(c);
    end

    if img == 4
       a = fliplr(a);
    end

    if img == 5
       a = fliplr(a);
       b = fliplr(b');
       d = fliplr(d)';
    end

    if img == 6
       b = fliplr(b)';
       c = fliplr(c);
       d = fliplr(d');
    end

    % Stitch image:
    IMG_ext(end,2:end-1) = a;
    IMG_ext(2:end-1,end) = b;
    IMG_ext(1,2:end-1) = c;
    IMG_ext(2:end-1,1) = d;

    clear a b c d

    % Calculate Spatial Contrast:
    SC(:,:,img) = modified_SC_on_a_box(IMG_ext);

    IMG_ext_mtx(:,:,img) = IMG_ext;
    clear IMG_ext
end
```

modified_SC_on_a_box.m

```
% Implemented 15.04.2016, MLA
% Works on extended images (n+2,m+2) with nan corners from a box.
% MODIFIED SPATIAL CONTRAST
% Developed by Siobhan Rockcastle

function [SC, mean_SC, sum_SC] = modified_SC_on_a_box(IMG)

SC = ( (abs(IMG(2:end-1,2:end-1) - IMG(1:end-2,2:end-1)) + ...
        abs(IMG(2:end-1,2:end-1) - IMG(3:end,2:end-1)) + ...
        abs(IMG(2:end-1,2:end-1) - IMG(2:end-1,1:end-2)) + ...
        abs(IMG(2:end-1,2:end-1) - IMG(2:end-1,3:end))) + ...
        ( (1/sqrt(2)) .* (...
        abs(IMG(2:end-1,2:end-1)-IMG(1:end-2,1:end-2)) + ...
        abs(IMG(2:end-1,2:end-1)-IMG(1:end-2,3:end)) + ...
        abs(IMG(2:end-1,2:end-1)-IMG(3:end,3:end)) + ...
        abs(IMG(2:end-1,2:end-1)-IMG(3:end,1:end-2)) ...
        ) ) )./ (4+2*sqrt(2));


% Special cases:
% Corner (1,1)
SC(1,1) = ( (abs(IMG(2,2) - IMG(1,2)) + ...
        abs(IMG(2,2) - IMG(3,2)) + ...
        abs(IMG(2,2) - IMG(2,1)) + ...
        abs(IMG(2,2) - IMG(2,3))) + ...
        ( (1/sqrt(2)) .* (...
        abs(IMG(2,2)-IMG(1,3)) + ...
        abs(IMG(2,2)-IMG(3,3)) + ...
        abs(IMG(2,2)-IMG(3,1)) ...
        ) ) )./ (4+3*sqrt(2)/2);


% Corner (1,end)
SC(1,end) = ( (abs(IMG(2,end-1) - IMG(1,end-1)) + ...
        abs(IMG(2,end-1) - IMG(3,end-1)) + ...
        abs(IMG(2,end-1) - IMG(2,end-2)) + ...
        abs(IMG(2,end-1) - IMG(2,end))) + ...
        ( (1/sqrt(2)) .* (...
        abs(IMG(2,end-1)-IMG(1,end-2)) + ...
        abs(IMG(2,end-1)-IMG(3,end)) + ...
        abs(IMG(2,end-1)-IMG(3,end-2)) ...
        ) ) )./ (4+3*sqrt(2)/2);


% Corner (end,1)
SC(end,1) = ( (abs(IMG(end-1,2) - IMG(end-2,2)) + ...
        abs(IMG(end-1,2) - IMG(end,2)) + ...
        abs(IMG(end-1,2) - IMG(end-1,1)) + ...
        abs(IMG(end-1,2) - IMG(end-1,3))) + ...
        ( (1/sqrt(2)) .* (...
        abs(IMG(end-1,2)-IMG(end-2,1)) + ...
        abs(IMG(end-1,2)-IMG(end-2,3)) + ...
        abs(IMG(end-1,2)-IMG(end,3)) ...
        ) ) )./ (4+3*sqrt(2)/2);

SC(end,end) = ( (abs(IMG(end-1,end-1) - IMG(end-2,end-1)) + ...
        abs(IMG(end-1,end-1) - IMG(end,end-1)) + ...
        abs(IMG(end-1,end-1) - IMG(end-1,end-2)) + ...
        abs(IMG(end-1,end-1) - IMG(end-1,end))) + ...
        ( (1/sqrt(2)) .* (...
```

```
        abs(IMG(end-1,end-1)-IMG(end-2,end-2)) + ...
        abs(IMG(end-1,end-1)-IMG(end-2,end)) + ...
        abs(IMG(end-1,end-1)-IMG(end,end-2)) ...
        ) ) )./ (4+3*sqrt(2)/2);

SC = round(SC);

