Predicting window view preferences using the environmental information criteria

Abstract

Daylighting standards provide an assessment method that can be used to evaluate the quality of window views. As part of this evaluation process, designers must achieve five environmental information criteria (location, time, weather, nature, and people) to obtain an excellent view. To the best of our knowledge, these criteria have not yet been verified and their scientific validity remains conjectural. In a two-stage experiment, a total of 451 persons evaluated six window view images. Using machine learning models, we found that the five criteria could provide accurate predictions for window view preferences. When one view was largely preferred over the other, the accuracy of decision tree models ranged from 83% to 90%. For smaller differences in preference, the accuracy was 67%. As ratings given to the five criteria increased, so did evaluations for psychological restoration and positive affect. Although causation was not established, the role of most environmental information criteria was important for predicting window view preferences, with nature generally outweighed the others. We recommend the use of the environmental information criteria in practice, but suggest some alterations to these standards to emphasize the importance of nature within window view design. Instead of only supporting high-quality views, nature should be promoted across all thresholds dictating view quality.

Keywords: Window view; Design; Standards; Machine learning; Nature

This is an Accepted Manuscript of an article published by Taylor & Francis in LEUKOS on 26th July 2022, available at: https://www.tandfonline.com/doi/10.1080/15502724.2022.2077753

1. Introduction

Views from windows are an essential element in any architectural design that draw daylight in and allow visual content to be seen out of the building (Tregenza and Wilson 2013). With the intended purpose of providing quality views, daylighting standards (SLL 2014; EN 17037 2018) provide a list of visual characteristics (Table 1). When these are present, designers can use them to assess the quality of the window view. Table 1 presents recommendations found in both the EN 17037 (EN 17037 2018), and Society of Light and Lighting (SLL) Guide 10 (SLL 2014). We have separated their recommendations into two sections: (a) visual features characterizing the view (e.g., horizontal layers (i.e. ground, landscape and sky), and distance of content); and (b) environmental information criteria. The latter section contains five different categories describing different facets of the view, which allegedly contribute to its overall quality.

Table 1. View quality assessment criteria used by the EN 17037 (EN 17037 2018) and SLL (SLL 2014) presenting: (a) the visual characteristics that need to be present within the view to be awarded a certain level of assessment, and (b) the environmental information criteria that can be used to determine whether other important features in the view are present

(a) Visual features

View quality assessment		**Horizontal layers**		**Content distance (m)**	
EN 17037	*SLL LG10*	*EN 17037*	*SLL LG10*	*EN 17037*	*SLL LG10*
-	Insufficient	-	Only foreground or sky	-	<6
Minimum	Sufficient	At least landscape	Landscape and another layout	≥ 6	≥ 14
Medium	Good	Landscape and another layer		≥ 20	≥ 28
High	Excellent	All three layers (foreground, landscape and sky)		≥ 50	≥ 54

(b) Environmental information criteria

Criteria	*Insufficient*	*Sufficient (Minimum)*	*Good (Medium)*	*Excellent (High)*
Location	-	□	□	□
Time	□	□	□	□
Weather	□	□	□	□
Nature	-	-	□*	□
People	-	-	*	□

*Either “nature” or “people” required for “good” to be achieved

Note for (b): “Insufficient”, “Sufficient”, “Good”, and “Excellent” are criteria used to denote window view quality in the SLL Lighting Guide 10, while “Minimum”, “Medium”, and “High” are used in the EN 17037.

The scientific evidence upon which the recommendations found in standards (SLL 2014; EN 17037 2018) were developed are not explicitly reported. Some criteria may have been motivated by scientific literature, while others could have been based on the professional judgment from the committee members. Although their origins cannot be determined, independent studies that exemplify their practical utility, but were not necessarily the basis for them, were identified in our literature review.

