## A preliminary study

The exploratory preliminary study was carried out to evaluate feasibility and to inform the experiment design of the main research. The objectives were:

- to acquire practice in recruiting participants, designing and running an experiment
- data analysis
- to test the experiment approach
- to obtain an indication of the correlation between the dependent variable and a range of potential independent variables which would inform the criteria in the main experiment

It was not the aim of the study to statistically test a hypothesis. The main part of the study consisted of brightness comparison of successive pairs of waiting images with similar reflectance values. The waiting room was selected as windowless space that many can relate to. The L*a*b* colour space was used as it is more perceptually uniform than other colour systems (Raynham, 2012) and has a luminance dimension, which is a variable of this research. Full details of this study are included in Appendix A.

This study indicated that there is a difference in brightness perception of chromatic and achromatic stimuli, with the chromatic image appearing brighter. There were a high number of contradicting responses when the same pair of images were presented in reverse order with the most inconsistent responses being for colours of a higher reflectance value. More saturated wall colours also indicated a higher brightness perception but they were least preferred by the participants.

Parts 2 and 3 of the preliminary study showed that colour has an emotional value that varies from person to person. 15 of the 17 participants chose a room with 3 walls coloured as the least they preferred to be waiting in. A Mann Whitney U and a median comparison test showed that the view angle of the room in the scene had no influence on the results.

## The main experiment

<omitted text>

For the outcome to be more useful and for a richer data set, the two-interval forced choice (2-IFC) brightness discrimination was combined with a multiple random staircase sequence. With the same simple discrimination task for the observer this method can measure the difference in perceived brightness between the chromatic and achromatic stimuli, providing magnitude as well as direction. The staircase method has been widely used in psychophysics studies since the 1960s (Cornsweet, 1962) in perception assessments of vision (García-Pérez, 1998) but also of pain (Gracely et al., 1988) and audition (Bernstein and Gravel, 1990, Bausenhart et al., 2015) for example. It is efficient and more precise as it is based on judgements of equality, thus limiting inter-personal variation attributed to inconsistent scaling across participants. The threshold value can be expressed in terms of physical units allowing direct comparisons of the criterion being measured.

In this experiment the chromatic stimulus (scene of room with coloured wall) would be the standard stimulus (fixed) and the achromatic stimulus would be the comparison stimulus (variable). The term forced choice implies that the response can be in the case of this research question, either interval 1: ‘the first room has more light’ or interval two: ‘the second room has more light’. The response, ‘the rooms have equal amount of light’ is not allowed and when unsure the observer makes an arbitrary choice. The staircase technique is based on the method of limits whereby the order and choice of stimuli is adaptive to the response. A response of the standard stimulus is brighter is followed by a comparison stimulus of higher luminance, and a response of the comparison stimulus is brighter is followed by a comparison stimulus of lower luminance. The aim is to determine a threshold i.e. perceptually indistinguishable brightness between of the standard and comparison stimuli, also referred to the point of subjective equality (PSE). The difference between the luminance of the standard and comparison stimulus at the point of threshold is the magnitude of difference in brightness

between the two. Therefore, while the equal brightness condition is not a possible response, it is inferred via staircase reversals ie. the point at which an observer changes opinion of which of the two stimuli appears brighter. The multiple random staircase method where staircases are interleaved and the order of viewing is randomised adds robustness as it excludes error resulting from the participant being able to predict a response from an apparent presentation pattern (Cornsweet, 1962, Gracely et al., 1988).

It was claimed that the 2-IFC leads to bias which can be towards both the first and second stimulus and loss of discrimination or detection sensitivity attributed to some or a combination of stimulus order, interval duration, temporal interval between the two stimul. Other factors such as observer experience have been mentioned. Bias and discrimination sensitivity in the use of the method in vision research was studied (eg. Yeshurun et al., 2008; Jäkel and Wichmann, 2006). Conflicting results have been reported and advantages for the first stimulus to be the standard and fixed stimulus have been suggested (Nachmias 2006) to deal with the issue of having to memorise the first stimulus for longer for example. There is no conclusion that one 2-IFC formula for dealing with bias and levelling sensitivity which works for one study will work for all. The varying reports on this matter may affirm this.