[n, m] = size(SC);
sum_SC = squeeze(sum(sum(SC)));
mean_SC = sum_SC/(n*m);
```

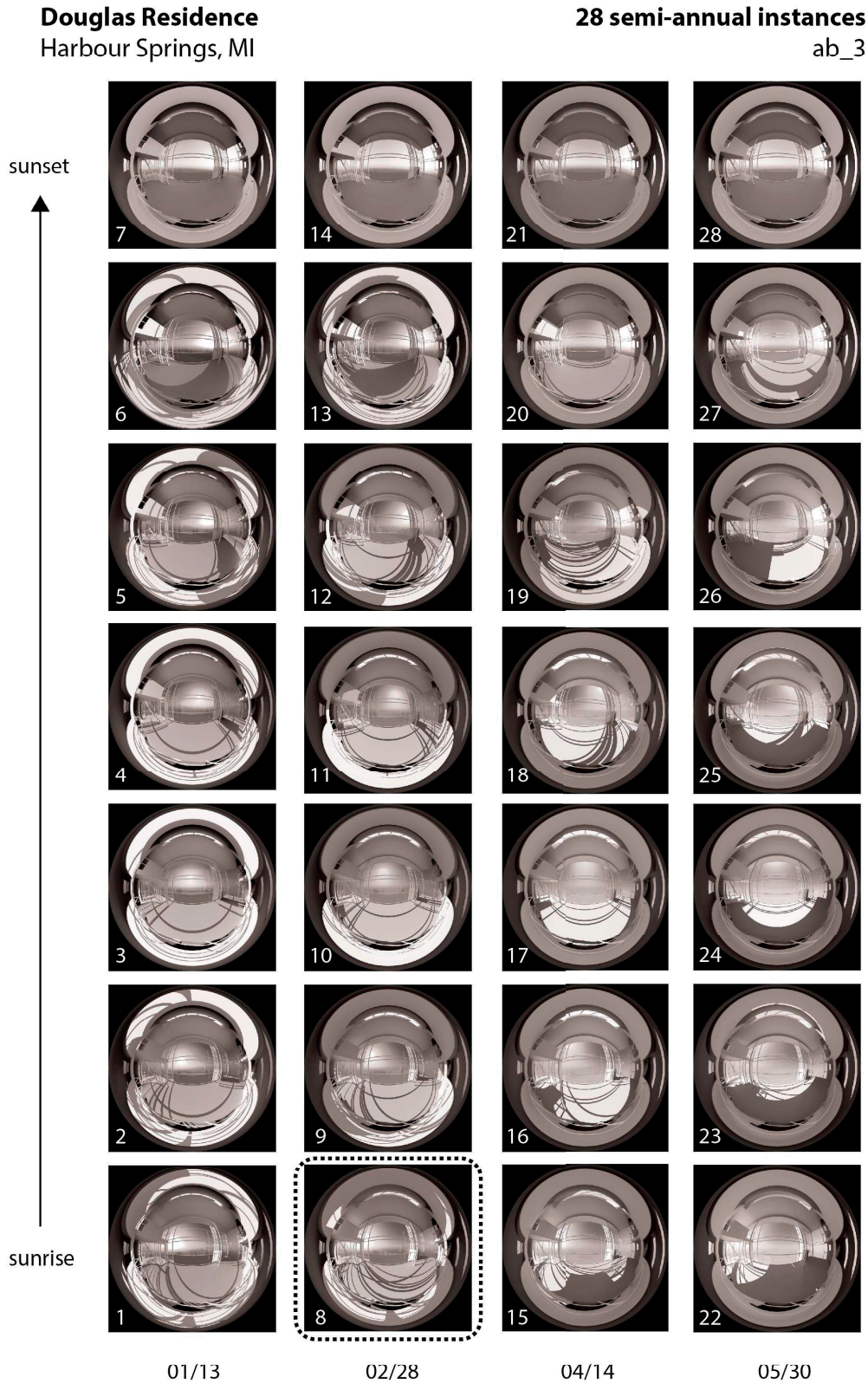


Figure A4.2 RAMM5 (mSC_5) annual results for the Douglas House

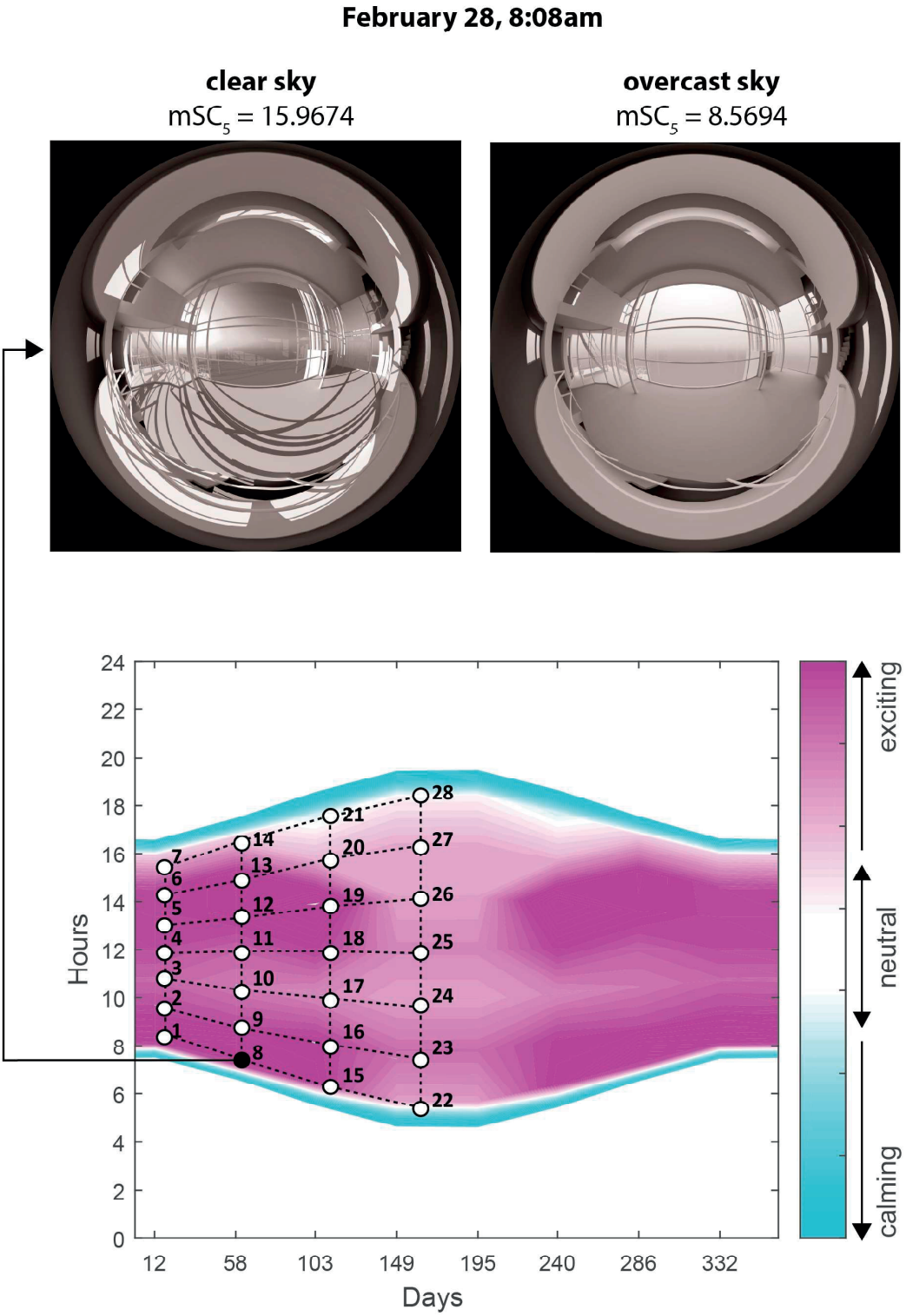


Figure A4.2 continued... High clear sky instance & corresponding overcast sky

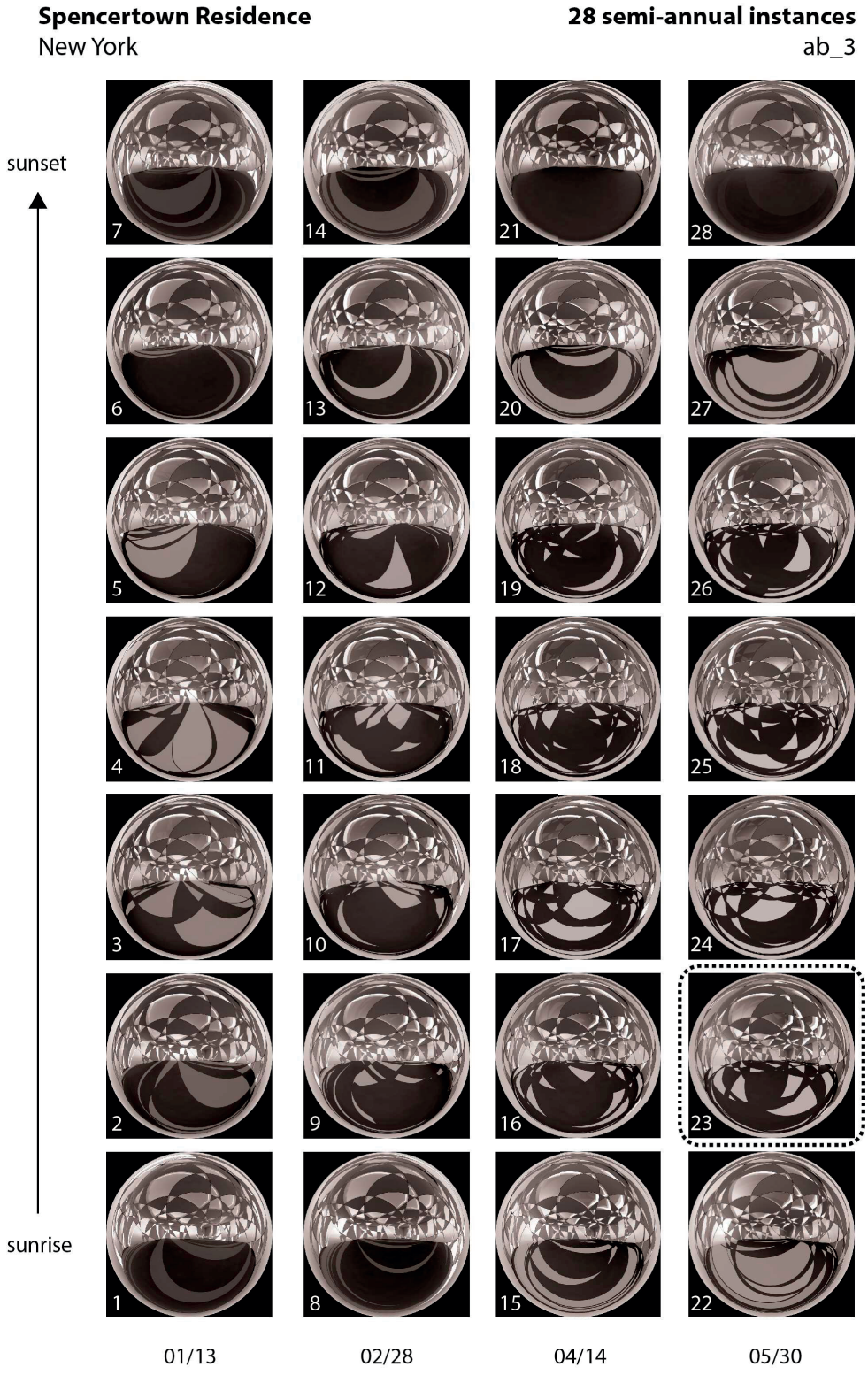


Figure A4.3 RAMM5 (mSC_5) annual results for the Serpentine Pavilion

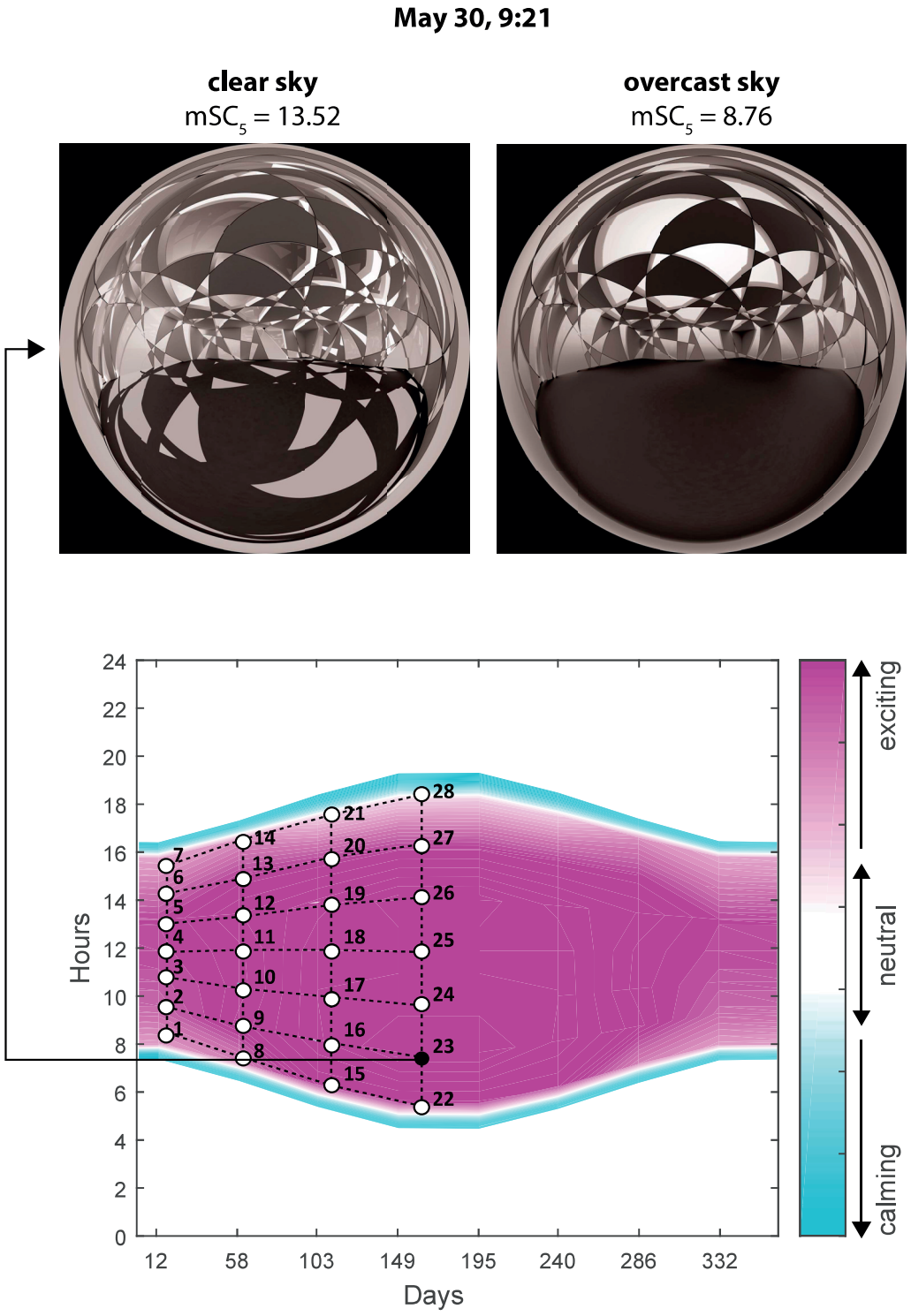


Figure A4.3 continued... High clear sky instance & corresponding overcast sky

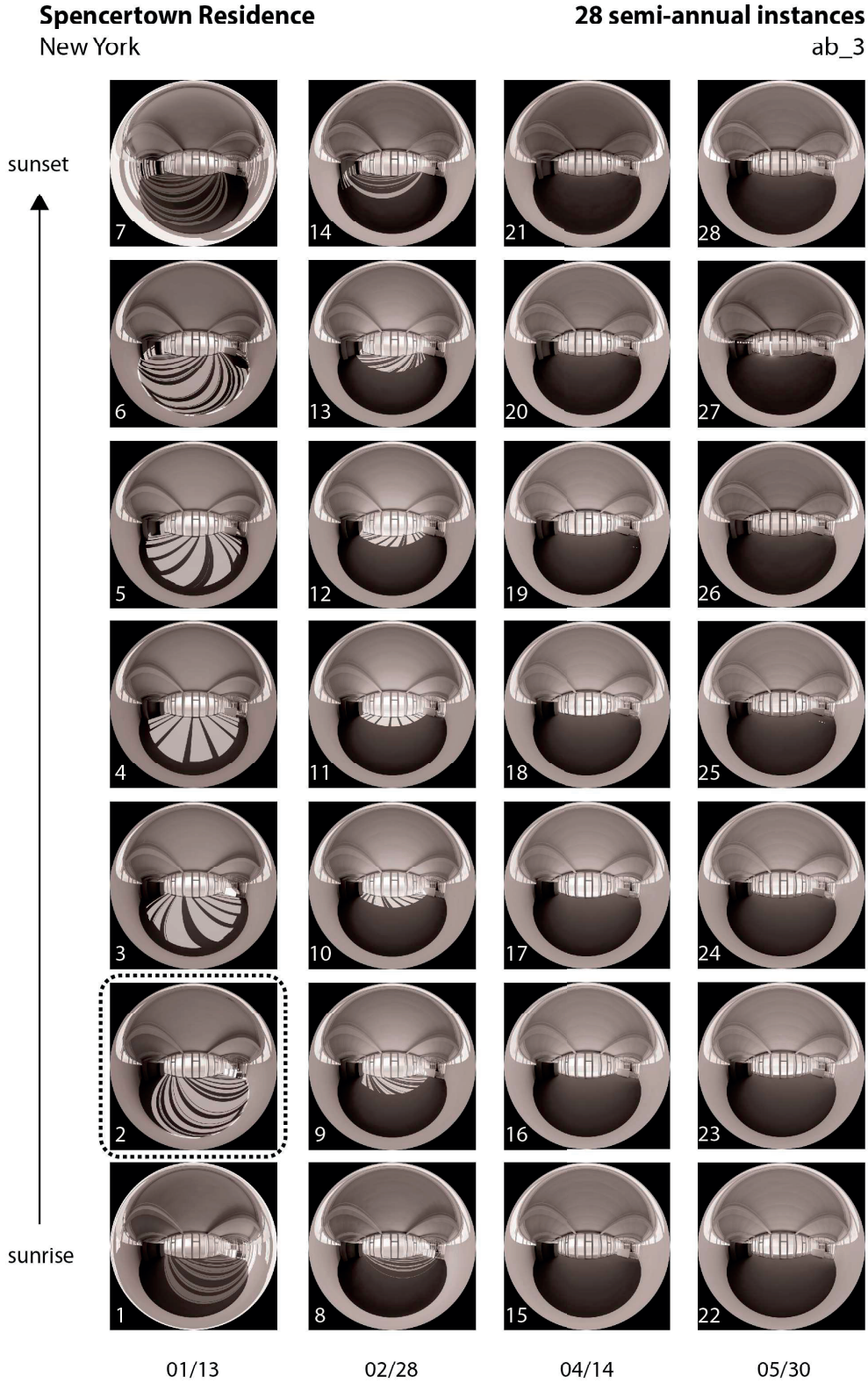


Figure A4.4 RAMM5 (mSC_5) annual results for the Spencertown House

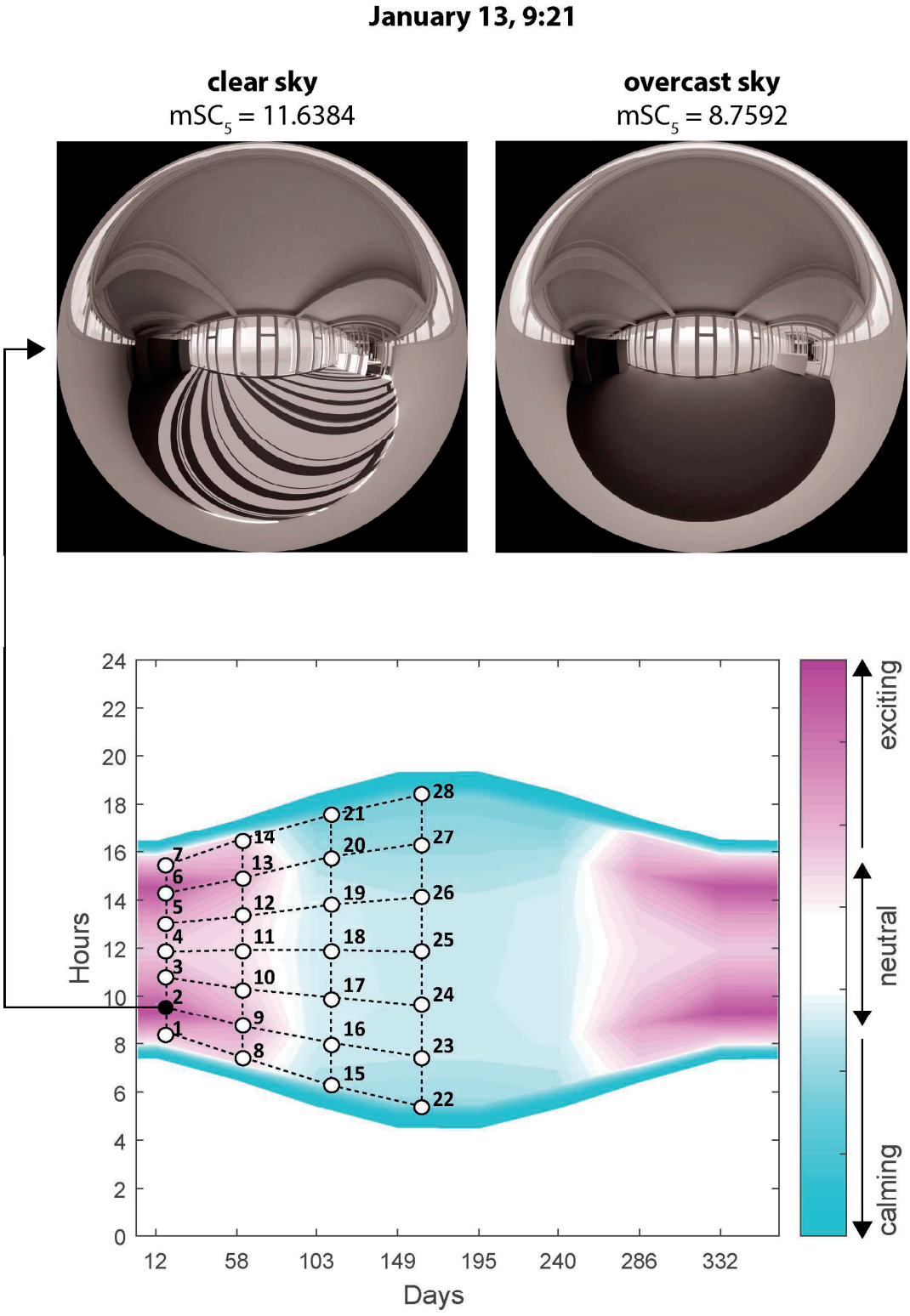


Figure A4.4 continued... High clear sky instance & corresponding overcast sky

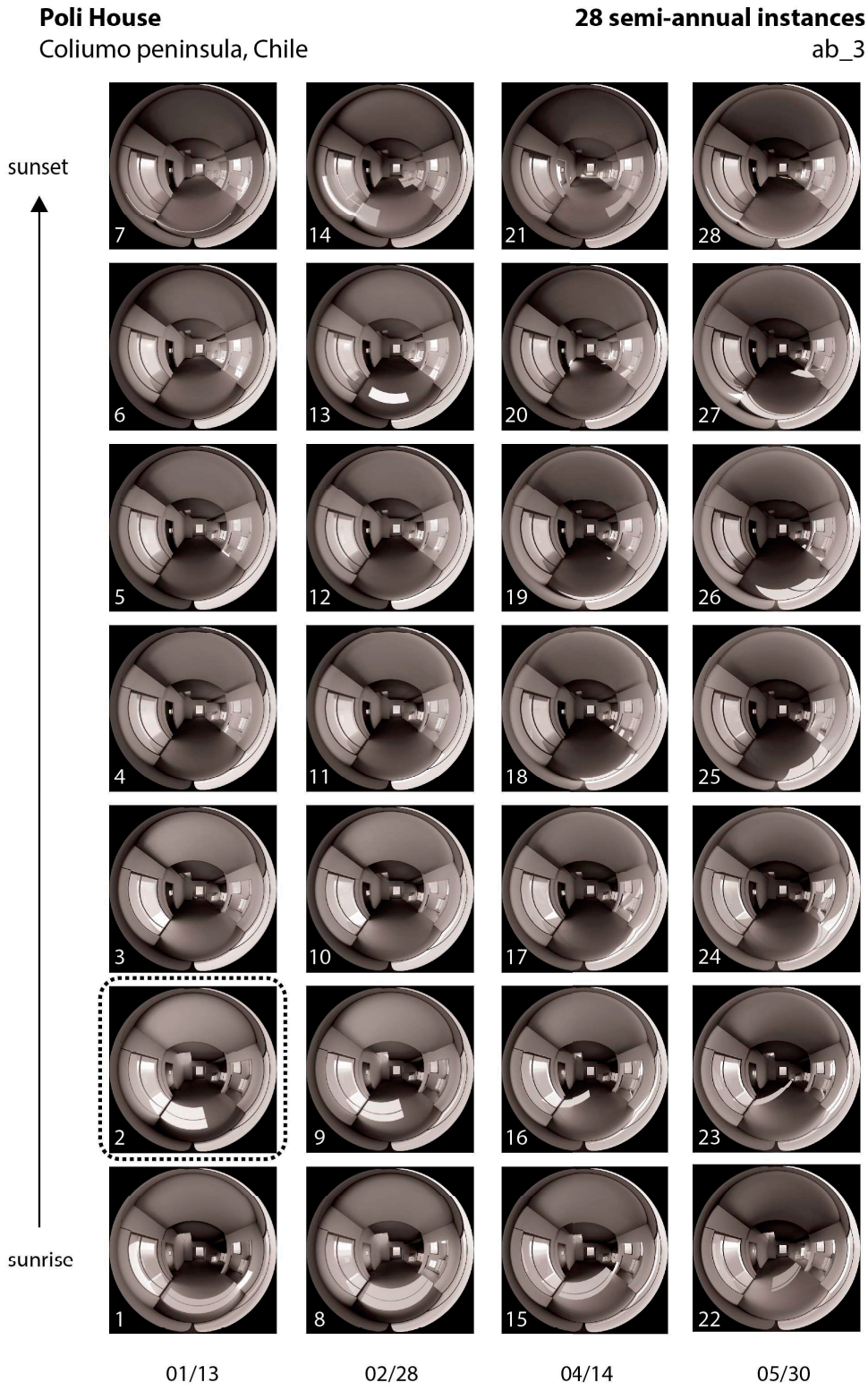


Figure A4.5 RAMM5 (mSC_5) annual results for the Poli House

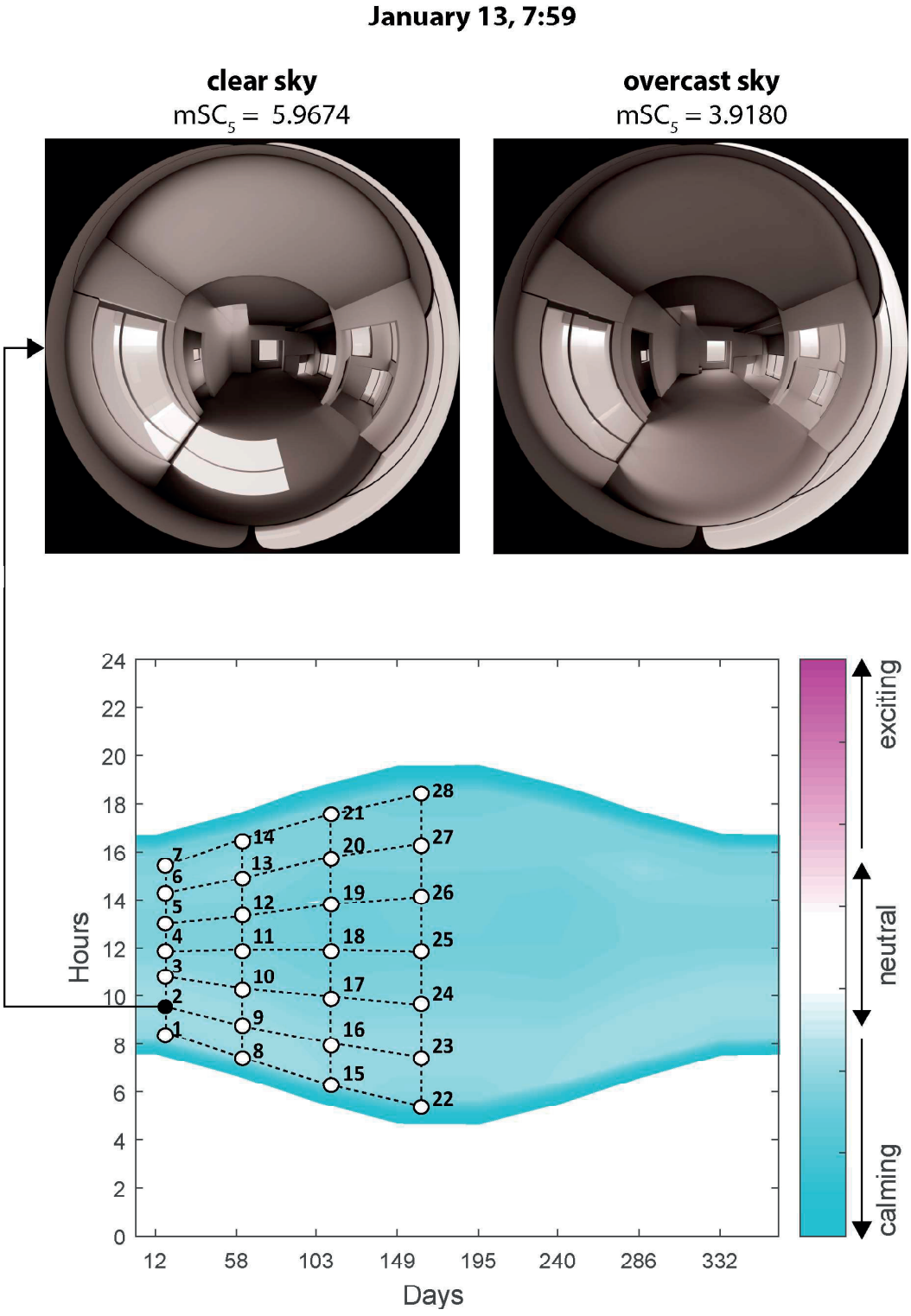


Figure A4.5 continued... High clear sky instance & corresponding overcast sky

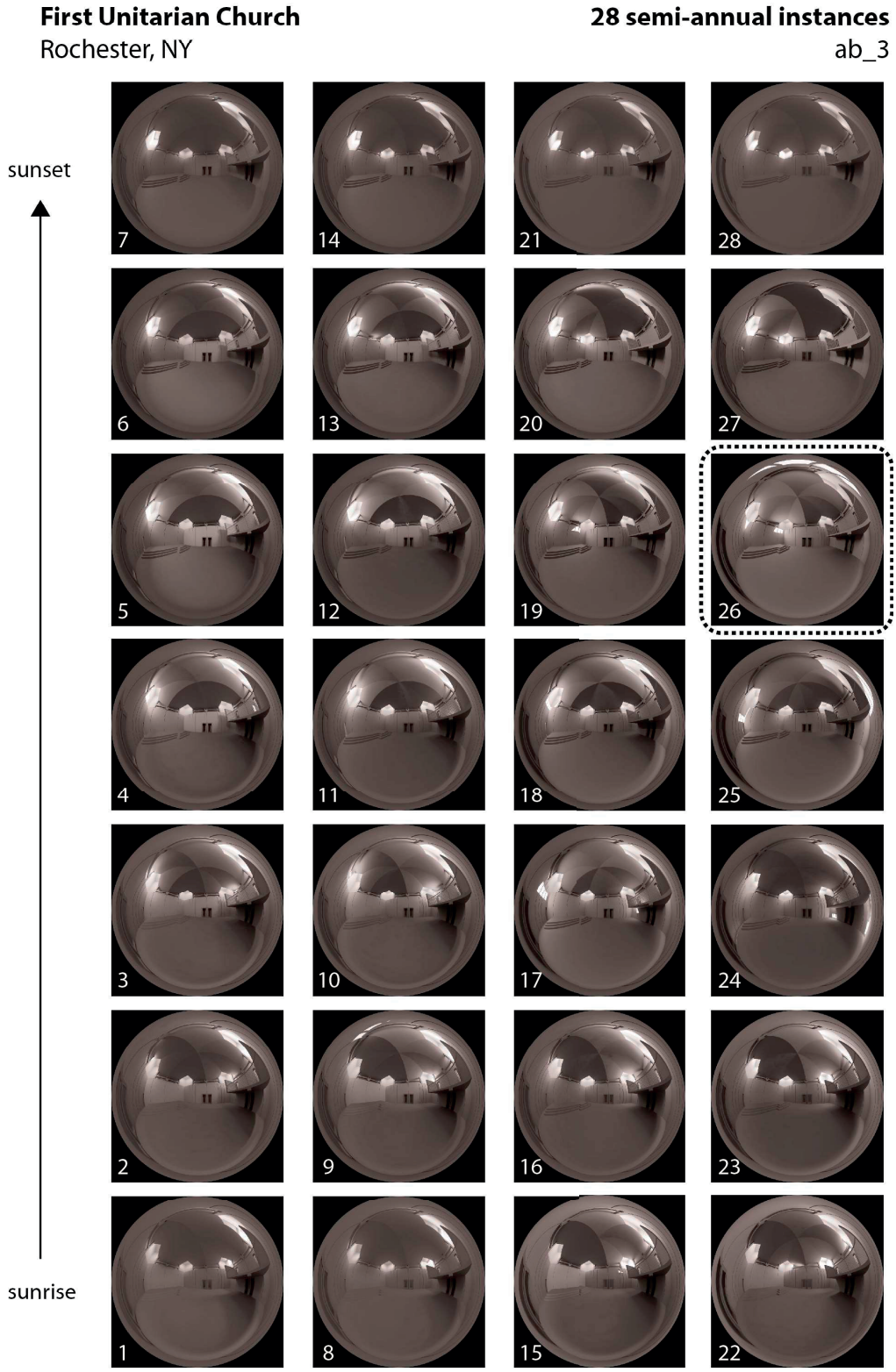


Figure A4.6 RAMM5 (mSC_5) annual results for the First Unitarian Church

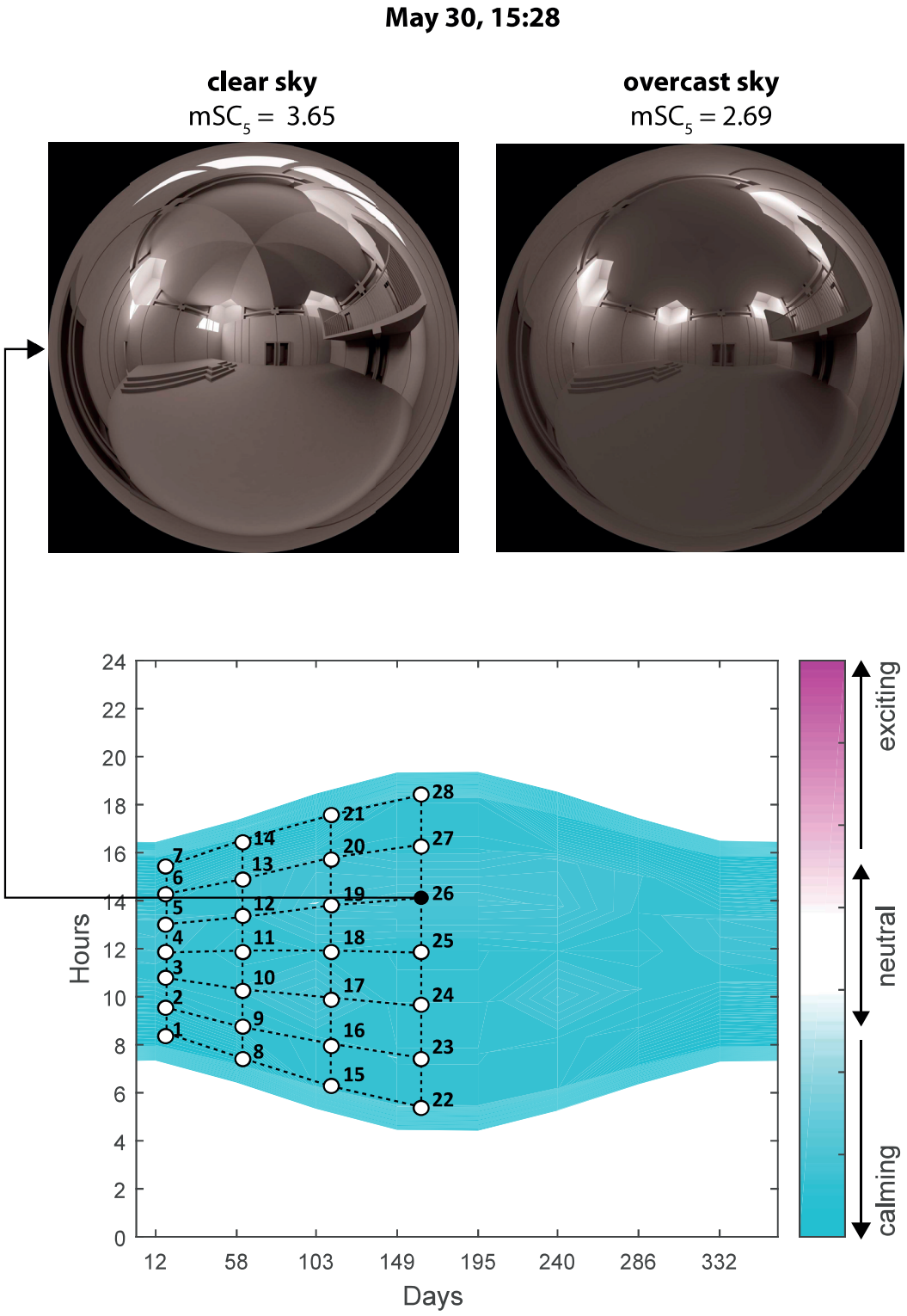


Figure A4.6 continued... High clear sky instance & corresponding overcast sky

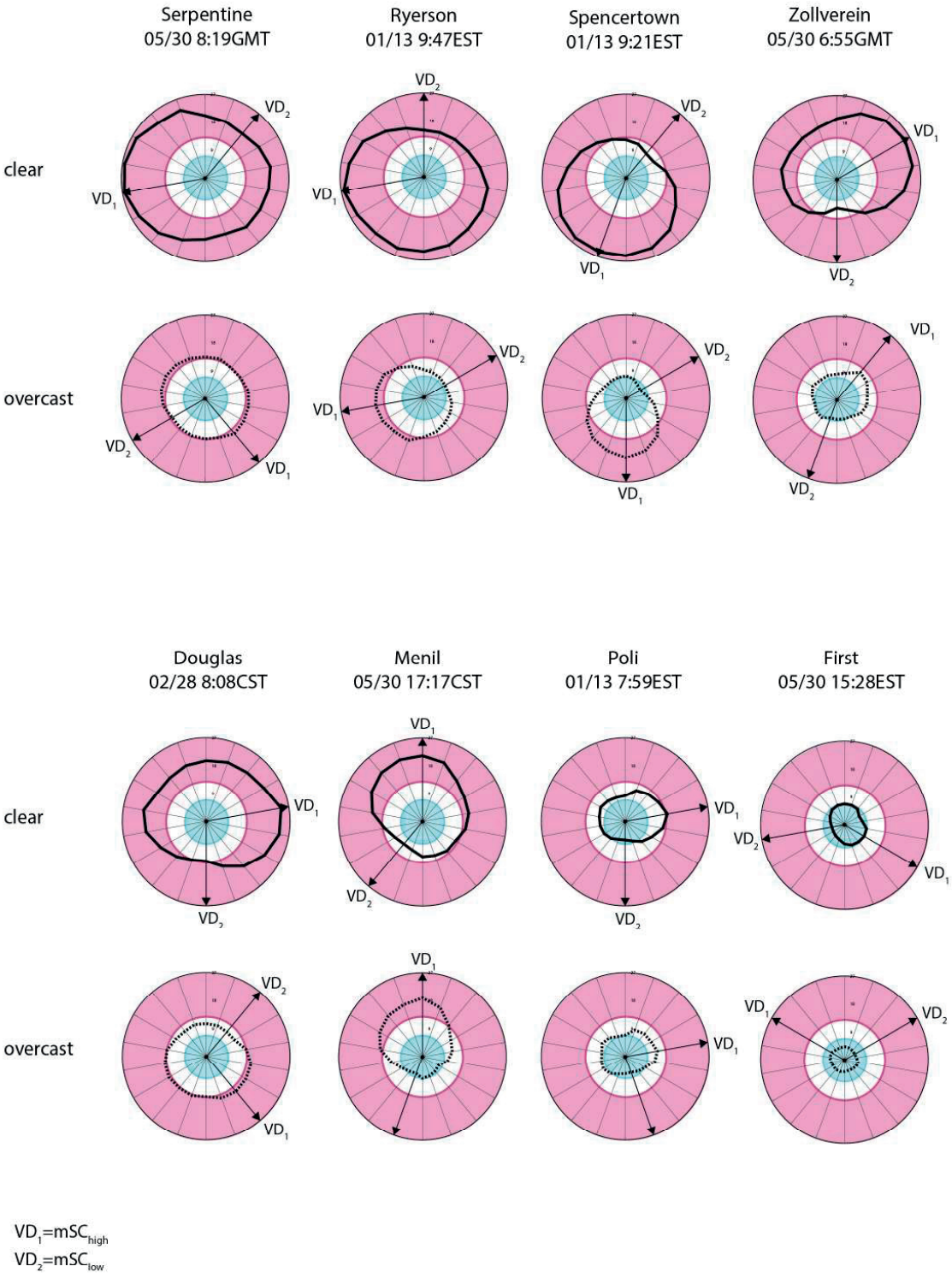


Figure A4.7 Selected view directions based on mSC_5 for each space, sky condition, and view direction

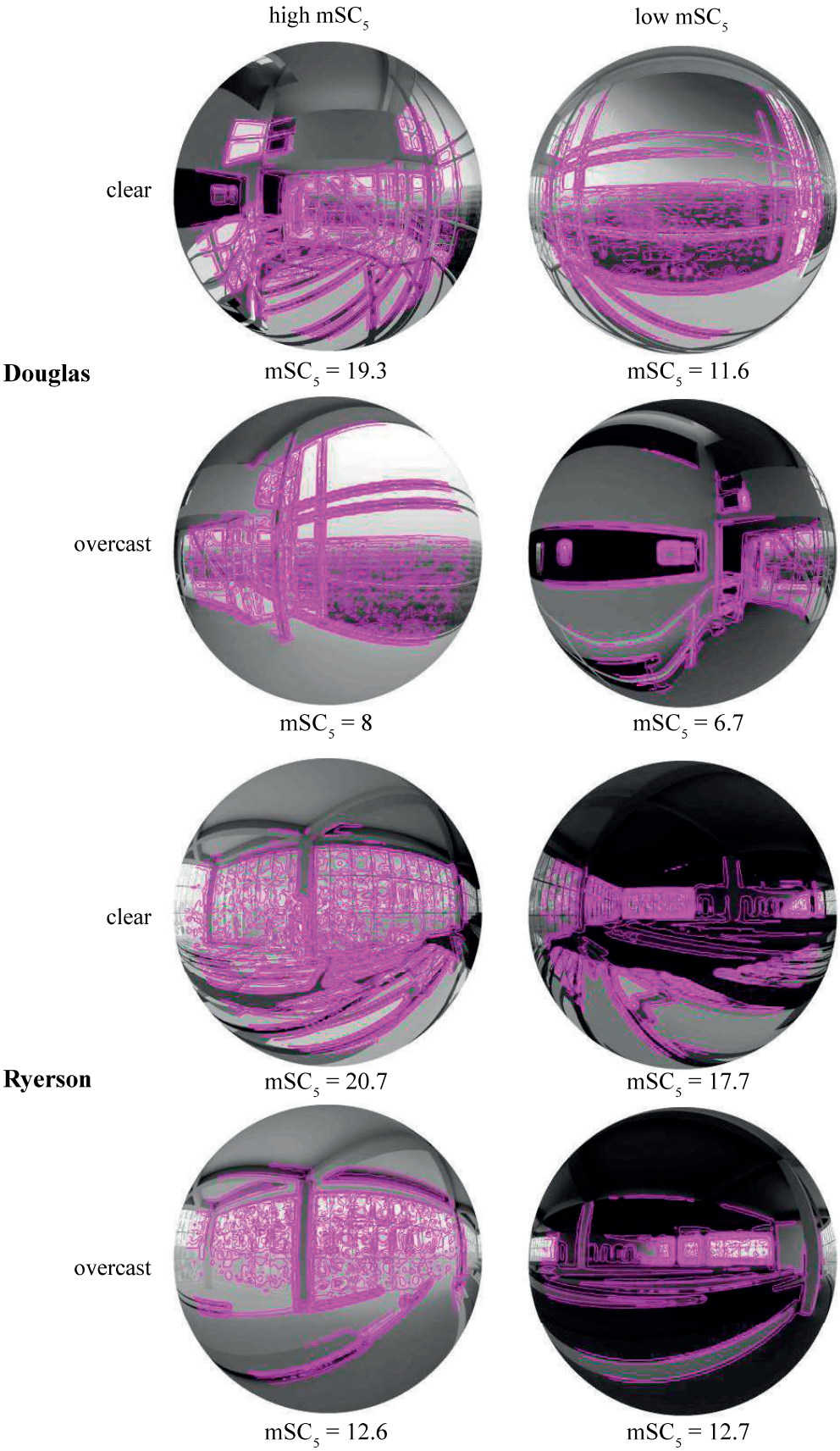


Figure A4.8 mSC_5 results for each space, sky, and view direction in the 180° group.

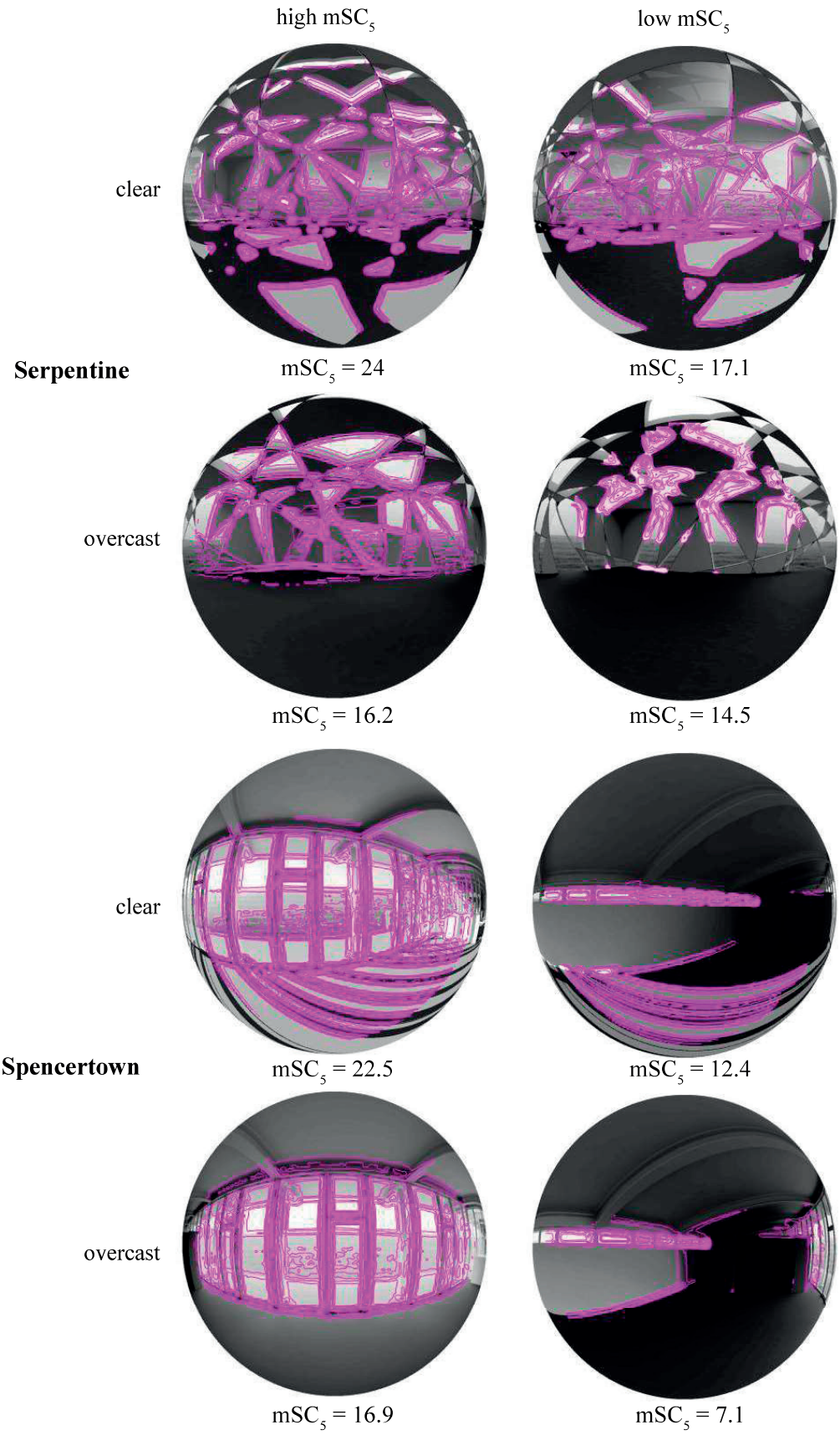


Figure A4.8 continued... mSC_5 results for each space, sky, and view direction in the 180° group.

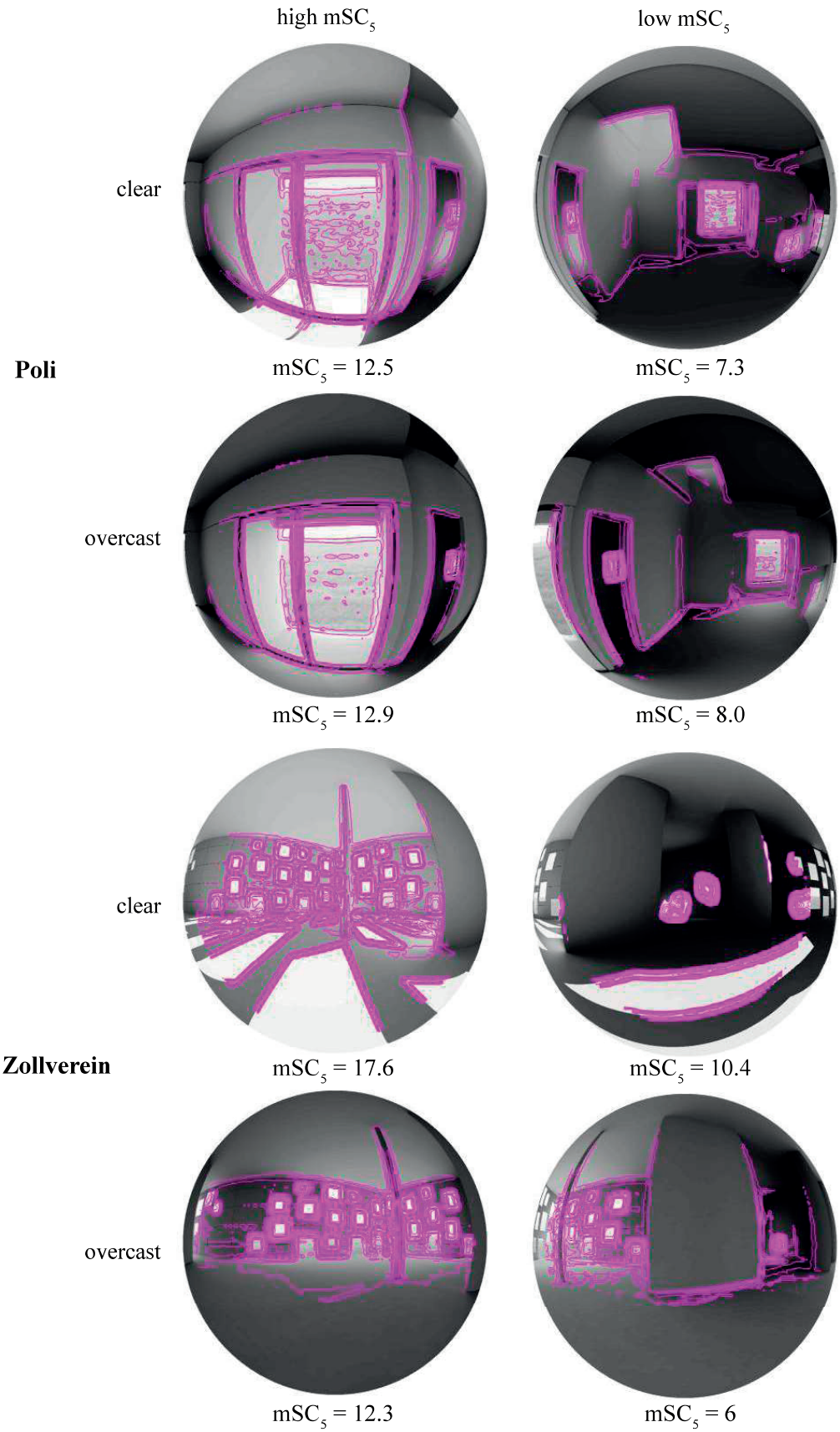


Figure A4.8 continued... mSC_5 results for each space, sky, and view direction in the 180° group.

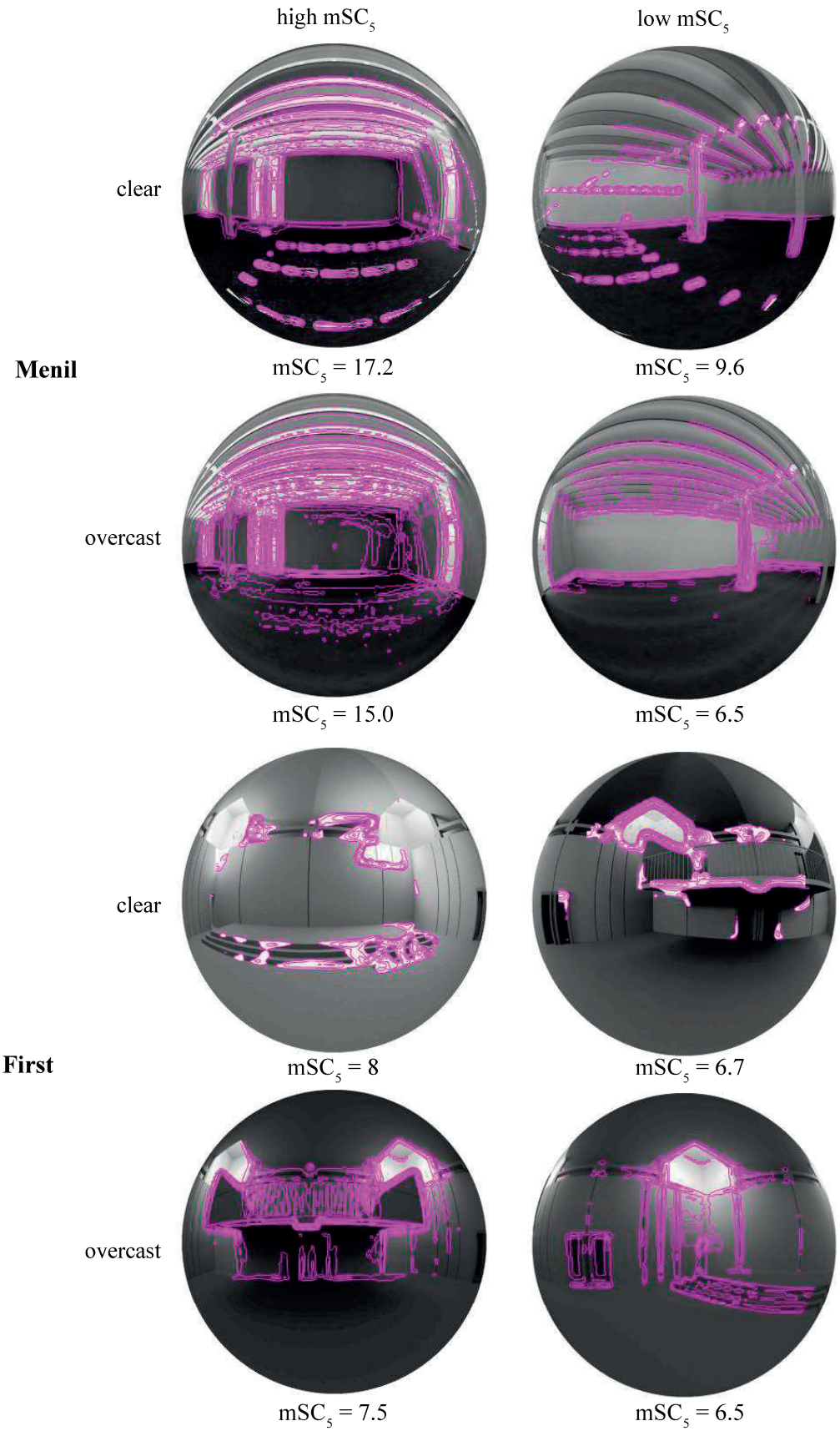


Figure A4.8 continued... mSC_5 results for each space, sky, and view direction in the 180° group.

a)

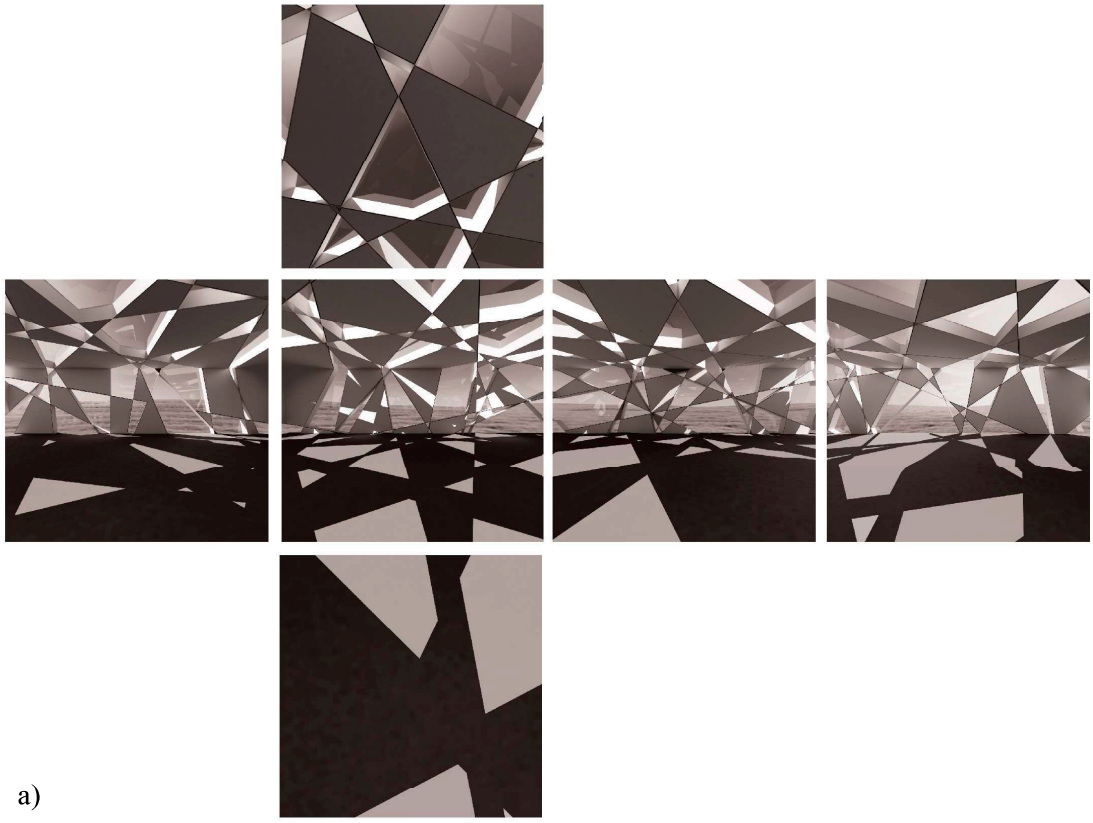
b)

Figure A4.9 Final renderings in cube map projection for Douglas a) clear sky & b) overcast

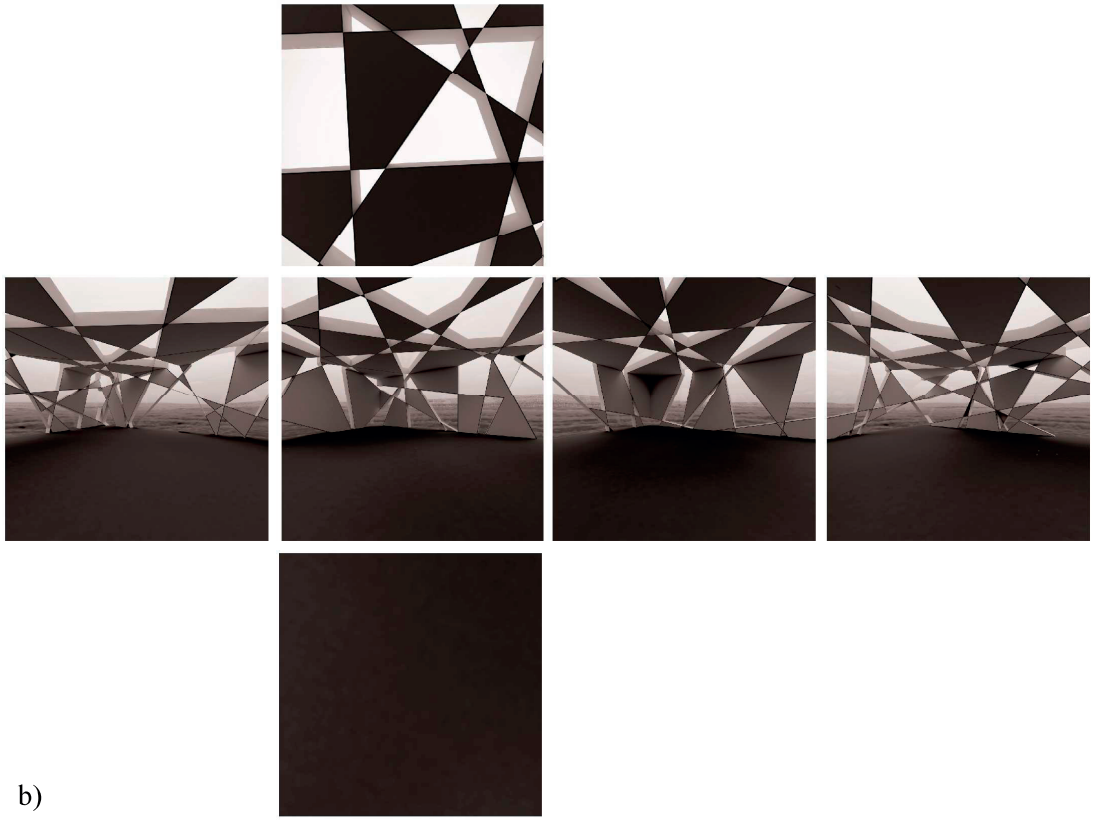
a)

b)

Figure A4.10 Final renderings in cube map projection for Ryerson a) clear sky & b) overcast



a)



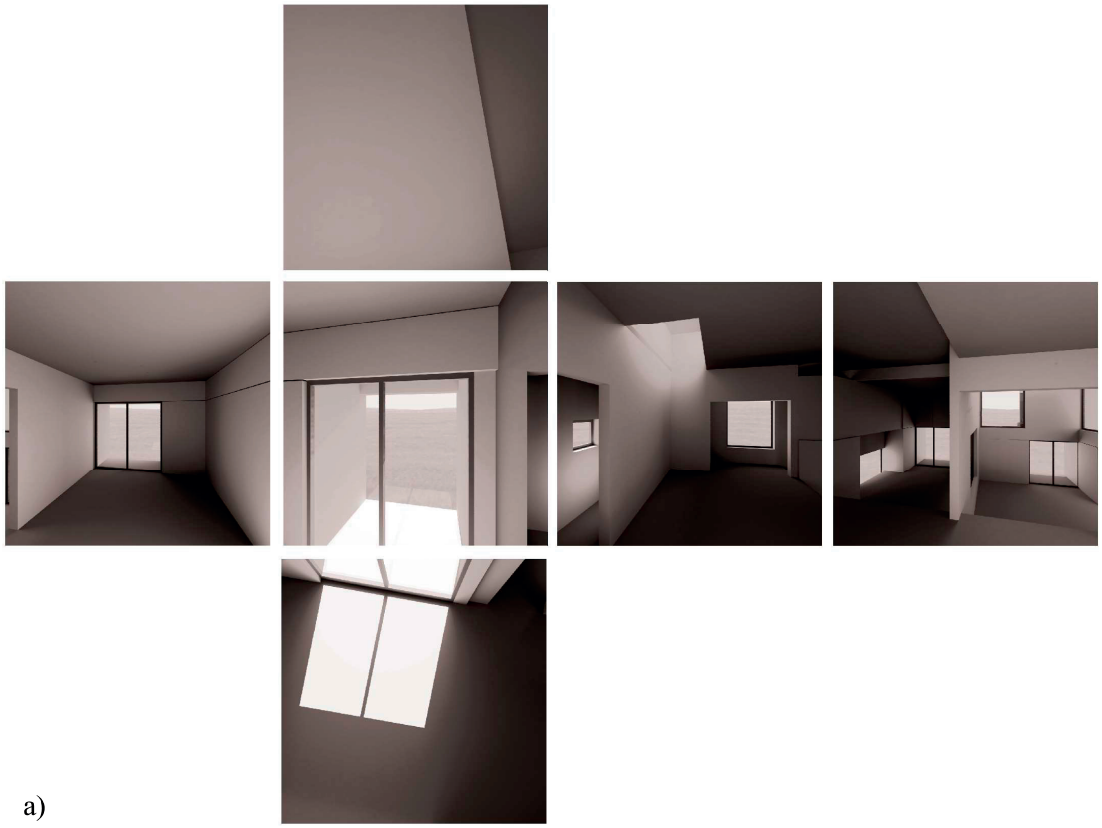
b)

Figure A4.11 Final renderings in cube map projection for Serpentine a) clear sky & b) overcast

a)

b)

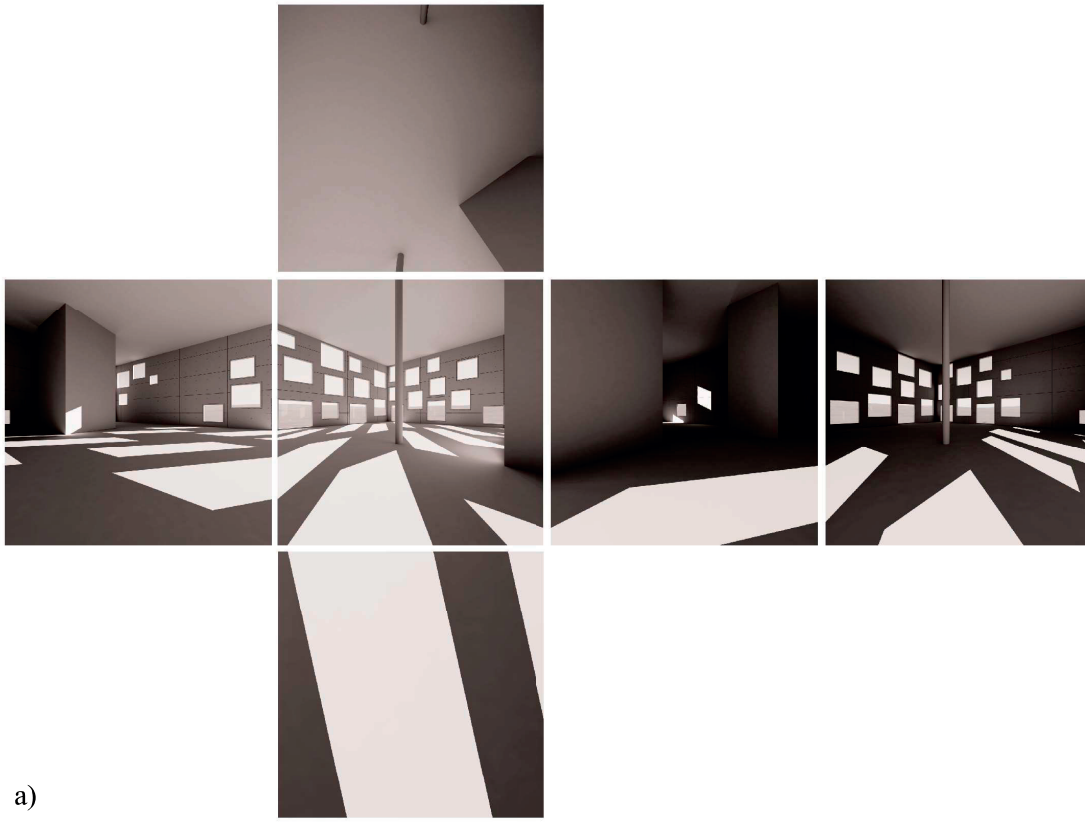
Figure A4.12 Final renderings in cube map projection for Spencertown a) clear sky & b) overcast



a)

b)

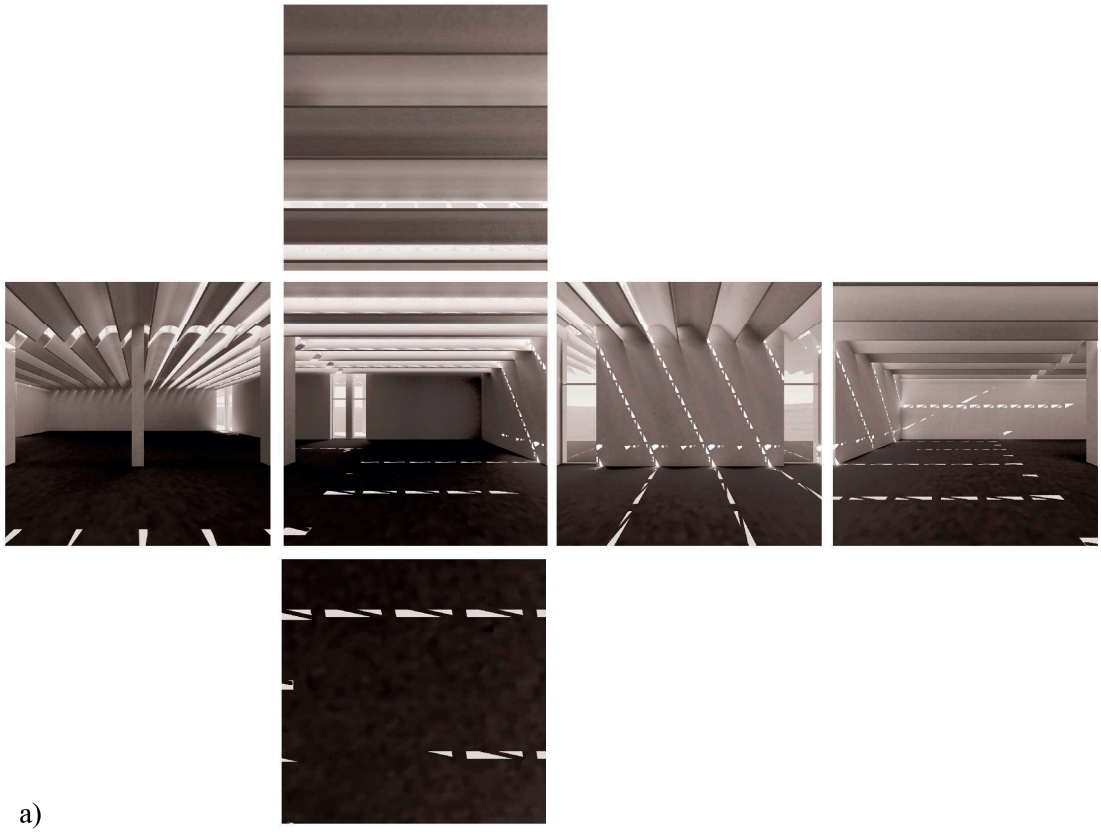
Figure A4.13 Final renderings in cube map projection for Poli a) clear sky & b) overcast



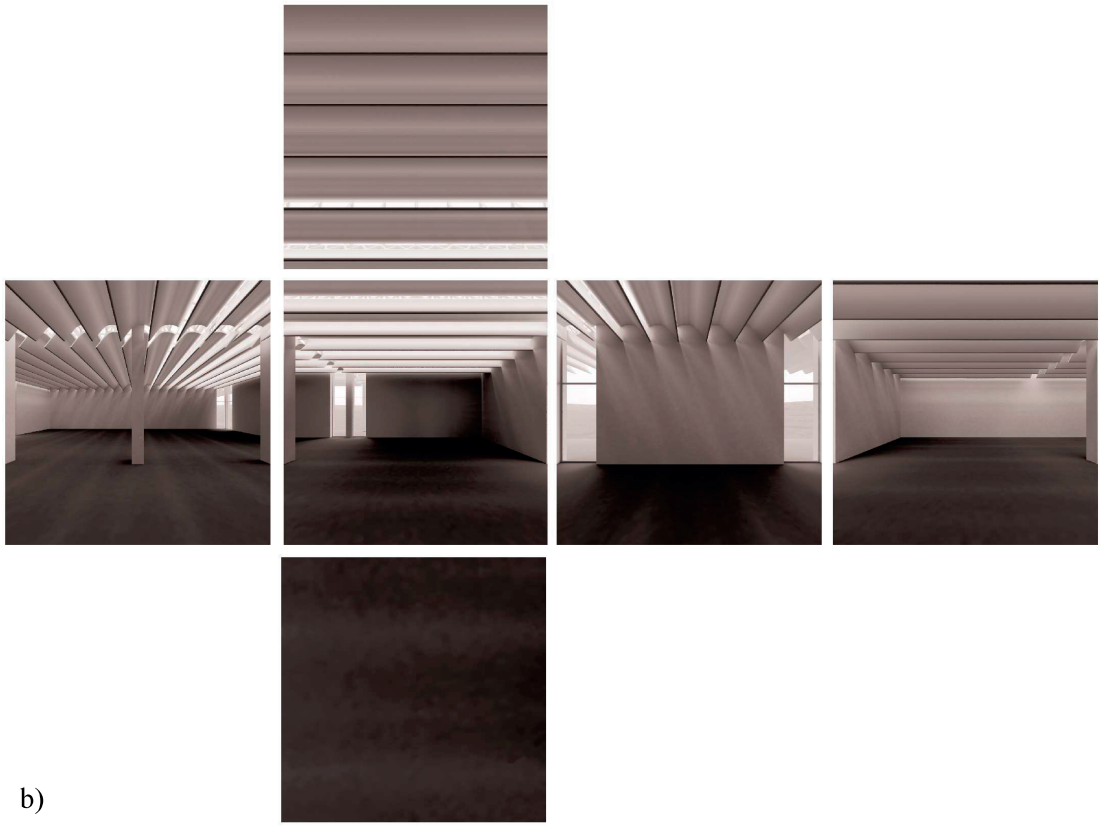
a)

b)

Figure A4.14 Final renderings in cube map projection for Zollverein a) clear sky & b) overcast

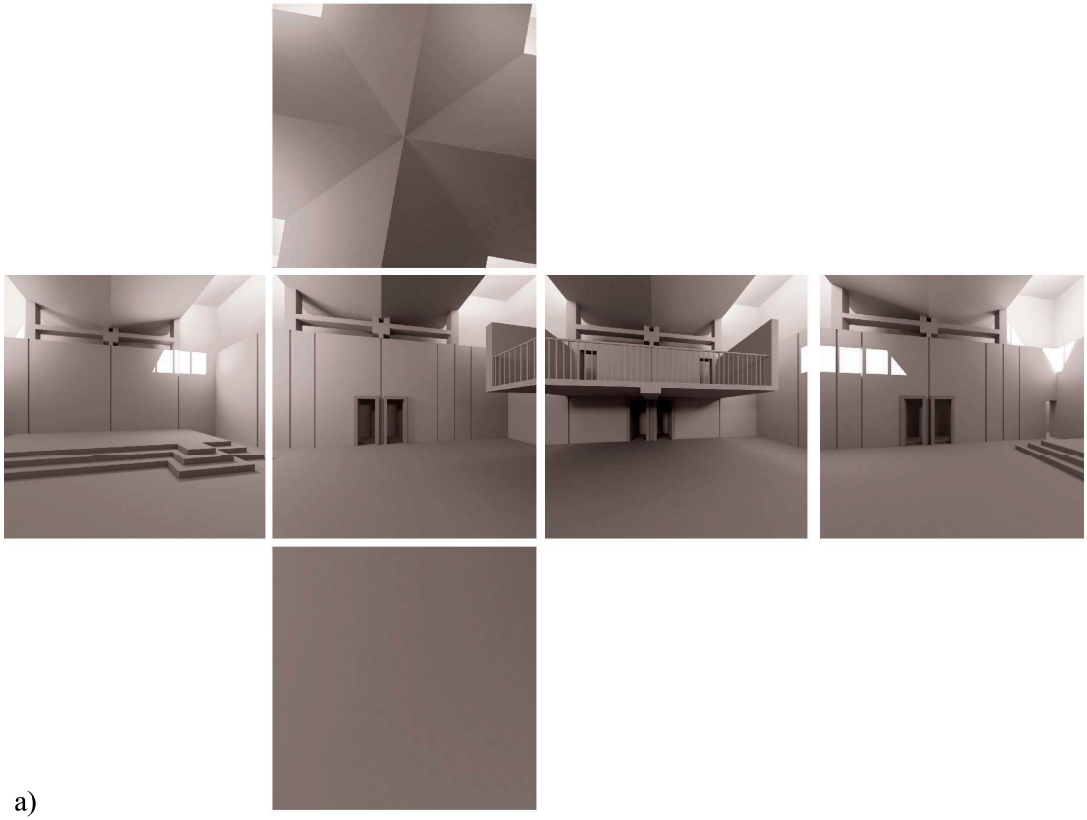


a)

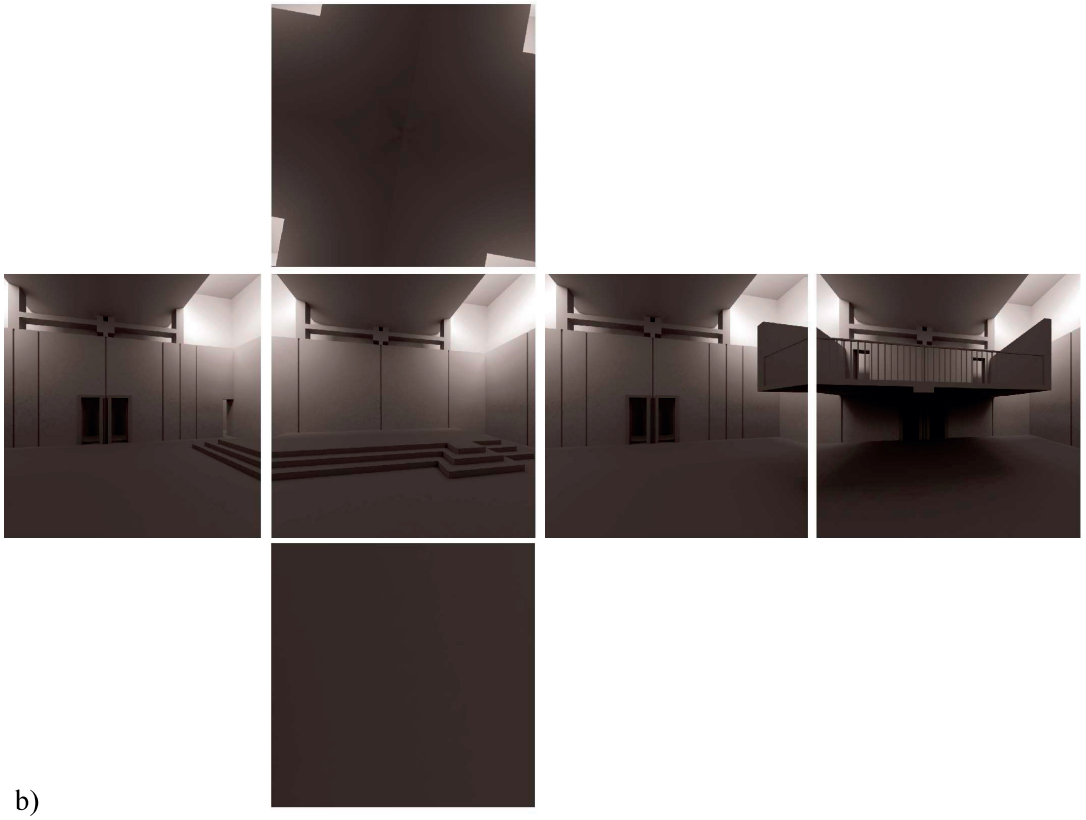


b)

Figure A4.15 Final renderings in cube map projection for Menil a) clear sky & b) overcast



a)



b)

Figure A4.16 Final renderings in cube map projection for First a) clear sky & b) overcast

ENAC - SCHOOL OF ARCHITECTURE, CIVIL AND ENVIRONMENTAL ENGINEERING
IA – INSTITUTE OF ARCHITECTURE
LIPID INTERDISCIPLINARY LABORATORY OF PERFORMANCE-INTEGRATED DESIGN)

EPFL-ENAC-IA-LIPID
LE 1111 (Bâtiment LE), Station 18
CH-1015 Lausanne - Switzerland

Phone : +41 21 69 30884
E-mail : kynthia.chamilothori@epfl.ch
Website : http://lipid.epfl.ch

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

INFORMATION SHEET FOR PARTICIPANTS IN RESEARCH STUDIES

You will be given a copy of this information sheet to keep.

Project title: Perception of daylit spaces in Virtual Reality

You are being asked to take part in a research study about your perceptual impression of spaces in virtual reality. Please read this form carefully and ask any questions you may have before agreeing to take part in the study.

Details of the study:

The purpose of this study is to understand how different parameters influence your perception of the interior of a daylit space in virtual reality.
If you agree to participate, you will be asked to wear a Virtual Reality (VR) Headset, in which you will be shown interiors of architectural spaces. We will ask you to fill in a short questionnaire verbally, while the researchers will be noting your answers. The questionnaire will include questions on your perception of the different scenes in the virtual reality headset, besides some personal information. We will also take some environmental measurements (such as air temperature and amount of light) and ask you to do a visual acuity test in the VR Headset. The experimental session will last about 20 minutes.

There are no anticipated risks in the participation of this study other than those encountered in day-to-day life. The use of the VR headset might induce a slight headache to sensitive participants for a few minutes after the completion of the experimental session. There is no benefit or compensation for your participation.

Taking part in this study is completely voluntary. You can withdraw your participation in this study at any time without giving a reason or facing negative consequences. Your answers will be confidential. All data will be collected and stored safely and reported in an anonymous form, in accordance with the CH Federal law on data protection ("Loi fédérale sur la protection des données" – RS 235.1). Only the principal investigator and/or the members of the Research Ethics Committee have access to the original data under strict confidentiality.

Any injury or damage occurring during or following the participation in the above-mentioned research project and for which the responsibility of the EPFL can be established, is covered by the general liability insurance of EPFL (insurance policy no. 30/5.006.824 of the Bâloise Insurance), in accordance with the terms and conditions of the insurance.

To complement the above-mentioned insurance, it is your responsibility to have adequate health insurance and accident insurance coverage. In case of doubt, just like in case of discomfort or undesirable effects related to your participation in this study, please contact the researchers directly.

The researchers conducting this study are Kynthia Chamilothori, Siobhan Rockcastle and Dr. Jan Wienold. Please ask us if there is anything that is not clear or if you would like to receive more information. If you have questions later, you may contact Kynthia Chamilothori and Siobhan Rockcastle at **{kynthia.chamilothori, siobhan.rockcastle}**@epfl.ch.

Figure A4.17 Information sheet for participants in the VR experiment

page 1

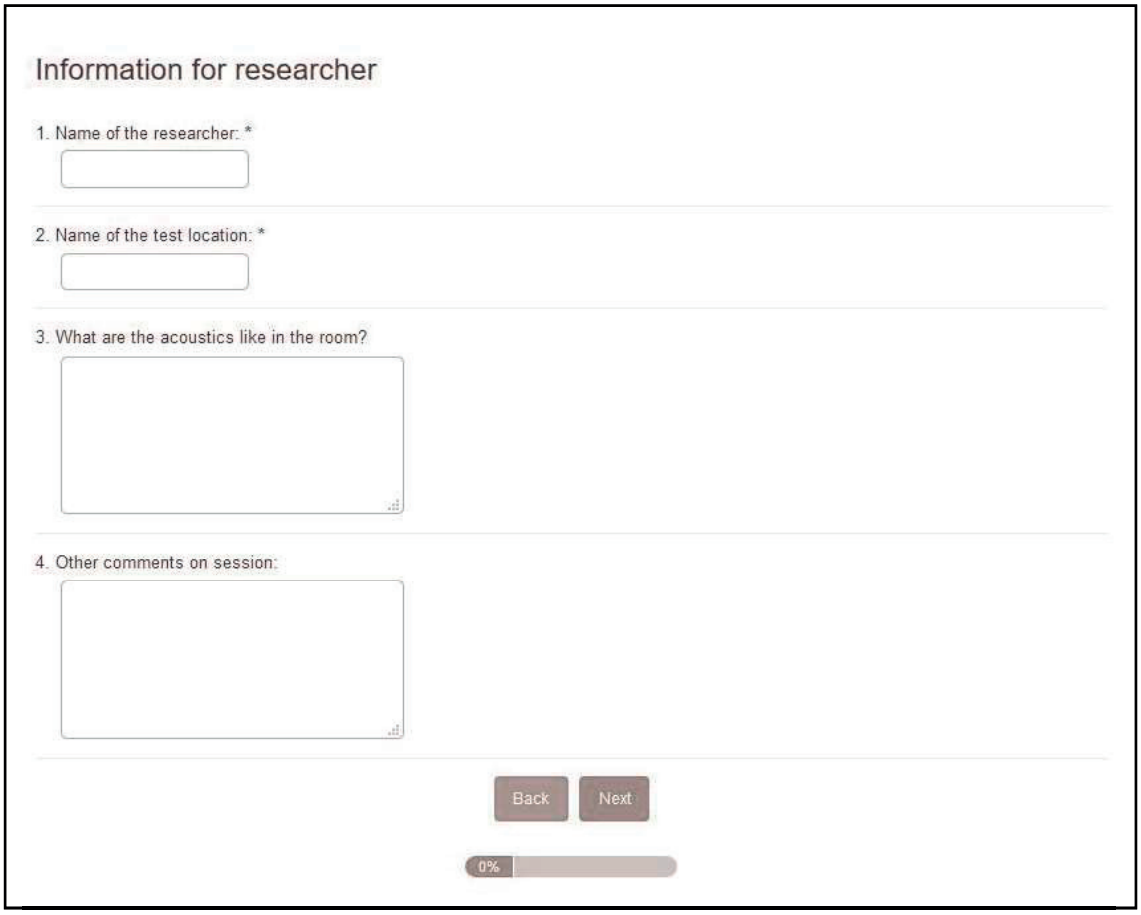
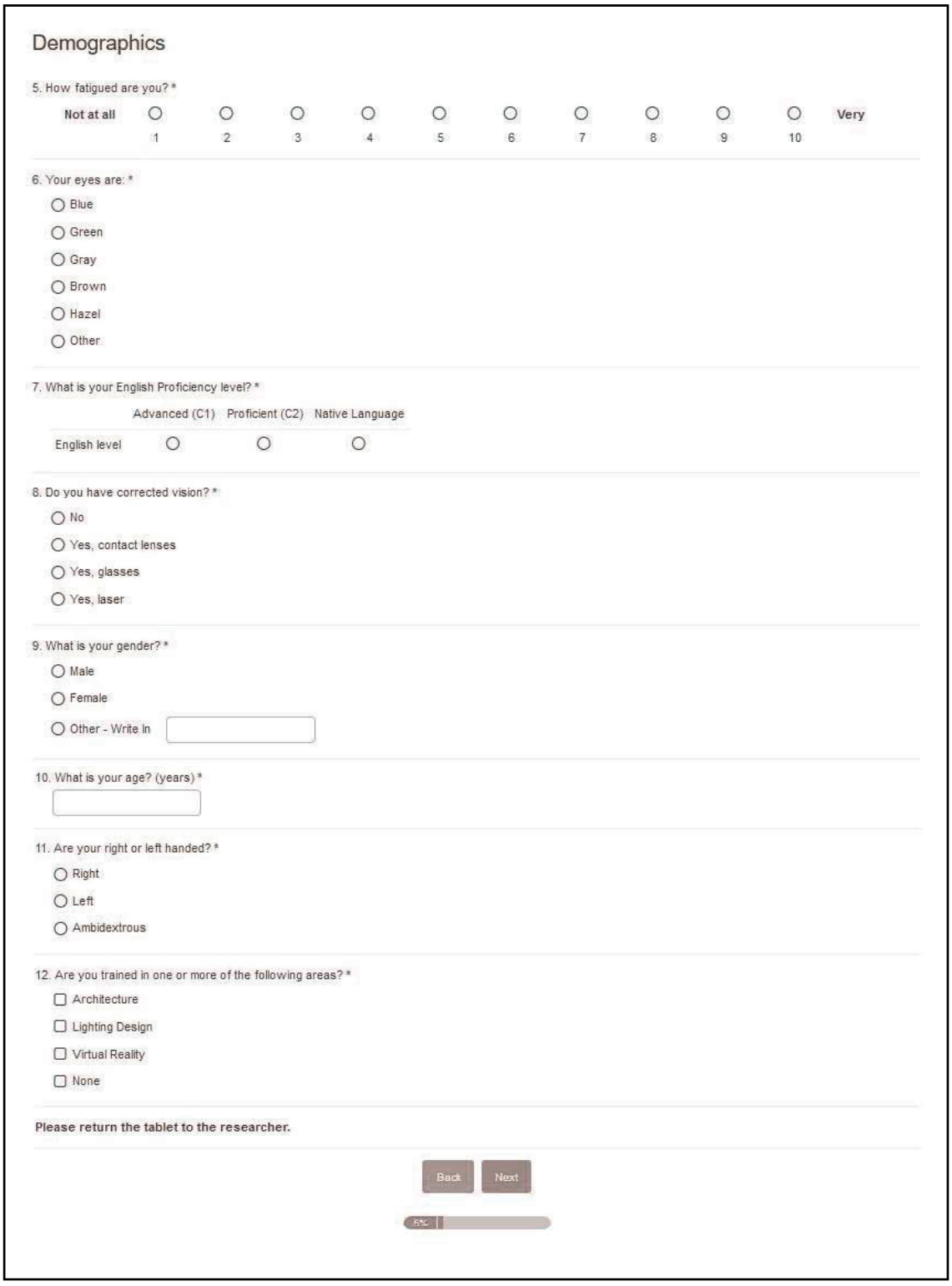


Figure A4.18 VR Survey (administered using Survey Gizmo on a tablet)

page 2



Demographics

5. How fatigued are you? *

Not at all ○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ 6 ○ 7 ○ 8 ○ 9 ○ 10 Very

6. Your eyes are: *

- ○ Blue
- ○ Green
- ○ Gray
- ○ Brown
- ○ Hazel
- ○ Other

7. What is your English Proficiency level? *

	Advanced (C1)	Proficient (C2)	Native Language
English level	○	○	○

8. Do you have corrected vision? *

- ○ No
- ○ Yes, contact lenses
- ○ Yes, glasses
- ○ Yes, laser

9. What is your gender? *

- ○ Male
- ○ Female
- ○ Other - Write In

10. What is your age? (years) *

11. Are your right or left handed? *

- ○ Right
- ○ Left
- ○ Ambidextrous

12. Are you trained in one or more of the following areas? *

- ☐ Architecture
- ☐ Lighting Design
- ☐ Virtual Reality
- ☐ None

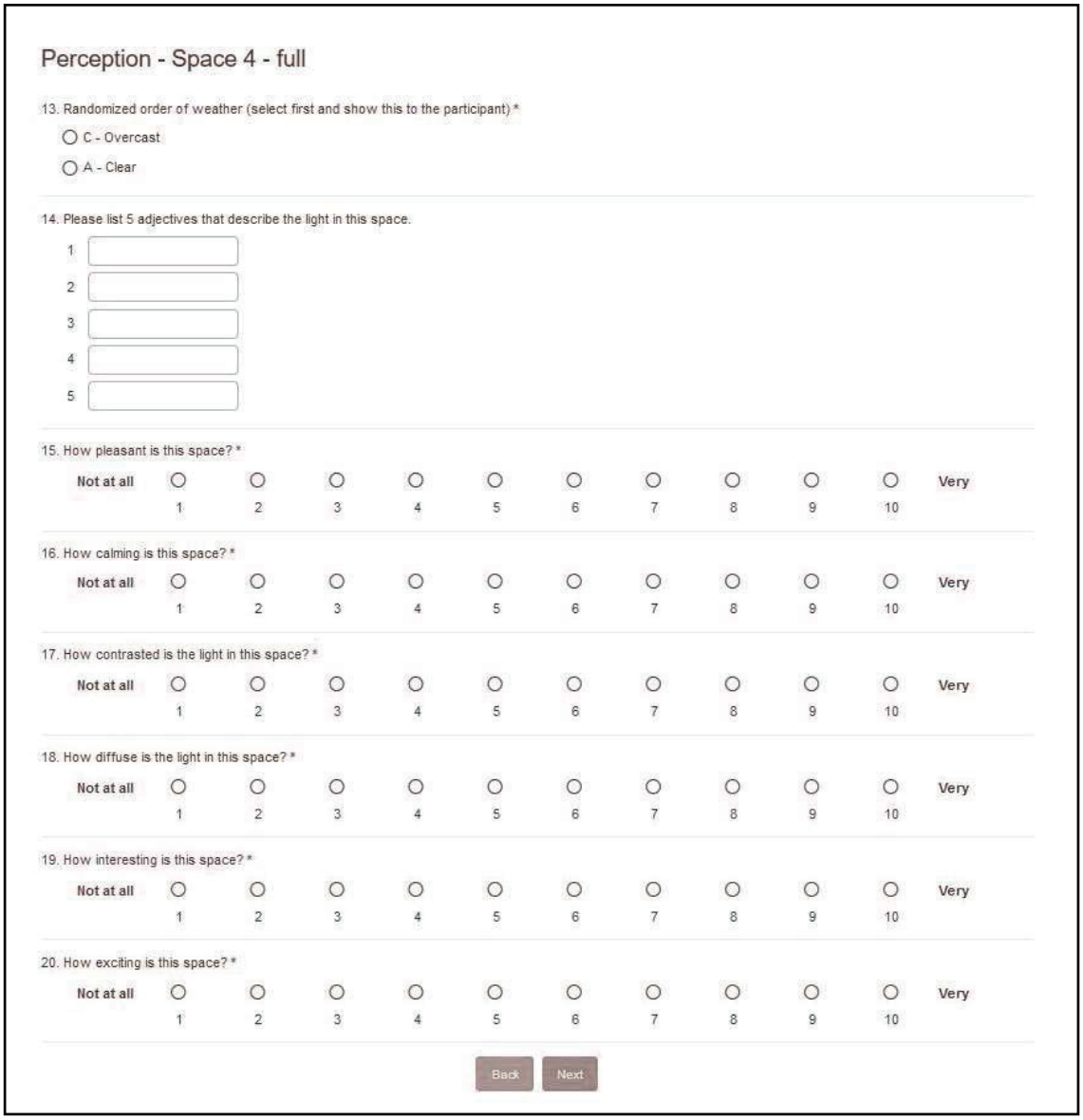
Please return the tablet to the researcher.

Back Next

5%

Figure A4.18 continued...

Questions delivered verbally for each projected scene
This example shows the page for Space 4 in the 360° group



Perception - Space 4 - full

13. Randomized order of weather (select first and show this to the participant) *

- C - Overcast
- A - Clear

14. Please list 5 adjectives that describe the light in this space.

1
2
3
4
5

15. How pleasant is this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

16. How calming is this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

17. How contrasted is the light in this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

18. How diffuse is the light in this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

19. How interesting is this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