Horizontal stratification describes the separation of view content into a maximum of three distinct layers (Markus 1967). To be eligible for high (or "Excellent") assessments, views must contain layers showing the foreground, landscape, and sky (Table 1a). Layers containing more sky have shown to better support psychological restoration when the view was located within an urban (built) landscape (Masoudinejad and Hartig 2018). Another feature is content distance that evaluates how far away content is located relative to the window. Views containing distant content lead to higher assessments and our previous work (Kent and Schiavon 2020) indeed showed that as content distance increased, so did visual satisfaction. A third criterion recommended in standards (SLL 2014; EN 17037 2018) is horizontal view angle. This is used to inform the size of the openings, determining how much access occupants have to outdoor view content. Although access is a salient facet for overall window view quality (Li and Samuelson 2020; Ko et al. 2022), in our work, horizontal view angle does not play a central role and therefore, was not included in Table 1. Table 1b lists additional criteria that broadly relate to environmental features of the view, namely: location, time, weather, nature, and people. The addition of a specific set of one or more criteria dictates the overall assessment level (e.g., the presence or ability to determine the location, time and weather, equate to a "sufficient" view), while the inclusion of more conditions (e.g., people and nature) are needed to achieve the higher ratings. The relative importance of the former could be traced to research findings showing that discernment of temporal information (e.g., time of day) was the most influential parameter influencing preferred window size (Butler and Biner 1989), whereas determining the weather and views showing people revealed significant, yet weaker, relationships. Predictions were derived from ratings for window size averaged across 14 different spaces, including, but were not limited to, lecture halls, residential rooms, and offices.

To achieve the highest assessments, the inclusion of nature (e.g., trees, plants, and other sources of greenery) is required. Amongst the five environmental criteria, this criterion has been documented widely (e.g., (Ulrich 1981; Kaplan 1993; Tennessen and Cimprich 1995; Kaplan 2001; Aries et al. 2010), whereby a literature review by Velarde *et al.* (Velarde et al. 2007) documented the many benefits nature brings and which features (e.g. fields, green vegetation, forests) were seminal for each reported health effect (e.g. increased parasympathetic responses). However, exposure to nature may not invariably guarantee the same or even any positive effects. If opportunities for regular contact are limited or spaces containing nearby nature fail to meet certain expectations, preferences toward nature will inevitably vary (Hadavi et al. 2015).

The environmental criteria appear as mutually inclusive binary options, indicating the presence or absence of any given criterion. This simplification formats the five criteria into a checklist, but does not address the granularity to which each need to be measured to verify performance. For example, standards do not indicate how accurate estimates for "time" need to be (e.g., minutes or hours), nor how measurements should be verified. Determining the precise time of day (i.e. within an hour interval) from changes in correlation color temperature and luminance within daylit views has yielded inaccurate temporal estimates (Granzier and Valsecchi 2014). Although observers were unable to correctly determine the time of day, the authors did not completely disregard daylight as a reliable temporal indicator. Observers were not aware of the view's location and often incorrectly interpreted visual cues (e.g., shadows) that helped orientate the position of the Sun to reveal the correct time of day. While the abovementioned showed that location facilitated time of day, dependencies across other environmental criteria have also been shown. Other than eliciting positive emotional responses, time was perceived as passing slower when images contained nature (Davydenko and Peetz 2017). Time is also an important proxy for location, whereby changes in foliage due to circannual effects (e.g., from fall to winter) can cause changes in mood (Brooks et al. 2017). The myriad linkages amongst these criteria could be complex but as of yet, no studies – to the best of our knowledge – have systematically evaluated their effects on window view.

Waczynska *et al.* (Waczynska et al. 2021) had shown that three criteria in daylighting standards (SLL 2014; EN 17037 2018) (i.e., relative window size, content distance, and number of horizontal layers) were unable to accurately characterize subjective ratings of view quality from 169 observers, but their work had some limitations. Low accuracy was defined by statistically significant differences that were found between estimated values of view quality, calculated from criteria recommended in these standards, and subjective ratings. While this raises questions to whether these criteria should be used in practice, not all the criteria in these standards were used, and the incorrect inferential analyses were used to gauge the overall predictive capacity of those that were. This warrants a more systemic evaluation of these recommendations, particularly the environmental information criteria. With the rise in studies measuring data that are richer both in size and complexity (e.g. number of measured variables), increased attention has been placed on analytical techniques that generally offer more accurate predictions using machine learning algorithms than conventional statistical (e.g. regression) models (Bzdok et al. 2018). Studies for indoor environmental quality (Graham et al. 2021; Kent et al. 2021), thermal comfort (Kim et al. 2018; Cheung et al. 2019), and window view design (Kim et al. 2022) are among those which have applied machine-learning algorithms, testing the predictive limits from data that sought to understand occupant satisfaction. To determine the predictive capacity of the environmental information criteria, machine learning algorithms were used to verify their ability to predict window view preferences in our current study.