<omitted text>

As suggested by Yeshurun et. al (2008) and by Fotios and Houser (2013) in this study each staircase was performed twice, once with the chromatic stimulus as the first interval and once with the chromatic stimulus as the second interval in order to identify and report for the bias. The coincidence of reversals in the staircase method can indicate intra-personal variation.

## Method

### Observers

33 volunteer participants, 18 males and 15 females aged 21 to 43 (mean = 30.75, standard deviation = 5.5 years) took part in the main experiment. 21 participants self-reported normal vision, 11 corrected and 1 amblyopia. None of the participants reported colour blindness. 24% of the participants undertook formal education in lighting, 27% in architecture, 9% in both architecture and lighting and 18% in both lighting and interior design. 21% of the participants had not received formal education in any of the aforementioned three disciplines. None of the participants of this experiment had participated in any of the pilot experiments or the preliminary study and all signed a consent form.

### Stimuli

The test (achromatic) and reference (chromatic) stimuli were scenes of a waiting room [FIG.1], previously used in the preliminary study. The room was modelled using SKETCHUP 2016 where 29 achromatic (M00 to M28) and 3 chromatic (C1, C2, C3) scenes were generated. In the achromatic scenes the colour of the wall facing the viewer was changed by 2 points of L* on the L*a*b* colour space from one scene to the next. The material maps were created using Adobe Photoshop CC 2015. The 32 scenes were rendered using V-RAY for sketch-up v2.0. This software has been previously used to compute scenes for lighting research experiments (Villa and Labayrade, 2012). The number of standard (chromatic) scenes was limited to three to keep the experiment time to a maximum of 45 minutes in order to avoid observer fatigue. The three colours selected for the material maps have the same position relative to the origin and the axes on the L*a*b* colour space. The fourth colour corresponding to this position in the +a,-b quadrant (pink) was the one of the four selected to be excluded since it was the one of the least favourite colours for pilot study participants and also the one that gave clearer answers of brightness perception. The colours were chosen to have a low chroma a,b value and a relatively high reflectance, which is typical of surface colours used in interiors. Furthermore, there is limited reference to such colours in research on brightness and colour. Details of all the scenes are included in Appendix B, FIG 9. Adobe Photoshop was used to measure the L*a*b* values of the coloured walls at three locations so that they may be compared to the screen measurements.

*FIG 1: Scene C1 (Reference) and Scene M12 (Comparison)*

## Apparatus

The stimuli were displayed on a wall mounted display screen NEC V421 (active screen area 930mm wide x 523mm high) with a refresh rate of 60Hz and with an image resolution of 1920 x 1080. The image size excluding a black border was 1820 x 1010.

Custom software written in visual basic was used to present each pair of scenes and to record the results. To characterise the chromacity and reflectance of the display a spectrometer Avantes Avaspec-2048 was used to measure each image displayed on the screen at the same three locations as indicated in FIG 2. The spectrum of a standard lamp was measured using the same equipment and compared to the spectrum of the measured reference lamp in order to obtain a calibration factor at every 5nm interval. These factors were applied to each of the spectral values at 5nm intervals. The X,Y and Z tristimulus values corresponding to each of the spectra measured was found using formulae [1][1] to [3] and these in turn converted to L*a*b* values using formulae [4] to [6]. The calculated L*a*b* values [FIG 9] were used in the analysis of the results.

The maximum, minimum and average luminance values in cd/m$^2$ of the display screen were found by measuring the luminance of separate Red, Green and Blue samples of different values displayed on the screen using a luminance meter Konica Minolta LS-110?. By interpolating between these values, the real luminance of a 4x 4 average pixel sample on each image was determined using custom software written in visual basic. The scene luminance varied between 1 and 113 cd/m$^2$ and the mean luminance of each scene is listed in FIG 9.

The illuminance at the eye was measured using a Konica Minolta illuminance spectrophotometer CL-500A and on average found to be 9 lux with a black screen and 12 lux when an image was projected on the display screen, which represents photopic conditions.