20. How exciting is this space? *

	Not at all										Very
	1	2	3	4	5	6	7	8	9	10	

Back Next

Figure A4.17 continued...

Seeking volunteers

for an experiment using virtual reality

The Laboratory for Interdisciplinary Performance-Integrated Design is seeking healthy adult participants for an experiment on **visual perception** in architecture using **virtual reality.**

Requirements for participation:

age: 20 - 50 years
vision: full vision (or correction with contact lenses, no glasses)
time commitment: 20 minutes
location: mobile (we come to a seminar room close to you), multiple participants in one location is preferred
language: english C1 proficiency or higher

schedule a session:

For more information, please contact:

Siobhan Rockcastle
siobhan.rockcastle@epfl.ch

Kynthia Chamilothori
kynthia.chamilothori@epfl.ch

Figure A4.19 Advertisement for the VR experiment

	Douglas				First				
	clear		overcast			clear		overcast	
	180°	360°	180°	360°		180°	360°	180°	360°
pleasant	8.0	8.0	8.0	8.0	pleasant	4.0	4.5	3.0	4.0
interesting	9.0	8.0	7.5	8.0	interesting	4.0	5.0	5.0	6.0
exciting	8.0	8.0	7.0	8.0	exciting	4.5	4.0	4.0	5.0
calming	7.5	7.0	7.0	7.5	calming	5.5	5.5	5.0	5.0
diffuse	***5.0****	***3.0****	8.0	7.0	diffuse	6.0	7.0	7.0	7.0
contrasted	8.0	8.0	***3.5****	***6.0****	contrasted	6.0	6.0	7.0	5.0
PIE	8.3	7.7	7.8	8.0	PIE	4.7	4.5	3.7	5.7

Table 8.7 – Continued…

	Menil				Poli				
	clear		overcast			clear		overcast	
	180°	360°	180°	360°		180°	360°	180°	360°
pleasant	7.0	7.0	6.0	6.0	pleasant	6.0	6.0	6.0	7.0
interesting	6.0	6.0	5.0	6.0	interesting	6.0	7.0	6.0	6.0
exciting	4.0	5.0	4.0	5.0	exciting	6.0	6.0	6.0	6.0
calming	6.0	6.0	6.0	7.0	calming	6.0	6.0	***5.0****	***6.0****
diffuse	5.0	5.0	***6.0****	***8.0****	diffuse	6.0	6.0	6.0	7.0
contrasted	7.0	7.0	6.0	5.0	contrasted	8.0	7.0	6.0	5.0
PIE	5.3	5.7	5.0	6.0	PIE	5.3	6.3	6.0	6.0

Table 8.7 – Continued…

	Ryerson				Serpentine				
	clear		overcast			clear		overcast	
	180°	360°	180°	360°		180°	360°	180°	360°
pleasant	7.0	7.0	7.0	7.0	pleasant	7.5	6.0	***8.0****	***7.0****
interesting	7.0	7.0	7.0	7.0	interesting	8.0	8.0	9.0	8.0
exciting	8.0	7.0	7.0	6.0	exciting	7.5	7.0	7.0	6.0
calming	7.0	5.0	7.0	8.0	calming	4.5	5.0	6.0	6.0
diffuse	5.0	5.0	8.0	8.0	diffuse	2.0	3.0	***7.0****	***8.0****
contrasted	8.0	8.0	***3.0****	***5.5****	contrasted	8.5	9.0	***6.5****	***4.0****
PIE	7.7	6.7	7.0	7.2	PIE	7.7	7.0	7.8	7.7

Table 8.7 – Continued…

	Spencertown				Zollverein				
	clear		overcast			clear		overcast	
	180°	360°	180°	360°		180°	360°	180°	360°
pleasant	7.0	7.0	***8.0****	***6.0****	pleasant	***6.0****	***4.0****	6.0	7.0
interesting	7.0	7.0	6.0	6.0	interesting	6.0	5.5	7.0	7.0
exciting	6.5	7.0	5.5	5.0	exciting	6.0	6.0	***5.0****	***6.0****
calming	7.0	6.5	***8.0****	***6.0****	calming	5.0	5.0	7.0	8.0
diffuse	***2.0****	***4.0****	8.0	7.0	diffuse	4.0	3.5	7.0	7.0
contrasted	8.5	8.0	4.5	5.0	contrasted	8.0	8.0	***4.0****	***6.0****
PIE	6.7	7.2	6.2	5.7	PIE	6.0	5.2	6.0	6.7

*$p \leq 0.10$, **$p \leq 0.05$

Table A4.20 Median values for each rating scale, space, sky (high vd only) and group (180° and 360°). Asterics indicate an instance when the KS test revealed a statistically significant difference between groups (either 180° or 360° for the same space and sky).

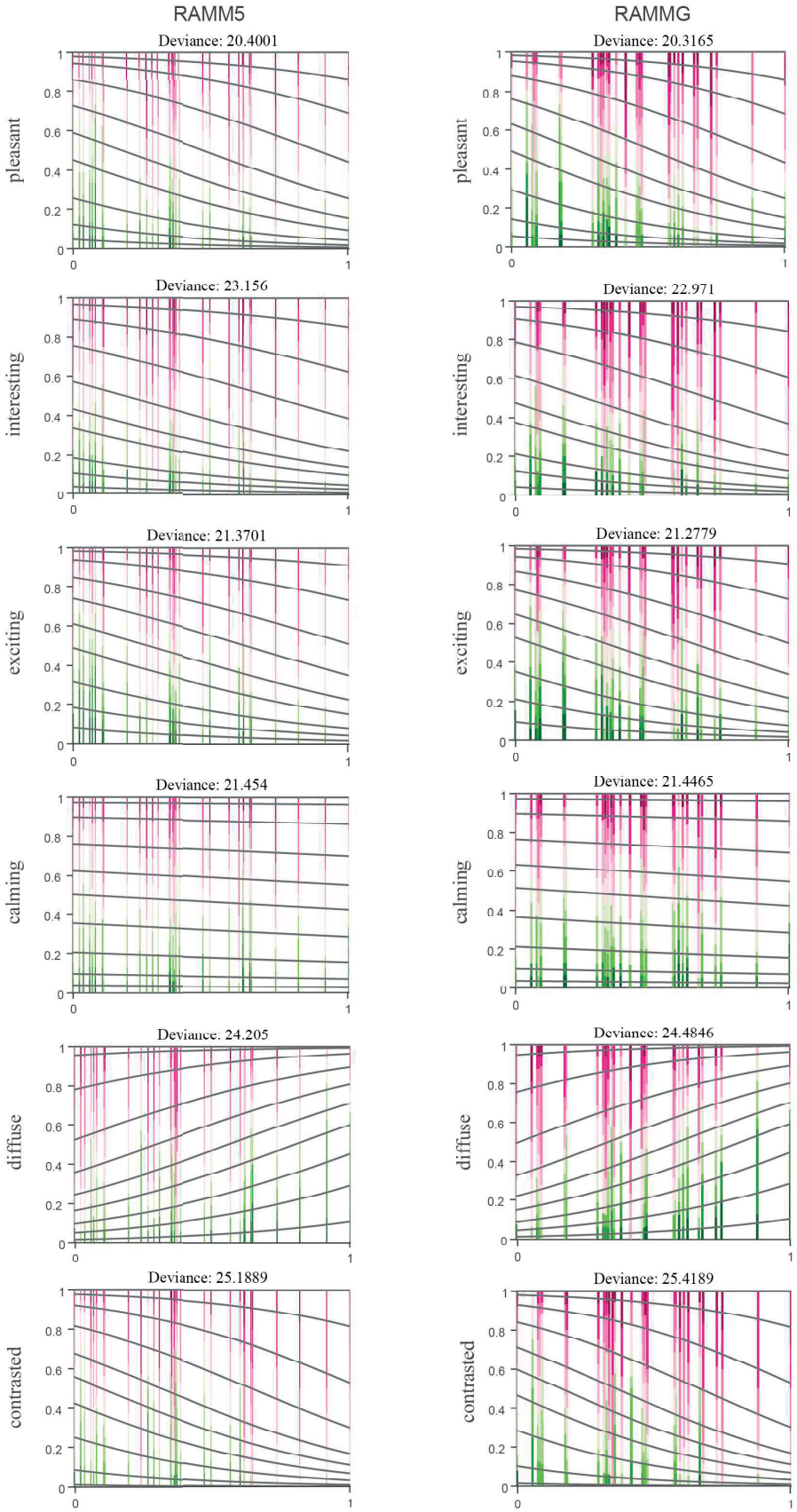


Figure A4.21 Logistic regression between subject ratings for each rating scale and a) RAMM5 or b) RAMMG for the 180° group.

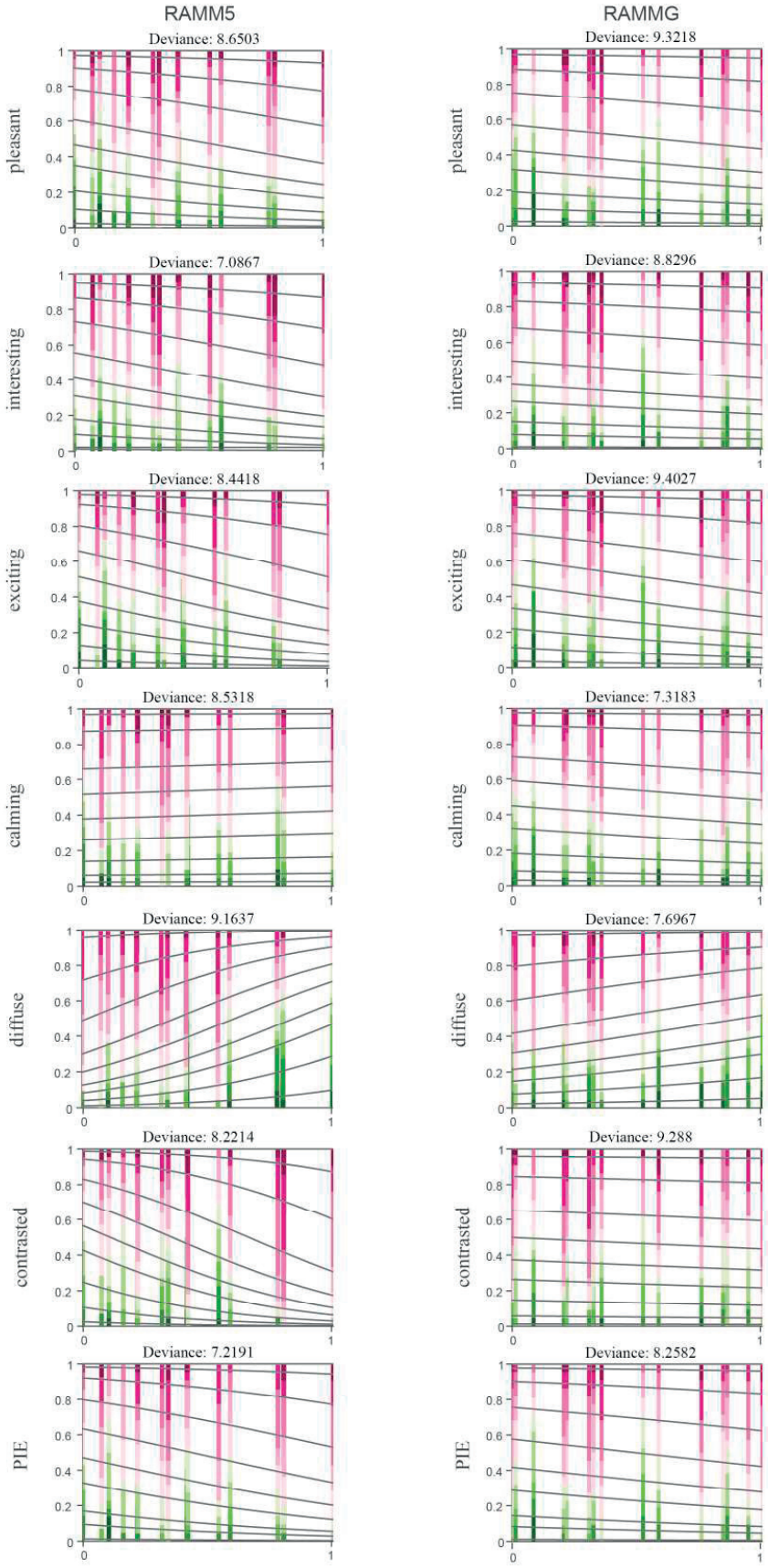


Figure A4.22 Logistic regression between subject ratings for each rating scale and a) RAMM5 or b) RAMMG for the 360° group.

A5 APPENDIX: CHAPTER 9

a)

			BMP					
			pcond +s		pcond -s		no tone-map	
	SKY	VD	mSC_5	RAMMG	mSC_5	RAMMG	mSC_5	RAMMG
Doug	clear	high	19.3	14.2	16.2	12.1	0.5	0.5
		low	11.6	11.1	10.2	9.6	0.9	1.1
	overcast	high	10.8	10.3	11.6	9.3	3.8	3.2
		low	11.2	9.2	7.5	6.9	8.4	8.9
First	clear	high	8.0	6.1	9.6	7.6	7.2	5.1
		low	6.7	6.3	6.9	6.4	4.8	4.2
	overcast	high	7.5	7.4	8.3	7.8	3.9	3.1
		low	6.5	5.9	6.7	6.8	3.1	3.4
Menil	clear	high	17.2	12.7	18.6	12.1	9.5	8.1
		low	9.6	8.8	6.3	5.5	7.6	7.4
	overcast	high	15.0	11.0	17.1	11.3	14.9	10.3
		low	6.5	6.3	4.3	4.3	5.3	5.4
Poli	clear	high	12.5	9.3	10.8	9.5	7.0	4.5
		low	7.3	9.1	5.3	5.3	5.9	6.6
	overcast	high	12.9	10.9	11.5	10.9	10.9	9.3
		low	8.0	6.2	5.6	4.3	6.2	4.7
Ryer	clear	high	20.7	13.6	17.4	12.1	2.9	1.7
		low	17.7	14.5	10.4	8.7	11.0	8.7
	overcast	high	12.6	10.4	13.1	11.3	12.2	10.8
		low	12.7	10.8	8.4	8.4	8.9	8.6
Serp	clear	high	24.0	17.6	20.7	14.8	1.1	1.2
		low	17.1	12.6	13.1	9.7	1.0	1.2
	overcast	high	16.2	12.3	15.0	10.8	1.9	2.0
		low	14.5	12.4	12.5	10.5	1.9	2.1
Spenc	clear	high	22.5	16.1	19.1	14.6	0.8	0.9
		low	12.4	9.9	8.4	6.7	4.5	3.3
	overcast	high	16.9	13.4	15.6	13.7	15.6	13.1
		low	7.1	7.4	5.4	5.5	5.8	6.0
Zoll	clear	high	17.6	12.9	20.7	15.4	0.7	0.7
		low	10.4	9.5	8.3	7.4	2.7	3.2
	overcast	high	12.3	9.6	12.9	10.2	11.6	9.5
		low	6.0	5.1	6.1	4.1	5.7	4.1

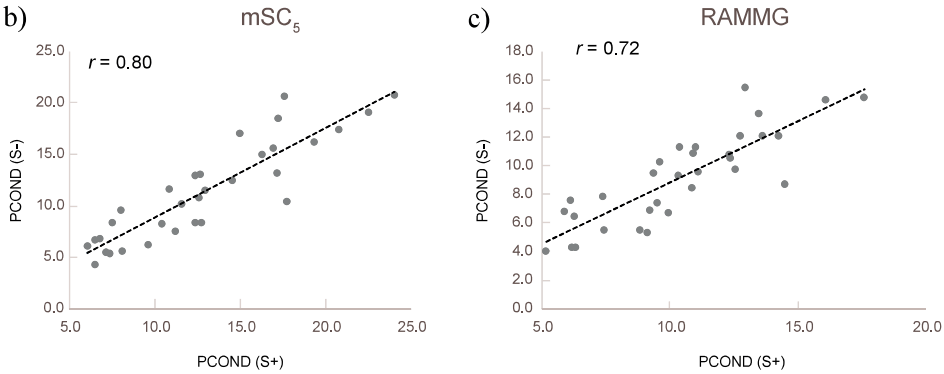
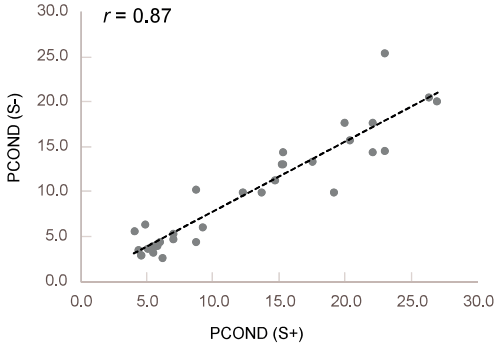


Figure A5.1 Results for a) mSC_5 and RAMMG ran on the BMP hemispherical renderings using the tone-mapping operator 'PCOND' with contrast sensitivity 'S' turned on (S+) and turned off (S-) alongside an un tone-mapped linear compression. Pearson correlation values (*r*) between results for b) mSC_5 and c) RAMMG each (with and without contrast sensitivity) are shown below.

a)

			HDR (*255)					
			pcond +s		pcond -s		nomap	
	SKY	VD	mSC_5	RAMMG	mSC_5	RAMMG	mSC_5	RAMMG
Doug	clear	high	23.0	17.1	14.5	10.9	0.1	0.0
		low	13.7	13.0	9.9	8.9	0.0	0.0
	overcast	high	15.2	12.3	13.0	10.1	6.6	5.4
		low	5.9	4.9	3.9	3.2	14.3	15.0
First	clear	high	8.8	6.7	10.3	8.1	6.4	4.2
		low	4.4	3.8	3.6	3.0	2.6	2.1
	overcast	high	4.9	3.2	6.3	4.6	1.9	1.0
		low	4.1	4.6	5.6	6.4	1.5	1.7
Menil	clear	high	22.1	13.8	17.7	10.6	16.1	12.8
		low	8.7	7.7	4.4	3.6	12.7	12.2
	overcast	high	20.0	12.7	17.7	10.5	22.8	14.4
		low	6.2	6.1	2.6	2.7	6.1	6.3
Poli	clear	high	14.7	12.1	11.3	11.7	12.4	7.7
		low	5.5	7.7	3.1	4.6	9.6	11.6
	overcast	high	12.3	12.1	9.9	11.3	13.0	12.2
		low	4.6	3.7	2.8	2.4	6.3	4.9
Ryer	clear	high	22.0	15.5	14.3	10.8	5.4	2.8
		low	6.0	5.4	4.5	4.0	19.7	15.2
	overcast	high	15.3	13.1	14.4	12.8	17.5	16.0
		low	5.5	7.1	3.9	5.4	8.4	9.6
Serp	clear	high	26.3	18.3	20.5	14.1	1.2	1.4
		low	19.2	14.0	10.0	7.2	0.9	1.0
	overcast	high	15.3	10.7	13.0	8.6	2.5	2.5
		low	12.3	9.9	10.0	7.9	2.6	2.7
Spenc	clear	high	26.9	20.5	20.0	16.4	0.7	0.5
		low	7.0	5.9	4.7	3.9	7.9	5.8
	overcast	high	20.4	17.2	15.7	14.8	22.4	20.2
		low	5.1	5.2	3.6	3.5	6.3	6.6
Zoll	clear	high	23.0	16.9	25.4	19.3	0.7	0.7
		low	9.3	8.0	6.0	5.7	5.2	6.3
	overcast	high	17.6	14.8	13.3	10.6	17.6	14.8
		low	7.0	4.8	5.2	3.3	7.7	5.6

b)



c)

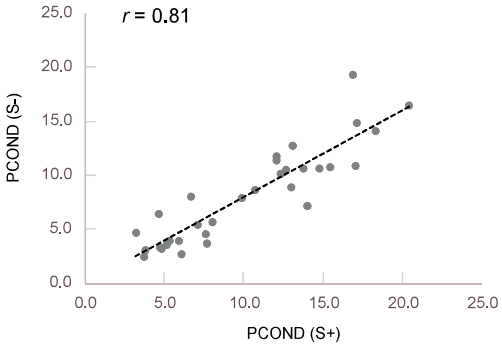


Figure A5.2 Results for a) mSC_5 and RAMMG ran on the HDR hemispherical renderings using the tone-mapping operator 'PCOND' with contrast sensitivity 'S' turned on (S+) and turned off (S-) alongside an un tone-mapped linear compression. Pearson correlation values (*r*) between results for b) mSC_5 and c) RAMMG each (with and without contrast sensitivity) are shown below. *All values are multipied by 255 to be in range with the BMP calculations in Figure A4.21.

a)

HDR vs. BMP

mSC_5 w/ PCOND (S+)

One-Way ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.022533	1	0.022533	0.000577	0.98091	3.995887
Within Groups	2420.486	62	39.0401			
Total	2420.508	63				

b)

HDR vs. BMP

RAMMG w/ PCOND (S+)

One-Way ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.297542	1	0.297542	0.017225	0.896009	3.995887
Within Groups	1071.008	62	17.27432			
Total	1071.306	63				

c)

PCOND (S+) vs. PCOND (S-)

mSC_5 w/ BMP

One-Way ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	35.54367	1	35.54367	1.473258	0.229436	3.995887
Within Groups	1495.806	62	24.1259			
Total	1531.349	63				

d)

PCOND (S+) vs. PCOND (S-)

RAMMG w/ BMP

One-Way ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	16.99114	1	16.99114	1.144348	0.288883	3.995887
Within Groups	920.5686	62	14.84788			
Total	937.5598	63				

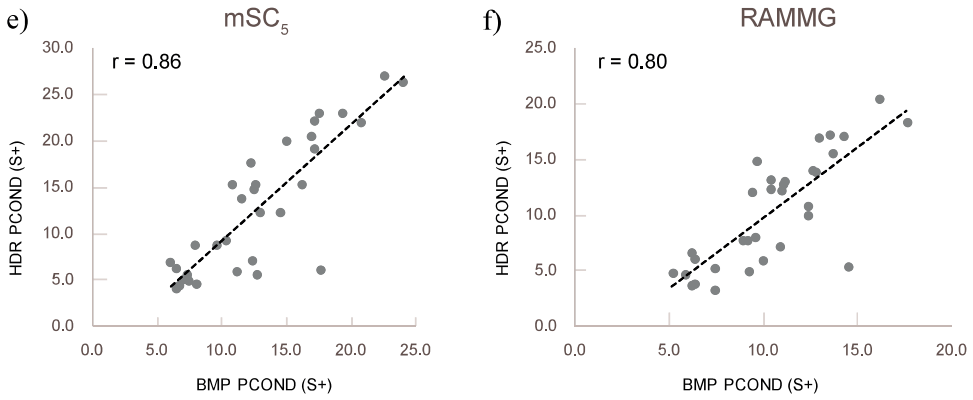


Figure A5.3 One-way ANOVA for a) HDR vs. BMP with PCOND (S+) on results for mSC_5, b) HDR vs. BMP with PCOND (S+) on results for RAMMG, c) PCOND (S+) vs PCOND (S-) on results for mSC_5 with BMP compression, and d) PCOND (S+) vs PCOND (S-) on results for RAMMG with BMP compression. Pearson Correlation Coefficient values for e) HDR vs. BMP with PCOND (S+) on results for mSC_5 and f) HDR vs. BMP with PCOND (S+) on results for RAMMG.