In our study, we aimed to determine if and how well the environmental information criteria can be used to assess window view preferences. We hypothesized that when window view preferences were larger (i.e. one view was greatly preferred over another), prediction accuracy would be higher. We evaluated the importance of each criterion in Table 1b when used to classify different views, and we also measured their relative impacts on psychological restoration and affect. In doing so, we provided informed recommendations that indicate whether these criteria should be used in mainstream daylighting standards.

2. Method

Our method utilized two stages (Fig. 1). In stage one, we asked a small group of participants (n= 30) to indicate which set of views they preferred. We assumed that visual content does not always dictate overall window view preferences, since preference can be sensitive to changes in context and experience (Warren et al. 2011). Preconceptions over certain visual features for one window view (e.g. a nearby open field) could create diverging responses (e.g. good or poor privacy) due to prior experiences elicited by window views with similar content, but are contextually different: for example, one was a busy school field, while the other was empty private land. Preferences toward window views may also be influenced by the inherent function of the space, altering the expectations of the occupants toward the window and its view (Dogrusoy and Tureyen 2007).

Rather than inferring preferences directly from visual content (e.g. views of nature are automatically preferred over those that may have satisfying urban features), we affirmed differences in preference before proceeding to the second stage. Using a much larger group of participants (n= 421) and a broader range of survey questions, the same views were evaluated in more granular detail within stage two. Views were evaluated to determine if the environmental criteria could accurately predict different views, according to their differences in preference determined in stage one. More details are reported below.