1 All formulae are included in Appendix C

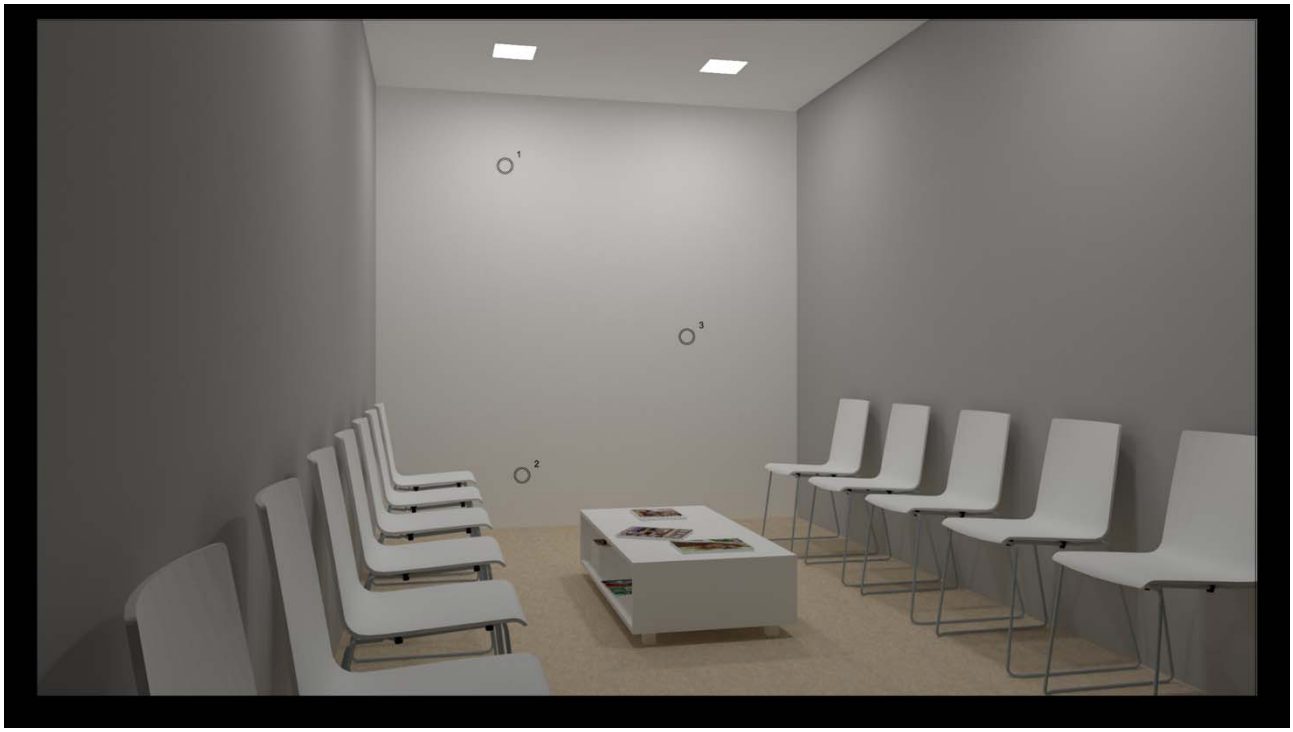



*FIG 2: Scene M28 showing the 3 spectra measurement points*

## The experiment question

Each participant was asked to compare pairs of scenes and determine which of the two had more light. The full instructions given to the participants are included in appendix B. The phrase ‘amount of light’ was specifically used in the instructions as the term brightness could be confused with how bright the colour appears to be. It was reported that brightness of a space equates to amount of light in a space. (Ishida and Ogiuchi, 2002)

## Procedure

The experiment was carried out in a meeting room at the basement level of Central House 14 Upper Woburn Place, London, WC1H 0NN. Blackout blinds to external windows were closed. The partition wall between the room and the corridor and a small return of the partition separating the experiment room from the adjacent room are glazed and some light entered the room from the artificially lit corridor. White paper was used to block the direct view of corridor recessed downlights.

An adaptation time of 10 minutes was allowed at the start of each experiment before the scenes were presented during which time the participants provided personal details and read the experiment instructions. Participants were seated 2.4m away from the screen with each scene subtending approximately 24º x 12º on the display. A preliminary test was performed with a pair of images before the start of the experiment to ensure that the task is understood. Achromatic image M06 (L*= 52) was used for this test.