IMAGE CREDITS

P 5

Figure 1.2

Chapel of St. Ignatius ©Steven Holl (found at from http://www.stevenholl.com/projects/st-ignatius-chapel on February 15, 2016)

Figure 1.3

Chapel of St. Ignatius ©Steven Holl (found at from http://www.stevenholl.com/projects/st-ignatius-chapel on February 15, 2016)

Chapel of St. Ignatius by Joe Mabel using a wikimedia creative commons license (found at http://pcad.lib.washington.edu/image/1913/ on March 27, 2017)

P 6

Figure 1.4

Domestic Astronomy © Phillippe Rahm (found at http://www.philipperahm.com/data/projects/domesticastronomy/ on March 27, 2017)

P 10

Figure 1.5

The Kogod Courtyard by AgnosticPreachersKid using a a Creative Commons Attribution-Share Alike 3.0 (found at https://commons.wikimedia.org/w/index.php?search=kogod+courtyard&title=Special:Search&go=Go&uselang=en&searchToken=6cy05yz6h7dusdbm8vykxemaa#/media/File:The_Kogod_Courtyard.JPG on March 27, 2017)

Figure 1.6

Chapel of St. Ignatius ©Steven Holl (found at from http://www.stevenholl.com/projects/st-ignatius-chapel on February 15, 2016)

P 11

Figure 1.7

Modern Wing at the Art Institute of Chicago by Renzo Piano (found at http://www.archdaily.com/24652/the-modern-wing-renzo-piano on March 27, 2017)

P 143

Figure 8.1

Serpentine Pavillion 2002 by Toyo Ito by Lwbalmondstudio.co using a Creative Commons Attribution-Share Alike 4.0 International License (found at https://commons.wikimedia.org/wiki/File:1019_SP_Ito_2002_photo05_3072x2048_300dpi.jpg on April 26, 2017)

Ryerson Student Learning Center by Zeidler Partnership Architects and Snøhetta ©Lorne Bridgman (found at http://www.archdaily.com/771491/ryerson-university-student-learning-centre-zeidler-partnership-architects-plus-snohetta on April 26, 2017)

Douglas House by Richard Meier ©Scott Frances. Ezra Stoller/ESTO (found at http://www.richardmeier.com/?projects=douglas-house-2 on April 26, 2017)
Spencertown Residence by Thomas Phifer and Partners ©HomeDSGN 2011-2017 (found at http://www.homedsgn.com/2011/04/12/spencertown-residence-by-thomas-phifer-and-partners/ on March 26, 2017)

Menil Collection by Renzo Piano ©Arup (found at http://www.arup.com/projects/the_menil_collection on march 26, 2017)

Zollverein School of Management by SANAA ©Thomas Mayer/Stiftung Zollverein (found at https://www.zollverein.de/angebote/sanaa-gebaeude-events-mit-ein-und-ausblicken on March 26, 2017)

First Unitarian Church by Louis Kahn (found at http://rudygodinez.tumblr.com/post/73247581100/louis-kahn-first-unitarian-church-rochester-ny-1962 on March 27, 2017)
Poli House by Pezo Ellrichshausen ©Cristobal Palma (found at https://arch.iit.edu/prize/mchap/selected-works/project/poli-house on March 27, 2017)
P 144

Figure 8.2

Zollverein School of Management by SANAA Zollverein School by p2cl using a Creative Commons Attribution-Share Alike 4.0 International License (found at http://www.architecturerevived.com/zollverein-school-of-management-and-design-essen-germany/ on March 27, 2017)

REFERENCES

Ackerman, M. (2010). *Cool Comfort: America's Romance with Air Conditioning.* Washington D.C.: Smithsonian Books.

Addington, M. (2003). Energy, body, building: rethinking sustainable design solutions. *Harvard Design Magazine, 18*, 1-6.

Adelson, E., Anderson, C., Bergen, J., Burt, P., & Ogden, J. (1984). Pyramid methods in image processing. *RCA Engineer, 29*.

Akyuz, A., Fleming, R., Riecke, B., Reinhard, E., & Bulthoff, H. (2007). Do HDR displays support LDR content? A psychophysical evaluation. *ACM Transactions on Graphics, 26*(3).

Allen, E., & Zalewski, W. (2009). *Form and Forces: Designing Efficient, Expressive Structures.* Hoboken: John Wiley and Sons.

Ámundadóttir, M. L. (2016). Light-driven model for identifying indicators of non-visual health potential in the built-environment. Lausanne: EPFL PhD Thesis.

Ámundadóttir, M., Lockley, S. W., & Andersen, M. (2016). Unified framework to evaluate non-visual spectral effectiveness of light for human health. *Lighting Research and Technology*.

Ámundadóttir, M., Rockcastle, S., Sarey Khanie, M., & Anderden, M. (2017). A human-centric approach to assess daylight in buildings for non-visual health potential, visual interest and gaze behavior. *Building & Environment, 113*(15), 5–21.

Andersen, M., Gagne, J., & Kleindienst, S. (2013). Interactive expert support for early stage full-year daylighting design: a user's perspective on Lightsolve. *Automation in Construction, 35*, 338–352.

Andersen, M., Guillemin, A., Amundadottir, M., & Rockcastle, S. (2013). Beyond illumination: An interactive simulation framework for non-visual and perceptual aspects of daylighting performance. *In Proceedings of IBPSA 2013.* Chambéry.

Andersen, M., Kleindienst, S., Yi, L., Lee, J., Bodart, M., & Cutler, B. (2008). An intuitive daylighting performance analysis and optimization approach. *Building Research and Information , 36*(6), 593-607.

Ashikhim, M. (2002). A tone mapping algorithm for high contrast images. In P. Debevec, & S. Gibson, *Eurographics Workshop on Rendering* (pp. 145-155). Pisa: The Eurographics Association.

Baker, N., Fanchiotti, A., & Steemers, K. (1993). *Daylighting in Architecture: A European Reference Book.* London: Routledge.

Banham, R. (1969). *Architecture of the Well-Tempered Environment* (2nd ed.). Chicago: Architectural Press.

REFERENCES

Banos, R., Botella, C., Rubio, I., Quero, S., Garcia-Palacios, A., & Alcaniz, M. (2008). Presence and emotions in virtual environments: the influence of stereoscopy. *CyberPsychology & Behavior, 11*(1).

Bean, A., & Bell, R. (1992). The CSP Index: A practial measure of office lighting quality as perceived by the office worker. *Lighting Research and Technology, 24*(4), 215-225.

Bishop, I., & Rohrmann, B. (2003). Subjective responses to simulated and real environments: a comparison. *Landscape Urban Planning, 65*(4), 261–277.

Boubekri, M., Hull, R., & Boyce, L. (1991). Impact of window size and sunlight penetration on office worker's mood & satisfaction: A novel way of assessing sunlight. *Environment & Behavior, 23*(4), 474-493.

Boyce, P. (2003). *Human Factors in Lighting.* London: Taylor & Francis.

Calabria, A., & Fairchild, M. (2003). Perceived image contrast and observer preference II. Empirical modelling of perceived image contrast and observer preference data. *The Journal of Imaging Science and Technology, 47*, 494-508.

Calatrava, S. (1997). *Santiago Calatrava's Creative Process.* Basel; Boston; Berlin: Birkhäuser Architecture.

Cauwerts, C. (2013, November). Influence of presentation modes on visual perceptions of daylit space. Louvain-la-Neuve: Universite Catholique de Louvain PhD Thesis.

Cauwerts, C., & Bodart, M. (2011). Investigation of 3D projection for qualitative evalutation of daylit spaces. *PLEA 2011 - 27th Conference of Passive and Low Energy Architecture*, (pp. 819-824). Louvain-la-Neuve.

Cauwerts, C., Bodart, M., & Labayrade, R. (2013). Assessing lighting appearance using pictures: Influence of tone-mapping parameters and lighting conditions in the visualization room. *Proceedings to CISBAT 2013.* Lausanne.

Cetegen, D., Veitch, J., & Newsham, G. (2008). View size and office illuminance effects on employee satisfaction. *Proceedings to Balkan Light* (pp. 243-252). Ljubljana: Proceedings to Balkan Light.

Chamilothori, K., Wienold, J., & Andersen, M. (2016). Daylight patterns as a means to influence the spatial ambiance: a preliminary study. Volos: 3rd International Congress on Ambiances.

Chamilothori, K., Wienold, J., & Andersen, M. (2017). Adequacy of immersive virtual reality for the perception of daylit spaces: comparison of real and virtual environments. *LEUKOS: The Journal of the Illuminating Engineering Society of North America*, submitted.

Chamilothori, K., Wienold, J., & Andersen, M. (forthcoming). Perceptual accuracy of tone mapping operators in immersive virtual reality. Planned for submission in 2018.

CHPS. (2006). *Design for High Performance Schools.* Collaborative for High Performance Schools.

CIE. (1926). *Commission Internationale de L'Eclairage Proceedings.* Cambridge: Cambridge University Press.

Cooper, G. (1998). *Air-Conditioning America: Engineers and the Controlled Environment, 1900-1960.* Baltimore: Johns Hopkins University Press.

Corbusier, L. (1991). *Precisions: On the Present State of Architecture and City Planning.* Cambridge: MIT Press; First English Language Edition.