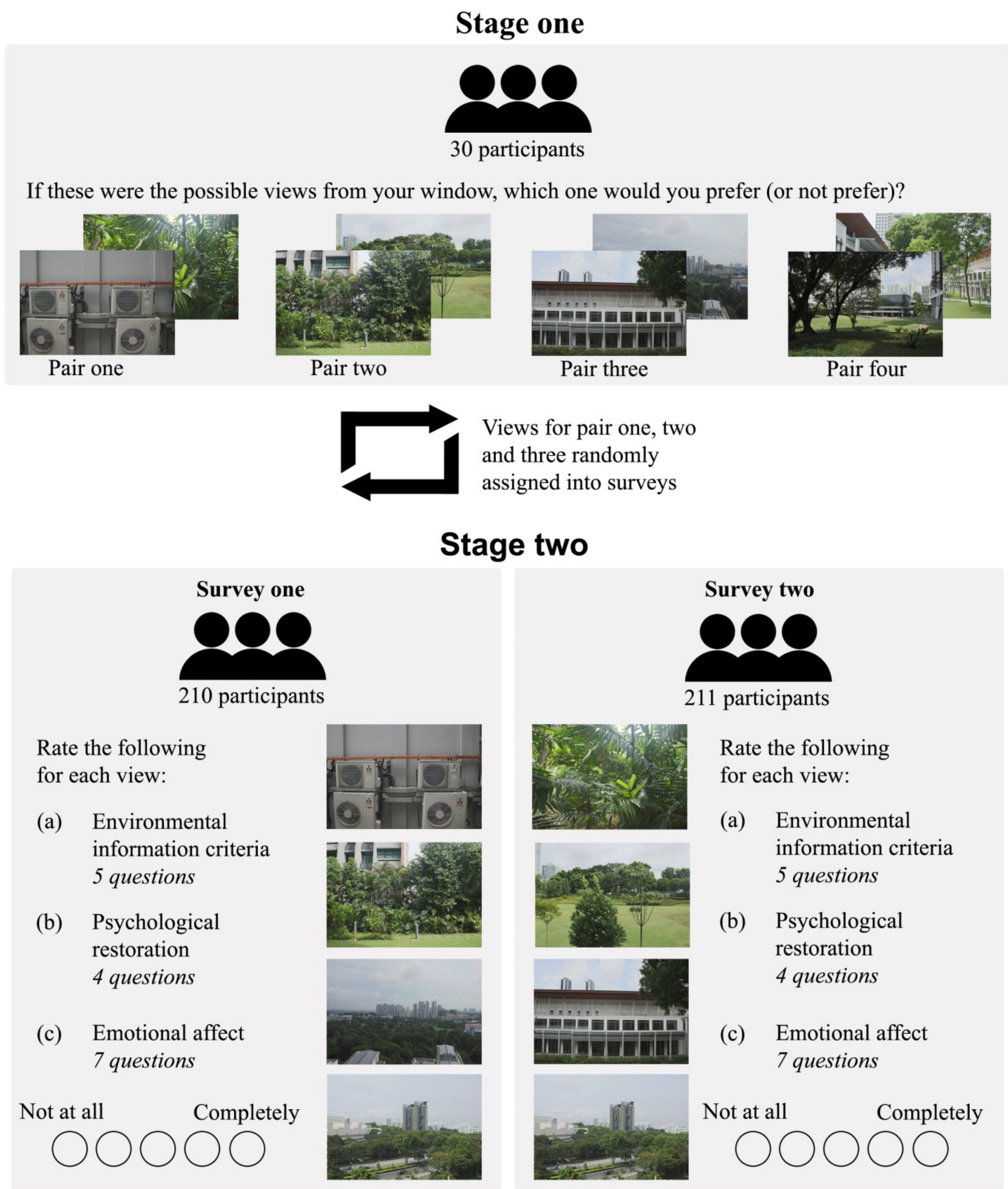

Fig. 1. A schematic overview of our research study method that composes of two stages. In each stage, we showed the surveys and the questionnaire items, the windows views, and the total number of participants that took part and gave their subjective ratings

2.1. Window view images

We used images to represent different views, which is a common method found in several studies (Tuaycharoen and Tregenza 2005; Brooks et al. 2017; Masoudinejad and Hartig 2018; Kent and Schiavon 2020). In total, nine different images (Kent and Schiavon 2020) were taken at the National University of Singapore campus on the same day from 09:00 am to 11:00 am. During this period, there was an intermediate sky and direct sunlight was avoided

 https://doi.org/10.1080/15502724.2022.2077753

when capturing the images. We used a Canon 7OD with an efs 10-55 mm f/3.5-5.6 STM lens. Eight of the views shown in the images were taken from actual windows. Since no other window with ideal content could be found, one view in pair 4 with a landscape distance of 64 m (Fig. 2) was not taken from a window, but was included for comparative purposes.

View selection was generally based on features recommended in Table 1a belonging to daylighting standards (SLL 2014; EN 17037 2018), considering views that had a different number of horizontal layers and varying content distance. Following the procedure outlined in the SLL Guide 10 (SLL 2014), we were able to identified the number of horizontal layers for each view. Presence of nature and urban content was also used to discern views, which took into account the vast amount of greenery abundant and often integrated within urbanized spaces in Singapore (Henderson 2013). Organized into four pairs, Fig. 2 presents views that were compared against each other in stage one. Our previous study (Kent and Schiavon 2020) showed that visual content ratings given to views in the first three pairs spanned across the full spectrum of satisfaction semantics anchored on a continuous scale, ranging from "very satisfied" to "very dissatisfied". Although visual content may not solely affect overall preference, the diverse content seen in these views was ideal for the purposes of our current work (i.e., they encapsulated view content quality varying from high- to low-end, and views within these extremes). When ordering the six views according to visual content ratings (see Appendix A, Fig. A1), images were systematically paired based on both differences in satisfaction and visual content (i.e. horizontal layers and content distance). We assumed this might also maximize the differences in preference across each pair of window views.