On each trial the participant viewed a pair of images successively, one chromatic (standard stimulus) and one achromatic (comparison stimulus), but otherwise identical and positioned in the same spatial location on the screen. Each stimulus of the pair (first and second interval) was presented for 4 seconds with a dark period of 0.5s in between the two images. There was no opportunity for repeated viewing of the scenes. Participants were asked to choose which of the two rooms in the scenes appeared to have more light, room one or room two. The experimenter inputted the results in the custom software and selected the successive stimuli manually from a list using this software, every time referring to the previous response for the relevant staircase.

To address bias towards the first or second visual stimulus, each of the three staircases were repeated with the order of the images reversed. All 6 staircases were interleaved with the staircase order randomised according to the following sequence: 1, 5, 3, 4, 2, 6, 4, 2, 6, 1, 5, 3, 1, 5, 3…. The luminance L* of the scenes varied from L* =46.63 to L*=75.32, the mean absolute luminance of the whole scene varied between 24cd/$m^2$ and 45.9 cd/$m^2$ and the mean absolute luminance of the varying wall varied between 13.0cd/$m^2$ and 62.6cd/$m^2$. The values of L* and absolute luminance for each of the consecutive scenes is referred to in FIG 9. A step size of 4 succeeding scenes was used until the first reversal, 2 scenes between the first and second and 1 image until termination. [Even though participants were asked to answer as the second image disappeared from the screen, some participants responded earlier]

The total number of staircases received from all 33 participants was 198. The termination criterion was the completion of five reversals. There was no distinction between an up-step and a down-step. Each session took between 30 and 45 minutes in total. The last 3 reversals were used for threshold estimation.

*FIG 3: Experiment setting*

## Supplemental tests

Following an initial analysis which showed extensive intra personal variation, 3 participants whose staircases represented 3 distinct trends and who were available for a second session, were asked to repeat the experiment. This time each reversal needed to be confirmed twice before a change in staircase direction was made following the method by Davoudian et.al (2013) in order to rule out unexpected results arising from inaccurate reporting on colour blindness, all participants were asked to carry out an Ishihara test.

## Pilot studies

A series of trials were carried out with two participants before the actual experiment sessions started and adjustments were made to the range of stimuli, the staircase starting and termination points, the step size and the time interval for which the two stimulus intervals were shown. The time taken to complete the experiment was also determined during these experiments. The target number of trails was to be less than 25 (Cornsweet, 1962). Despite these trials the first two experiments highlighted a weakness in the experiment, whereby the range of stimuli was not wide enough for 5 reversals to be reached. These 2 experiments were treated as a pilot and excluded from the main analysis. Stimuli of higher luminance were prepared for subsequent experiments.