Cuttle, C. (2010). Towards the third stage of the lighting profession. *Lighting Research & Technology*, 73-93.

Cuttle, C. (2015). *Lighting Design: Perception-Based Approach.* New York: Routledge.

Demers, C. (2007). A classification of daylighting qualities based on contrast and brightness analysis. *Conference Proceedings of the American Solar Energy Society*, (pp. 243-252). Cleveland, Ohio.

Drago, F., Myszkowski, K., Annen, T., & Chiba, N. (2003). Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum, 22*(3).

Drago, F., Myszkowski, K., Annen, T., & Chiba, N. (2003). Adaptive logarithmic mapping for displaying high contrast scenes. In P. Brunet, & D. Fellner, *Proceedings of Eurographics* (pp. 419-426).

Dubois, C., Demers, C., & Potvin, A. (2007). The influence of daylighting on occupants: comfort and diversity of luminous ambiances in architecture. Cleveland: Proceedings of the American Solar Energy Society.

Dubois, M., Cantin, F., & Johnsen, K. (2007). The effect of coated glazing on visual perception: a pilot study using scale models. *Lighting Research & Technology, 39*(3), 283-304.

Durand, F., & Dorsey, J. (2002). Fast bilateral filtering for the display of high-dynamic-range images. *Proceedings of ACM SIGGRAPH* .

Engelke, U., Stokkermans, M., & Murdoch, M. (2013). Visualizing lighting with images: converging between the predictive value of renderings and photographs. San Francisco: proceedings of SPIE 8651.

Flynn, J., Hendrick, C., Spencer, T., & Martyniuk, O. (1979). A guide to methodology procedures for measuring subjective impressions in lighting. *Journal of IES*, 95-110.

Franz, G., Von Der Heyde, M., & Bulthoff, H. (2005). An empirical approach to the experience of architectural space in virtual reality - exploring relations between features and affective appraisals of rectangular indoor spaces. *Automation in Construction, 14*, 165-172.

Gagne, J., Andersen, M., & Norford, L. (2011). An interactive expert system for daylighting design exploration. *Building and Environment*, 2351-2364.

Hendrick, C., Martyniuk, O., Spencer, T., & Flynn, J. (1977). Procedures for investigating the effect of light on impression: simulation of a real space by slides. *Environment and Behavior, 9*, 491-510.

REFERENCES

Heschong, L. (1979). *Thermal Delight in Architecture.* Cambridge: The MIT Press.

Heschong, L., van den Wymelenberg, K., Andersen, M., Digert, N., Fernandes, L., Keller, A., . . . Tanteri, M. (2014). *Approved Method: IES Spatial Daylight Autonomy (sDA) and Annual Sunlight Exposure (ASE).* IES standard LM-83-12.

Hess, R., Bradley, A., & Piotrowski, L. (1983). Contrast-coding in amblyopia. I. Differences in the neural basis of human amblyopia. *Proceedings of Royal Society of London Series B (217)*, (pp. 309-330). London.

Heydarian, A., Carneiro, J., Gerber, D., Becerik-Gerber, B., Hayes, T., & Wood, W. (2015a). Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations. *Automation in Construction, 4*, 116–126.

Heydarian, A., Pantazis, E., Carneiro, J., Gerber, D., & Becerik-Gerber, B. (2015b). Towards understanding end-user lighting preferences in office spaces by using immersive virtual environments. *Proceedings of the 2015 International Workshop on Computing in Civil Engineering.* Austin.

Holl, S. (1999). *The Chapel of St. Ignatius.* New York: Princeton Architectural Press.

Holl, S. (2011). *Color Light Time.* Zurich: Lars Muller Publishers.

Holl, S. (2016.). *Steven Holl - Chapel of St. Ignatius*. Retrieved February 15, 2016, from http://www.stevenholl.com/projects/st-ignatius-chapel

Hopkinson, R. (1970). Glare from windows. *Construction Research & Development Journal, 2*, 98-105.

http://www.diva-for-rhino.com. (2009). Retrieved from DIVA-for-Rhino.

http://www.radiance-online.org/. (2009). Retrieved from Radiance - A Validated Lighting Simulation Tool.

http://www.rhino3d.com. (2007). Retrieved 2010, from Rhinoceros.

http://www.surveygizmo.com. (2015). Retrieved from Survey Gizmo.

https://www.mathworks.com/. (2012, September). Retrieved from Mathworks.

Hubalek S., B. M. (2010). Office workers' daily exposure to light and its influence on sleep quality and mood. . *42*(1).

IES. (1966). *IES Lighting Handbook, Fourth edition.* New York: Illuminating Engineering Society.

Inanici, M. (2003). Transformation of high dynamic images into virtual lighting laboratories. *IBPSA.* Eindhoven.

IPCC. (2014). *Climate change 2014: synthesis report.* Geneva: Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Jakubiec, J., & Reinhart, C. (2012). The 'Adaptive Zone' - a concept for assessing discomfort glare through daylit spaces. *Lighting Research and Technology*, 149-170.

Keighley, E. C. (1973). Visual requirements and reduced fenestration in offices: a study of multple apertures and window areas. *Building Science, 8*, 321-333.

King-Smith, P., & Kulikowski, J. (1975). Pattern and flicker detection analysed by threshold summation. *Journal of Pysiology*, 519-548 vol 249 (3).

Kleindienst, S., & Andersen, M. (2012). Comprehensive annual daylight design through a goal-based approach. *Building Research & Information*, vol 40 (2) 154-173.

Kleindienst, S., Bodart, M., & Andersen, M. (2008). Graphical representation of climate based daylight performance to support architectural design. *LEUKOS*, 39-61.

Konis, K., Lee, E., & Clear, R. (2011). Visual comfort analysis of innovative interior and exterior shading systems for commercial buildings using high resolution luminance images. *LEUKOS, The Journal of the Illuminating Engineering Society of North America, 7*(3).

Kuang, J., Yamaguchi, H., Johnson, G., & Fairchild, M. (2004). Testing HDR image rendering algorithms. *RIT Scholar Works.*

Kuliga, S., Thrash, T., Dalton, R., & Hölscher, C. (2015). Virtual reality as an empirical research tool - Exploring user experience in a real building and a corresponding virtual model. *Comput. Environ. Urban*, 363–375.

Lam, W. (1986). *Sunlighting as Formgiver for Architecture* . New York: Van Nostrand Reinhold Company.

Larson, W., & Shakespear, R. (1997). *Rendering with Radiance.* San Francisco: Morgan Kaufmann.

Liljefors, A. (1997). Lighting and color terminology. *Paper Presented at a CIE Discussion.* Stockholm: Comission Internationale de l'Eclairage.

Lo, W., & Steemers, K. (2009). The art of brightness and darkness: A critical investigation on daylighting quality. *PLEA2009 - 26th Conference on Passive and Low Energy Architecture.* Quebec City.

Loe D, M. K. (1994). Appearance of lit environment and its relevance in lighting design: Experimental study. *26*(3).

Mahdavi, A., & Eissa, H. (2002). Subject evaluation of architectural lighting via computationally rendered images. *Journal of the Illuminating Engineering Society, 31*, 11-20.

Mardaljevic, J. (2000). Simulation of annual daylighting profiles for internal illuminance. *Lighting Research and Technology Vol 32 (3)*, 111-118.

Mardaljevic, J., Heschong, L., & Lee, E. (2009). Daylight metrics and energy savings. *Lighting Research & Technology, 41*(3), 251-283.

REFERENCES

Matekovic, K., Neumann, L., Neumann, A., Psik, T., & Purgathofer, W. (2005). Global contrast factor: a new approach to image contrast. Aire-la-ville: Computational Aesthetics in Graphics, Visualizationg, and Imaging.

Michelson, A. (1927). *Studies in Optics.* Chicagp: University of Chicago Press.

Millet, M. (1996). *Light Revealing Architecture.* New York: Van Norstrand Reinhold.

Montgomery, D. C. (2009). *Design and Analysis of Experiments.* Hoboken: John Wiley & Sons.

Moon, P., & Spencer, D. (1942). Illumination for a non-uniform sky. *Illuminating Engineering Vol 37 (10)*, 797-826.

Moscoso, C., Matusiak, B., & Svensson, O. (2016). Impact of window size and room reflectance on the perceived quality of a room. *Journal of Arch. & Planning Research, 32*(4), 294-306.

Murdoch, M., & Stokkermans, M. (2014). Effects of image size and interactivity in lighting visualization. *Proceedings of SPIE - The International Society for Optical Engineering 9014.*

Nabil, A., & Mardaljevic, J. (2006). The useful daylight illuminance paradigm: A replacement for daylight factors. *Energy and Buildings Vol 38 (7)*, 905-913.

Ne'eman, E., & Hopkinson, R. (1970). Critical minimum acceptable window size: a study of window design and provision of view. *Lighting Research & Technology, 2*(1), 17-27.