In our current study, pair one contained a landscape only with nature in one view (trees) and urban (air-conditioning condensers) in the other. The labeling process for images helped distinguish views and not to positively or negatively connote preferences (e.g., the urban view overtly does not represent a wider range of urban views containing more sophisticated architecture). Pairs two and three had three layers each, but content distance varied. Finally, pair four had the same number of layers and content (mixture of nature and urban) in both images. Across the first three pairs, we anticipated that preferences would generally be higher for one view. But the magnitude of these differences in preference was equivocal. Due to similarities in the view content (e.g. trees, buildings) for pair four, these were designated as a null condition (i.e. anticipated differences in preference were minimal) (Fotios 2019). Unanticipated differences can be used to diagnose extraneous issues (e.g., the view pairings did not accurately reflect the actual preferences given by participants).

The first three pairs would later be included in stage two and an additional view was also used. Therefore, each survey contained four window view images and in total, both surveys had seven. One image originated from the first three pairs used in stage one, excluding the window views used in pair four. The previously unused null condition window view (Fig. 2) formed the fourth image, which was used in both surveys. Both surveys featured the same set of questions intended for two independent participant groups. The additional fourth image, appearing in both surveys, served as a null condition (Fotios 2019) and we expected that similar ratings would be given to the exact same view for each question.