APPENDIX B

## Experiment scene details

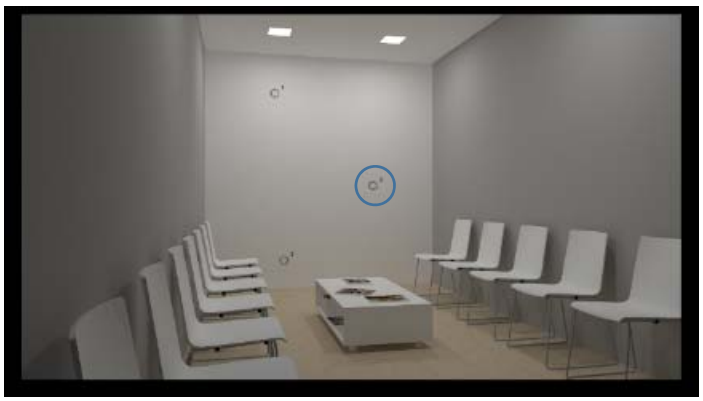
| 1 | 2 | 2 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| Scene No | Scene | L*a*b* values of material map | Material map | L*a*b* values of scene measured in Photoshop ** | L*a*b* values of scene based on spectra measured from the screen** | $C_{ab}$ based on scene screen measurement (6). Eq [7] | a,b hue - angle based on screen measurement (6) Eq. [8] | Mean scene luminance [cd/m²] | Mean luminance of varying wall [cd/m²] |
| | | | | | Step size | | | Step size [cd/m²] | Step size [cd/m²] |
| C1 |  | 80*-25*-25 | | 50*-17*-16* | 64.27*-8*-13* | 15.26 | 31.6 | 34.7 | 36.1 |
| C2 |  | 80*25*25 | | 50* 20* 20* | 63.75* 14* 16* | 21.26 | 41.18 | 33.2 | 39.2 |
| C3 |  | 80*-25*25 | | 50*-17* 18* | 63.16* -9* 13* | 15.81 | -34.69 | 34.3 | 36.3 |
| | | | | | | | | | |
| M00 |  | 45*0*0* | | 26* 0* 1* | 46.63* 7* 1* | | | 24.0 | 13.0 |
| | | | | | +0.88 | | | +0.1 | +1.0 |
| M01 |  | 47*0*0* | | 27* 0* 1 | 47.51* 6* 1* | | | 24.1 | 14.0 |
| | | | | | +0.53 | | | +0.4 | +0.7 |
| M02 |  | 49*0*0* | | 28* 0* 1* | 48.04* 6* 1* | | | 24.5 | 14.7 |
| | | | | | +1.19 | | | +0.3 | +1.3 |
| M03 |  | 51*0*0* | | 30* 0* 1* | 49.23* 5* 1* | | | 24.8 | 16.0 |
| | | | | | +0.77 | | | +0.1 | +1.9 |
| M04 |  | 53*0*0* | | 31* 0* 1* | 50.00* 5* 1 | | | 24.9 | 17.9 |
| | | | | | +0.74 | | | +0.3 | +0.5 |
| M05 |  | 55*0*0* | | 33* 0* 1* | 50.57* 5* 1* | | | 25.2 | 18.4 |
| | | | | | +1.23 | | | +0.2 | +0.9 |
| M06 | | 57*0*0* | | 34* 0* 1* | 51.98* 4* 1* | | | 25.4 | 19.3 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| |  | | | | +2.17 | | | +0.3 | +1.2 |
| M07 |  | 59*0*0* | | 35* 0* 1* | 54.15* 4* 2* | | | 25.7 | 20.5 |
| | | | | | +0.34 | | | +0.2 | +1.0 |
| M08 |  | 61*0*0* | | 36* 0* 1* | 54.49* 4* 2 | | | 25.9 | 21.5 |
| | | | | | +0.40 | | | +0.2 | +1.1 |
| M09 |  | 63*0*0* | | 38* 0* 1* | 54.89* 4* 1* | | | 26.1 | 22.6 |
| | | | | | +0.43 | | | +1.0 | +1.5 |
| M10 |  | 65*0*0* | | 39* 0* 1* | 55.33* 4* 1* | | | 27.1 | 24.1 |
| | | | | | +1.10 | | | +0.4 | +1.6 |
| M11 |  | 67*0*0* | | 41* 0* 1* | 56.43* 4* 1* | | | 27.5 | 25.7 |
| | | | | | +1.26 | | | +1.7 | +2.7 |
| M12 |  | 69*0*0* | | 42* 0* 2* | 57.69* 3* 2* | | | 29.2 | 28.4 |
| | | | | | +1.31 | | | +1.2 | +2.3 |
| M13 |  | 71*0*0* | | 43* 0* 2* | 59.00* 3* 2* | | | 30.4 | 30.7 |
| | | | | | +1.26 | | | +1.0 | +1.8 |
| M14 |  | 73*0*0* | | 45* 0* 2* | 60.27* 2* 3* | | | 31.4 | 32.5 |
| | | | | | +0.55 | | | +0.6 | +1.6 |
| M15 |  | 75*0*0* | | 46* 0* 2* | 60.82* 2* 2* | | | 32.0 | 34.1 |
| | | | | | +0.86 | | | +1.0 | +1.8 |
| M16 |  | 77*0*0* | | 48* 0* 2* | 61.68* 2* 2* | | | 33.0 | 35.9 |
| | | | | | +1.67 | | | +0.7 | +2.2 |
| M17 |  | 79*0*0* | | 49* 0* 2* | 63.35* 2* 2* | | | 33.7 | 38.1 |
| | | | | | +0.97 | | | -0.3 | +0.2 |
| M18 |  | 81*0*0* | | 50* 0* 2* | 64.32* 2* 3* | | | 33.4 | 38.3 |
| | | | | | +1.14 | | | +0.8 | +2.1 |
| M19 | | 83*0*0* | | 52* 0* 2* | 65.46* 1* 3* | | | 34.2 | 40.4 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| |  | | | | +0.55 | | | +0.5 | +1.5 |
| M20 |  | 85*0*0* | | 53* 0* 2* | 66.01* 1* 3* | | | 34.7 | 41.9 |
| | | | | | +2.08 | | | +1.4 | +2.6 |
| M21 |  | 87*0*0* | | 55* 0* 2* | 68.10* 1* 3* | | | 36.1 | 44.5 |
| | | | | | +1.45 | | | +1.1 | +2.3 |
| M22 |  | 89*0*0* | | 56* 1* 2* | 69.55* 0* 3* | | | 37.2 | 46.8 |
| | | | | | +0.95 | | | +1.3 | +2.7 |
| M23 |  | 91*0*0* | | 58* 1* 2* | 70.49* 1* 3* | | | 38.5 | 49.5 |
| | | | | | +1.59 | | | +1.2 | +2.8 |
| M24 |  | 93*0*0* | | 59* 0* 2* | 72.08* 1* 3* | | | 39.7 | 52.3 |
| | | | | | +1.71 | | | +1.6 | +3.2 |
| M25 |  | 95*0*0* | | 61* 1* 3* | 73.79* 0* 3* | | | 41.3 | 55.5 |
| | | | | | -0.17 | | | +2.3 | +3.5 |
| M26 |  | 97*0*0* | | 63* 1* 3* | 73.61* 1* 3* | | | 43.6 | 59.0 |
| | | | | | +2.66 | | | +1.0 | +2.8 |
| M27 |  | 99*0*0* | | 64* 1* 3* | 76.28* 0* 4* | | | 44.6 | 61.8 |
| | | | | | -0.95 | | | +1.3 | +0.8 |
| M28 |  | 100*0*0* | | 65* 1* 3* | 75.33* 0* 4* | | | 45.9 | 62.6 |
| | | | | | | | | - | - |