Newhsam, G., Seetzen, H., Veitch, J., Chaudhuri, A., & Whitehead, L. (2002). Lighting quality evaluations using images on a high-dynamic range display. *Proceedings of the ARCC / EAAE Conference on Architectural Research.*

Newsham, G., Cetegen, D., Veitch, J., & Whitehead, L. (2010). Comparing lighting quality evaluations of real scenes with those from high dynamic range and conventional images. *Transactions on Applied Perception*.

Newsham, G., Marchand, R., & Veitch, J. (2002). Preferred surface luminance in offices: A pilot study. Salt Lake City: IESNA Annual Conference.

Newsham, G., Richardson, C., Blanchet, C., & Veitch, J. (2005). Lighting quality research using rendered images of offices. *Lighting Res. Technol.*, 93-115.

Ochsendorf, J. (2013). *Guastavino Vaulting: The Art of Structural Tile* . New York: Princeton Architectural Press.

Omidfar, A., Niermann, M., & Groat, L. (2015). The use of environmental aesthetics in subjective evaluation of daylight quality in office buildings. *Proceedings of IES Annual Conference.*

Osterhaus, W. (2005). Discomfort Glare Assessment and Prevention for Daylight Applications in Office Environments. *Solar Energy Vol 79 (2)*.

Ourousoff, N. (2007, November 19). A delicate glass roof with links to the past. *New York Times*.

Ozdemir, A. (2010). The effect of window views' openness and naturalness on the perception of rooms' spaciousness and brightness: A visual preference study. *Scientific Research and Essays, 5*(16), 2275-2287.

Pallasmaa, J. (2005). *The Eyes of the Skin: Architecture and the Senses* (2nd ed.). Academy Press.

Parpairi, K., Baker, N., Steemers, K., & Compagnon, R. (2002). The luminance differences index: a new indicator of user preferences in daylit spaces. *Lighting Research and Technology, 34*(1), 53-68.

Pavel, M., Sperling, G., Riedl, T., & Vanderbeek, A. (1987). Limits of visual communication: the effect of signal-to-noise ratio on the intelligibility of American Sign Language. *Opt. Soc. Am. A, 4*(12), 2355 - 2365.

Rahm, P. (2017, January 14). *Winter House*. Retrieved from Philippe Rahm: http://www.philipperahm.com/data/projects/winterhouse/index.html

Rea, M. S. (2000). *IESNA lighting handbook: Reference and Application.* New York: Illuminating Engineering Society of North America.

Rea, M., & Ouellette, M. (1984). DBR Report No. 1160 An Assessment of the Blackwell Visual Task Evaluator, Model 3X. Ottowa : National Research Council of Canada.

Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J. (2002). Photographic tone reproduction for digital images. *SIGGRAPH '02 Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, (pp. 267-276). San Antonio.

Reinhard, E., Ward, G., Pattanaik, S., & Debevec, P. (2006). *High Dynamic Range Imaging.* San Francisco: Morgan Kaufmann.

Reinhart, C., Mardaljevic, J., & Rogers, Z. (2006). Dynamic Daylight Performance Metrics for Sustainable Building Design. *Leukos Vol 3 (1)*, 1-25.

Rizzi, A., Algeri, G., Medeghini, D., & Marini, A. (2004). A proposal for contrast measure in digital images. *Second European Conference on Color in Graphics, Imaging and Vision.* Aachen: Proceedings to CGIV .

Rizzi, A., Simone, G., & Cordone, R. (2008). A modified algorithm for perceived contrast measure in digital images. *Proceedings of the 10th International Symposium on Multispectral Colour Science,* Barcelona.

Rockcastle, S. (2011). *Daylight variability and contrast-driven architectural effect.* Cambridge: MIT SMArchS Thesis.

Rockcastle, S., & Andersen, M. (2012). Dynamic annual metrics for contrast in daylit architecture. *SimAUD Conference Proceedings.* Orlando.

Rockcastle, S., & Andersen, M. (2013a). *Annual Dynamics of Daylight Variability and Contrast: A Simulation Based Approach to Quanitfying Visual Effects in Architecture.* London: Springer-Verlag.

REFERENCES

Rockcastle, S., & Andersen, M. (2013b). Celebrating contrast and daylight varibaility in contemporary architectural design: A typological approach. *Proceedings to LUX EUROPA.* Krakow, Poland.

Rockcastle, S., & Andersen, M. (2014). Measuring the dynamics of contrast and daylight variability in architecture: A proof-of-concept methodology. *Building & Environment, 81*, 320–333.

Rockcastle, S., & Andersen, M. (2015). Human perceptions of daylight composition in architecture: a preliminary study to compare quantitative contrast measures with subjective user assessments in HDR renderings. *International Conference of the International Building Performance Simulation Association*, Hyderabad.

Rockcastle, S., Amundadottir, M., & Andersen, M. (2016). Contrast measures for predicting perceptual effects of daylight in architectural renderings. *Lighting Research & Technology, 0* (Online First), 1-22.

Rockcastle, S., Amundadottir, M., & Andersen, M. (2017). A simulation-based workflow to assess human-centric daylight performance. *Proceedings to simAUD 2017*, Toronto.

Rockcastle, S., Chamilothori, K., & Andersen, M. (2017). An experiment in virtual reality to measure daylight-driven interest in rendered architectural scenes. *Proceedings to the 15th International Conference of IBPSA*, San Francisco.

Ryan, R. (1995, November). Isolated Luminosity. *The Architectural Review*.

Salters, B., Murdoch, M., Sekulovksi, D., Chen, S.H., & Seuntiens, P. (2012). An evaluation of different setups for simulating lighting characteristics. San Francisco: Proceedings to SPIE 8291, Human Vision and Electronic Imaging XVII.

Sarey Khanie, M. (2015). Human responsive daylighting in offices: a gaze-driven approach for dynamic discomfort glare assessment. Lausanne: EPFL PhD Thesis.

Sarey Khanie, M., Stoll, J., Einhauser, W., Wienold, J., & Andersen, M. (2015). Development of a gaze-driven methodology for estimating luminance values in the field of view. *Lighting Research and Technology, accepted for publication*.

Seetzen, H., Whitehead, L., & Ward, G. (2003). High dynamic range display using low and high resolution modulators. Proceedings of the Society for Information Display International Symposium.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill Book Company, Inc.

Simone, G., Pedersen, M., & Hardeberg, J. Y. (2012). Measuring perceptual contrast in digital images. *Journal of Visual Communication and Image Representation, 23*(3), 491-506.

Simone, G., Pedersen, M., Hardeberg, J. Y., & Rizzi, A. (2009). Measuing Perceptual Contrast in a Multi-Level Framework. *7240.* San Jose: Human Vision and Electronic Imaging XIV.

Smolders, K. C., de Kort, Y. A., & van den Berg, S. M. (2013). Daytime light exposure and feelings of vitality: Results of a field study during regular weekdays. *Journal of Environmental Psycology, 36*.

Steane, M. A. (2011). *The Architecture of Light: Recent Approaches to Designing with Natural Light.* New York: Routledge.

Steane, M. A., & Steemers, K. (2004). "Environmental Diversity in Architecture." *Environmental Diversity in Architecture.* Ed. M.A. Steane, Ed. K. Steemers. New York: Routledge.

Stokkermans, M. G., Murdoch, M. J., & Engelke, U. (2012). Preference for key parameter of tone mapping operator in different viewing conditions. *Proceedings to Experiencing Light 2012.* Eindhoven.

Tadmor, Y., & Tolhurst, D. (2000). Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. *Vision Research 40*, 3145-3157.

Tiller, D., & Veitch, J. (1995). Perceived room brightness: Pilot study on the effect of luminance distribution. *Lighting Research & Technology, 27*(2).

Tremeau, A. (2000). Color contrast parameters for analysing image differences. *Proceedings of the University of Derby Color Image Science 2000 Conference* (pp. 11-23). Derby: Proceedings to Digital Image Science.

U.N. (1987). *Our Common Future: the World Commission on Environment & Development.* Oxford: New York: Oxford University Press.

USGBC. (2014). *LEED Reference Guide for Building Design and Construction, Version 4.* USGBC.

Veitch, J., & Newsham, G. (1998). Lighting quality and energy-efficiency on task performance, mood, health, satisfaction and comfort. *Journal of the Illuminating Engineering Society*, 107-129.

Veitch, J., & Newsham, G. (2000). Preferred luminous conditions in open plan offices: Research amd practice recommendations. *Lighting Research & Technology, 32*(4).

Velds, M. (1999). *Assessment of Lighting Quality in Office Rooms with Daylighting Systems.* Netherlands: Ph.D. Thesis, Delft University of Technology.