 https://doi.org/10.1080/15502724.2022.2077753

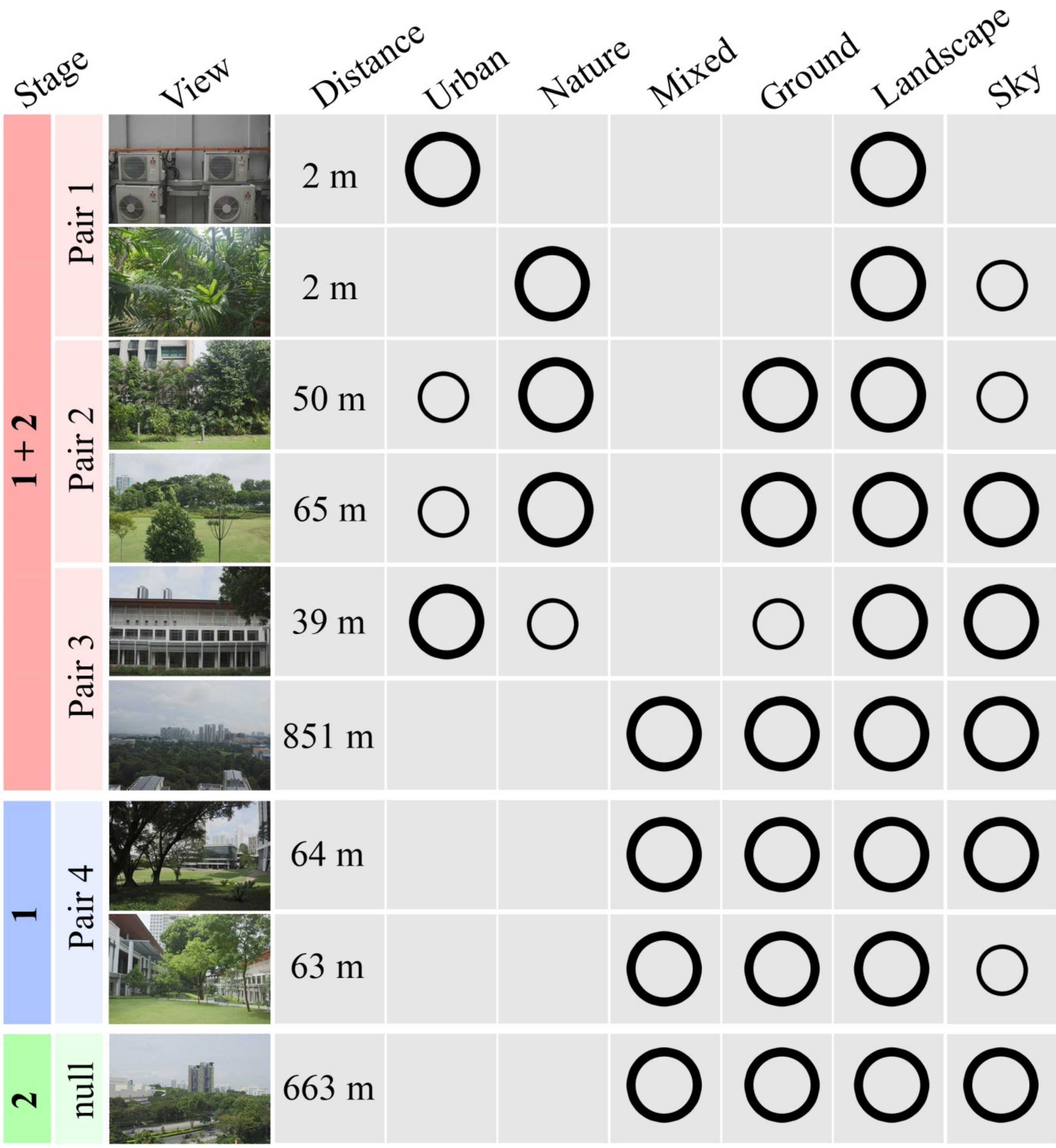

Fig. 2. The window view images utilized in our study are presented in a matrix-style table organized according to: the stage (one or two) they were implemented, and the pairings or null condition. For each window view image, the table shows the calculated landscape distance, whether the image contained nature or urban features or had mixed (both), and the identifiable ground, landscape, and sky layers. Note: prominent features (e.g. a large visible sky layer) are denoted by large circles, while less conspicuous features (e.g. a small visible sky layer) have smaller circles

2.2. Procedure

Stages one and two were both conducted using online surveys distributed through the UC Berkeley Qualtrics platform. Procedures were approved by the institutional review board (CPHS: 2020-04-13235).

Image pairs were evaluated using side-by-side comparisons (Tuaycharoen and Tregenza 2005; Salesses et al. 2013), requiring participants to select which they would prefer to have as their window view. Only one pairing appeared at a time and once a response was given, the next pair appeared. The order that the four image pairs (Fig. 2) appeared was randomized (Field and Hole 2011). To counterbalance any ordering effects associated with the small number of pairs available (Krosnick and Alwin 1987; Poulton 1989), evaluations were repeated once. In other words, once the four pairs had been evaluated, evaluations made to the same pairings were repeated again in another order. Since it was conceivable that participants may not always prefer either view for any given pair, a "neither" criterion was available to avoid forcing them to indicate their preference. Stage one took approximately 3-minutes to complete.

In stage two, two different online surveys were used to evaluate the seven images. A view from each pair (Fig. 2) was included in one survey and the remaining three images were included in the other. Besides the null condition that was included in both, the allocation of the three images into each survey was random. The surveys were sent worldwide, inviting anybody above the age of 18-years to take part. In the survey prelude, participants were requested to select one, but not both surveys and also to read the additional instructions before consenting to take part. Participants rated each view separately. To frame context around their responses (Tuaycharoen and Tregenza 2005), the following was provided before participants were asked to answer the 15 questions (Masoudinejad and Hartig 2018):

> "*Imagine that you are at work, you are mentally tired and in need of rest. To help clear your thoughts before carrying on your daily tasks, you have some time to sit down and during that period you look outside the window and this is your view.*"

All ratings were given on a 5-point unipolar scale ranging from "Not at all" to "Completely" (Masoudinejad and Hartig 2018). Questions were phrased so that the same semantics could be consistently applied. The surveys consisted of three groups of questions that evaluated different dimensions of the view (Appendix B, Fig.B1). Five questions were the environmental information criteria (Table 1b). In three of the five questions, participants assessed whether they were able to determine: the general "location" (e.g., nearby the city-center, suburbs, forest, etc.), the approximate "time of day" (e.g., early afternoon) without the use of a clock or watch, and the "weather" condition. Due to the binary assessment method used in standards (Table 1b), it was difficult to generate granular measurements for "people" and "nature" without modifying the original semantics. In other words, views containing very little (e.g., single potted plant) or an abundance (e.g., forest) of greenery both meet the criterion for nature. Similarly, images showing views of sidewalks absent of persons won't satisfy the criterion for people, but obviously have the capacity to meet this requirement. To avoid nonsensical questions, the criterion "people" was changed to "movement", requiring participants to rate potential changes occurring from anthropogenic examples (e.g., traffic or people walking). Presence of "nature" was modified to "how connected they felt to nature" (Mayer et al. 2009).