*** Refers to measurements taken at point 3 as indicated on FIG 2.*

*FIG9: Details of all the scenes available for the experiment*



Date:
Participant number:
Start time:
Completion time:
Illuminance:

## Experiment on brightness perception

Please complete the personal details section and read the experiment instructions below.

**Personal details**

1. Age:
2. Gender:
3. Have you completed formal education in any of the following? (tick as appropriate)
   - lighting ○
   - interior design ○
   - architecture ○
   - none of these ○
4. Which of the below best describes your vision? (tick as appropriate)
   - normal ○
   - corrected ○
   - abnormal ○ please specify: _____
   - colour blind ○ please specify: _____

**Experiment instructions:**

You will be shown a series of scenes of a waiting room, on the screen in front of you, one after the other in pairs. Each image will be shown for 4 seconds. You are asked to compare which of the two rooms appears to have more light, room 1 or room 2 and as the second image disappears from the screen say '1' if the room in the first scene has more light or '2' if the room in the second scene has more light. Your choice is to be based on the appearance of the whole scene. Responses will be recorded by the experimenter.

Before each pair of images is shown you will be notified to look at the screen. Please note that it will not be possible to view the scenes a second time before your selection.

Prior to the start of the experiment a trial will be performed, where you can ask questions. If your eyes feel tired at any time during the session, please let the experimenter know and there will be a short pause.

Thank you for taking part in this study.

**CONSENT**

I consent to take part in this experiment on brightness perception and I know that I can leave at any point.

_____ (signature)

_____ (name in block capitals)

_ th July 2016

APPENDIX C

## Formulae

| | | |
|---|---|---|
| $X = \sum \varphi\,(\lambda)\bar{x}(\lambda)$ | | [1] |
| $Y = \sum \varphi\,(\lambda)\bar{y}(\lambda)$ | | [2] |
| $Z = \sum \varphi\,(\lambda)\bar{z}(\lambda)$ | | [3] |
| $L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16$ | | [4] |
| $a^* = 500\left[\left(\frac{X}{X_n}\right)^{\frac{1}{3}} + \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right]$ | | [5] |
| $b^* = 200\left[\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} + \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}\right]$ | | [6] |
| $C_{ab} = \sqrt{(a^{*2} + b^{*2})}$ | | [7] |
| $h_{ab} = \tan^{-1}\left(\frac{a^*}{b^*}\right)$ | | [8] |