Villa, C., & Labayrade, R. (2010a). Calibrating a display device for subjective visual comfort tests: selection of light simulation programs and post-production operations. Vienna: Proceedings of the International Commission on Illumination CIE 2010 "Lighting Quality and Energy Efficiency".

Villa, C., & Labayrade, R. (2010b). Psychovisual assessment of tone-mapping operators for global appearance and color reproduction. (pp. 189-196). Joensuu: Proceedings of Colour in Graphics Imaging and Vision.

Villa, C., & Labayrade, R. (2013). Validation of an Online Protocol for Assessing the Luminous Environment. *Lighting Research & Technology, 45*(4), 401-420.

REFERENCES

Villa, C., & Labayrade, R. (2015). Psychovisual evaluations of many luminous environments on the internet. *Lighting Research & Technology, 47*(4), 405-418.

Vogels, I. (2008). Atmosphere metrics: Development of a tool to quantify experienced atmosphere. In J. Westerink, M. Ouwerkerk, T. Overbeek, & W. Pasveer (Eds.), *Probing Experience: From Assessment of User Emotions and Behaviour to Development of Products* (pp. 25-41). Springer.

Waldram, J. (1954). Studies in interior lighting. *Transaction of the Illuminating Engineering*, (pp. 95–124). London.

Ward, G. (1994). The RADIANCE Lighting Simulation and Rendering System. *Proceedings of '94 SIGGRAPH Conference* (pp. 459-472). Orlando: Proceedings of '94 SIGGRAPH conference.

Ward, G., Rushmeier, H., & Piatko, C. (1997). *A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes.* San Francisco : Lawrence Berkeley National Lab, University of California.

Weston, H. (1961). Rationally recommended illumination. *Transactions of the Illuminating Engineering Society.* London.

Whitehead, L., Ward, G., Stuerzlinger, W., & Seetzen, H. (2005). *US Patent No. 6,891,672.*

Wienold, J. (2009). Dynamic daylight glare evaluation. *Proceedings of International IBPSA Conference.*

Wienold, J., & Christofferson, J. (2006). Evaluating methods and development of a new glare prediction model for 152 daylight environments with the use of CCD cameras. *Energy and Buildings, 38*(7), 743-757.

Wymelenberg, K., & Inanici, M. (2009). A study of luminance distribution patterns and occupant preference in daylit offices. *PLEA2009 - 26th Conference on Passive and Low Energy Architecture.* Quebec City.

Wymelenberg, K., & Inanici, M. (2010). The Effect of Luminance Distribution Patterns on Occupant Preference in a Daylit Office Environment. *LEUKOS: The Journal of the Illuminating Engineering Society of North America, 7*(2), 103 - 122.

Yoshida, A., Blanz, V., Myszkowski, K., & Price, R. (2005). Perceptual evaluation of of tone-mapping operators with real-world scenes. *Electronic Imaging* (pp. 192-203). San Jose: International Society of Optics and Photonics .

Zeitzer, J. M., Dijk, D. J., Kronauer, R. E., Brown, E. N., & Czeisler, C. A. (2000). Sensitivity of the human circadian pacemaker to nocturnal light: melatonin phase resetting and suppression. *3*.

Zumthor, P. (2006). *Atmospheres: Architectural Environments - Surrounding Objects.* Basel: Birkhäuser.

GLOSSARY OF ACRONYMS & TERMS

2D	Two-dimensional
3D	Three-dimensional
3DS	3D Studio
AIE	Acceptable Luminance Extent
ANOVA	Analysis of Variance
ASE	Annual Sunlight Exposure
cd/m2	a unit of luminance
cDA	Continuous Daylight Factor
DF	Daylight Factor
DGI	Daylight Glare Index
DGIw	Daylight Glare Index
DGP	Daylight Glare Probability
DGPs	Daylight Glare Probability Simplified
DOG	Difference of Gaussians
DSP	Daylight Saturation Percentage
DXF	Drawing Exchange Format
EPFL	Ecole Polytechnique Federale de Lausanne
HDR	High Dynamic Range
KS	Kolmogorov-Smirnov two-sample test
KW	Kruskal-Wallis test or Kruskal-Wallis H test
LD-Index	Luminance Differences Index
LDR	Low Dynamic Range
LEED	Leadership in Energy and Environmental Design
LIPID	Laboratory for Interdisciplinary Performance-Integrated Design
LUX	a unit of illuminance
mSC_5	Modified Spatial Contrast, also RAMM5
nvR_D	Non-Visual Direct Response Model
OBJ	a type of object file
OCT	Octree - complied 3D scene file in Radiance
p value	probability of obtaining a result equal to or more extreme than what was actually observed
PCC	Pearson Correlation Coefficient
pcond	A tone-mapping algorithm
PIE	Average Pleasant, Interesting & Exciting composite rating
pinterp	A Radiance command to interpolate a new view from an existing image
r	Pearson Correlation Coefficient
RAD	Radiance file
RAMM1	The first pixel subsampling level in the RAMMG approach
RAMM5	The fifth pixel subsampling level in the RAMMG approach, also mSC_5
RAMMG	An acronym for the authors Rizzi, Algeri, Medeghini & Marini
Rhino	Rhinoceros
RMS	Root Mean Square
RSC	An acronym for the authors Rizzi, Simone & Cordone
SC	Spatial Contrast
sDA	Spatial Daylight Autonomy
SKP	Sketchup
TMO	Tone Mapping Operator
UDI	Useful Daylight Illuminance
VR	Virtual Reality

Siobhan Francois Rockcastle

Born 06.14.1985 in Minneapolis, Minnesota, U.S. Citizen
Avenue Recordon 15, Lausanne 1004 Switzerland
T: +41 78 622 33 70 | siobhan.rockcastle@gmail.com
www.siobhanrockcastle.com

EDUCATION

2013-current
Ecole Polytechnique Fédérale de Lausanne, *PhD Student*
Thesis: Measuring Perceptual Dynamics in Daylit Architecture
Director: Marilyne Andersen
School of Architecture, ENAC, Lausanne, Switzerland
Expected to finish in Spring, 2017

2009-2011
Massachusetts Institute of Technology, *SMArchS, Building Technology*
Thesis: Daylight Variability and Contrast-Driven Architectural Effect
Advisors: Marilyne Andersen, Terry Knight & Sheila Kennedy
SA+P, Cambridge, MA [SMArchS Thesis Award]

2003-2008
Cornell University, *B.Arch*
Thesis: Pig's Eye Park, Reoccupying the Dump
Advisors: Kevin Pratt & Leyre Asensio Villoria
AAP, Ithaca, NY [Alpha Rho Chi Bronze Medal]

TEACHING

2013 – current
Ecole Polytechnique Fédérale de Lausanne, *Teaching Assistant*
Architecture and Sustainability: Critical Approaches, Doctoral Level
Participated in the planning of lecture content, readings, and projects and lectured on case-study analysis and historical contexts for environmental performance evaluation.

Solar Decathlon Summer Workshop, Master Level
Led a team of 6 students in the documentation and analysis of case-study projects, supervised the early-stage design of concepts for EPFLs upcoming 2017 competition entry and lectured on daylight integration in architecture.

Space and Light: the Lighting Project, Master Level
Lectured on daylight performance software and led workshops in geometrical modelling and integration of performance evaluation software.

2011 – 2012
Northeastern University, *Teaching Fellow & Visiting Assistant Professor*
Architecture & Energy Systems, Bachelor Level
Independently planned and taught the required course for professional accreditation and lectured on architectural lighting, HVAC, acoustics, passive design, and energy management.

Comprehensive Design Studio, Bachelor Level
Coordinated 6 adjunct faculty (across 6 sections, composed of 70 student in total), led curriculum development for the terminal Bachelor-level studios, and

independently taught a section of 13 students. This course integrated technical lecture-based courses and promote the integration of structural, mechanical, constructive, and passive performance strategies.

Daylight Design Seminar, Master Level
Independently taught 12 master students advanced techniques in daylight modelling, simulation, and visualization to encourage a reiterative design process for sustainable lighting design.

2009 - 2010 **Massachusetts Institute of Technology,** *Teaching Assistant*
Comprehensive Design Studio, Master Level
Assisted instruction with Professors Andrew Scott, Sheila Kennedy & Cristina Parreño and introduced the use of environmental simulation software to teach digital modelling and energy/lighting evaluation.

2005 - 2009 **Cornell University,** *Teaching Associate*
Undergraduate Design Studios 101 & 102
Co-taught 65 undergraduate first year architecture students alongside 5 full-time teaching assistants under the supervision of Professors Vincent Mulcahy and Alex Mergold.

Summer College Introduction to Architectural Design Studio
Taught one section of 14 high school seniors in a 6 week intensive summer workshop under the supervision of Professors Vincent Mulcahy and Alex Mergold.

PROFESSIONAL

2014 - current **Independent Daylight Design Consultant,** Lausanne, Switzerland
- Campus Expansion, Ecole Hôtelière de Lausanne, Lausanne, Switzerland
- Bugaboo Corporate Headquarters, MSR Architecture, Amsterdam, Netherlands
- Automata Pavilion, TEN x TEN, Chicago, Illinois
- Fribourg Master Plan Competition, DEVspace, Fribourg [Honorable Mention]

2011- 2013 **Kennedy & Violich Architecture,** Boston, MA
Designer & Project Manager (April 2012 – January 2013)
- Minneapolis Riverfront Development Initiative, Minneapolis, MN
- Sunlight Delivery Fixtures for 3M Architectural Markets, Minneapolis, MN
- Bending Wood Vault for the Beaver Country Day School, Boston, MA

Design Consultant (February 2011 – September 2011)
- Solar-Powered Soft Rockers for MIT's 150th Anniversary, Cambridge, MA
- Sunlight Luminaires for 3M Architectural Markets, Minneapolis, MN
- Tozzer Library Renovation, Cambridge, MA

2010 – 2011 **Sekisui House & MIT,** Boston, MA & Tokyo, Japan
Research Assistant (January - February)
- Sustainable Housing Prototypes for future development in 2050, Tama, Japan

2010 **Snøhetta,** New York, NY
Intern Architect (May - August)
- Ryerson University Student Learning Center, Toronto **[Award of Excellence]**
- SFMOMA Competition, San Francisco, CA

2009 **Epiphyte Lab,** Ithaca, NY
Intern Architect (January - May)
- HSU Residence Ithaca, NY

2004-2008 **MSR Architecture,** Minneapolis, MN & Hyattsville, MD
Intern Architect (Intermittent)
- Urban Outfitters Corporate HQ Philadelphia, PA
- Farragut Street Row Homes, Hyattsville, MD

2005 – 2007 **Cornell Solar Decathlon Team,** Ithaca, NY & Washington D.C.
Architecture Team Co-Leader
- Concept through construction for Cornell's 2007 entry in the U.S. Department of Energy sponsored Solar Decathlon, exhibited (10/2007) on the National Mall in Washington, D.C.

2003 **Gensler,** New York, NY
Intern Architect (May - August)
- Thornton Tomasetti Offices, New York, NY

PUBLICATIONS

Books

2013 First Author, Rockcastle S., Andersen M., *Annual Dynamics of Daylight Variability and Contrast: A Simulation-Based Approach to Quantifying Visual Effects in Architecture*, Springer Brief in Computer Science, London, 2013. [http://infoscience.epfl.ch/record/181819]

2011 Research Assistant & Contributor, *re: New Town: Adaptive Urbanism & the Low-Carbon Community,* (Routledge, 2011)

Peer Reviewed Journal Papers

2016 Co-Author, Ámundadóttir M.L., Rockcastle S., Sarey Khanie M., Andersen M., *A Human-Centric Approach to Assess Daylight for Non-Visual Health Potential, Visual Interest and Gaze Behavior,* Building & Environment (published online first October 1st, 2016). [https://infoscience.epfl.ch/record/221590?ln=en]

2016 First Author, Rockcastle S., Ámundadóttir M.L., Andersen M., *A Comparison of Contrast-Based Measures for Predicting Perceptual Effects of Daylight in Architectural Renderings,* Lighting Research and Technology (published online first, April 15, 2016). [https://infoscience.epfl.ch/record/217503]

2014 First Author, Rockcastle S., Andersen M., *Measuring the Dynamics of Contrast and Light Variability in Architecture: A Proof of Concept Methodology*, Building and Environment, vol 81, November 2014, 320-333. [http://infoscience.epfl.ch/record/199718]

Peer Reviewed Conference Papers

2017 First Author, Rockcastle S., Ámundadóttir M.L., Andersen M., *Using Virtual Reality to Measure Daylight Driven Interest in Rendered Architectural Scenes,* Proceedings of IBPSA 2017 International Building Performance Simulation Association conference, San Francisco, 2017. (accepted)

2017 First Author, Rockcastle S., Chamilothori K., Andersen M., *A Simulation-Based Workflow to Assess Human-Centric Daylight Performance,* IBPSA 2017 - Symposium on Simulation for Architecture and Urban Design, Toronto, May 22-24, 2017. (accepted)

2016	First Author, Rockcastle S., Ámundadóttir M.L., Andersen M., *Predicting Visual Interest in Daylit Architectural Renderings: An Experimental Simulation-Based Approach,* SimAUD 2016 - Symposium on Simulation for Architecture and Urban Design, London, May 14-16, 2016. [http://infoscience.epfl.ch/record/215977]
2016	Co-Author, Pastore L., Rastogi P., Rockcastle S., Monari H., Rueff G., Andersen M., *Assessing the impact of contemporary urbanization on bioclimatic features of historic architecture through a two-step simulation process,* PLEA 2016 - Symposium on Passive and Low Energy Architecture, Los Angeles, July 11-13, 2016 [https://infoscience.epfl.ch/record/218598]
2015	First Author, Rockcastle S., Andersen M., *Human Perceptions of Daylight Composition in Architecture: A Preliminary Study to Compare Quantitative contrast Measures with Subjective User Assessments in HDR Renderings,* Proceedings of IBPSA 2015 International Building Performance Simulation Association conference, Hyderabad, India, December 7 – 9, 2015. [http://infoscience.epfl.ch/record/209191]
2013	Co-Author, Andersen M., Guillemin A., Amundadottir M. & Rockcastle S., *Beyond illumination: An interactive simulation framework for non-visual and perceptual aspects of daylighting performance,* Proceedings of IBPSA 2013 – International Building Performance Simulation Association conference, Chambéry, France, August 26-30, 2013. [http://infoscience.epfl.ch/record/188744]
2013	First Author, Rockcastle S., Andersen M., *Celebrating Contrast & Daylight Variability in Contemporary Architectural Design: A Typological Approach,* Proceeding to LUX EUROPA, Krakow, September 17-19, 2013. [http://infoscience.epfl.ch/record/186396]
2012	First Author, Rockcastle S., Andersen M., *Dynamic Annual Metrics for Contrast in Daylit Architecture,* Proceedings of SimAUD 2012 – Symposium on Simulation for Architecture and Urban Design, Orlando, March 26-30, 2012. [http://infoscience.epfl.ch/record/175191] [Best Paper Award]

EXHIBITIONS & ANIMATIONS

2015	*The Dynamics of Shadow: Architecture of Natural Light in Extreme Latitudes,* Animation for the 'Dynamics of Darkness in the North', Reykjavik, Iceland
2012	*Dynamic Daylight,* Department of Architecture, Northeastern University, Boston, MA
2011	*Cloud Canopy,* Kennedy & Violich Architecture & 3M, Saint Paul, MN
2011	*Soft Rockers,* Kennedy & Violich Architecture & MIT, Cambridge, MA

AWARDS, FELLOWSHIPS & GRANTS

2015	*Honorable Mention*, Fribourg urban design competition, Fribourg
2014 - 2018	*Velux Research Fellowship*, 100% PhD project funding, Velux Stiftung, Zurich
2013	*Architizer A+ Finalist*, Architecture & Learning Category, Soft Rocker, KVA matX
2013	*Architizer A+ Special Mention*, Architecture & Fabrication, Beaver Wood Vault, KVA
2012	*Best Paper Award*, SimAUD conference, Orlando
2011-2012	*Teaching Fellowship,* full-time visiting assistant prof., School of Arch., Northeastern
2011	*SMArchS Thesis Prize*, for 'Daylight Variability and Contrast-Driven Arch. Effect,' MIT
2011	*Award of Excellence*, Ryerson SLC, Snøhetta, Canadian Architect
2010-2011	*Samuel A. Marx Fellowship,* full tuition award for outstanding scholastic merit, MIT
2010	*Holcim Forum Student Poster Selection*, for ' channels for learning'

2008	*Alpha Rho Chi Bronze Medal*, for leadership and professional merit, Cornell
2008-2009	*Robert James Eidlitz Traveling Fellowship*, energy infrastructure Iceland, Cornell
2007	*Lifecycle Building Challenge*, Hon. Mention C. Lambur, K. Pratt & D. Cupkova
2007	*AIA NYS Student Design Award*, for 'growing Gibraltar,' Cornell University
2004	*First Place Baird Prize*, for 'pop-up cinema' in 1st year design competition, Cornell

SKILLS

Highly experienced in a range of 3D digital modelling and scripting software (Rhinoceros, Autodesk Revit, AutoCAD, Digital Project, Grasshopper, 3D Studio Max), fluent in graphic and illustration software (Adobe Illustrator, Adobe InDesign, Adobe Photoshop), intermediate proficiency in technical computing (Matlab), fluent in environmental performance simulation software (Lightsolve, DIVA for Rhino, Design Builder), experienced in digital fabrication tools (laser cutting, dual-axis digital foam cutting, 5-axis CNC cutting).

STUDENT ADVISING

PhD Theses

2016 Katie Heinrich, secondary supervision of PhD project, CITA, Royal Danish Academy of Fine Arts

Diploma Theses

2015 Odile Alliman & Emilie Revaz, Mise en scène: le théâtre et la ville, Diploma of Architecture, EPFL

2014 Madeleine Deshaires, Feeling Rice: Pavilion de la Chine, Diploma of Architecture, EPFL

Master Semester Projects

2015 Hélène Monari & Guillaume Rueff, Assessing the Bioclimatic Features of Historic Architecture in Contemporary Cities, Master in Energy Management & Sustainability, EPFL