Four questions were adapted from previous studies that had evaluated psychological restoration from window view images (Masoudinejad and Hartig 2018), variations in architectural design (Lindal and Hartig 2013), and urban street vegetation (Lindal and Hartig

2015): namely, restoration, fascination, being-away, and preference. For the fourth item (i.e. preference), respondents were asked to evaluate the following: “I like this window view”; this question being distinctly different from preference ratings provided in stage one. Finally, seven questions (i.e. happy, strained for time, relaxed, sad, stressed, and busy) measured different aspects of affect. These questions were featured in a survey used to determine differences in emotional response when images containing either nature or urban content were compared (Davydenko and Peetz 2017). Once participants had finished providing all responses for one view, they provided the same responses for the remaining six images that appeared in a randomized order. Participation was strictly voluntary and involvement was not remunerated. Stage two took approximately 7-minutes to complete.

2.3. Participants

In stage one, participants residing and working in the same office building in Singapore took part. While we could not record cultural background in detail, participants that took part in stage one were from and/or had lived in Asia, Europe, and America. Participants were affiliated to our research institution. Some were research staff and others were not (e.g. administration and technical personnel). Although work profession varied, participants were unaware of the study objectives, having no expertise in (day)lighting and view research, design or practice. Our aims and hypotheses were also concealed from participants to blind them from any expected outcomes.

Table 2. Summary of the demographics recorded across the preference survey in stage one, and two surveys in stage two, showing age, gender, location, and total number of participants

Demographics	Value	Stage 1	Stage 2	
		Preference survey	*Survey 1*	*Survey 2*
Age	Mean (SD)	36 (7.47)	36 (11)	37 (12)
Gender (%)	Male	23	45	47
	Female	7	43	47
	Not specified	-	11	4
	Other	-	1	1
Location (%)	Africa	-	1	1
	Asia		26	23
	Europe		43	38
	North America		23	31
	South America		2	2
	Oceania		4	3
	Other		1	2
Total	*N*	30	210	211

In stage two, 550 participants had consented to take part in the experiment. Since it is common that some participants may consent, but then opt out before providing any responses or completing all questions in survey research (Brick and Kalton 1996), missing data needed to be excluded. The resultant dataset contained 421 participant responses. A summary of the demographic features recorded from participants is shown in Table 2. Across the surveys, the total number and breakdown of each demographic feature are relatively similar.

2.4. Statistical analyses

To analyze the data in stage one, we used the chi-squared (χ^2) test (Field et al. 2012) to compare the difference in view preference across each pair of images. Since the aim was to identify what views were preferred, "neither" (i.e. no preference) responses were excluded. We used a threshold of $p \leq 0.005$ to declare differences that were statistically significant and calculated the effect size (r) to estimate their magnitude. Our threshold is more stringent than the typical $p \leq 0.05$ in order to increase the reproducibility of scientific results (Benjamin et al. 2018). To interpret the latter, three thresholds denoting practical significance (i.e. small, moderate, and large: $r > 0.20$, 0.50, and 0.80, respectively (Ferguson 2009)) were utilized.

For the views utilized in stage two, different supervised machine learning algorithms were used. Using decision trees (Breiman et al. 1984), random forest (Breiman 2001), and neural networks (McCulloch and Pitts 1943), we determined how well the five environmental information criteria (Table 1b) could classify each of the three pairs of window view images (Fig. 2). All three approaches are appropriate when classifying binary outcomes using several predictor variables (Kotsiantis 2007; Statnikov et al. 2008). Decision trees are generally easy to understand models, while random forest and neural networks are more complex approaches (Guidotti et al. 2019). The application of supervised machine learning algorithms requires training an algorithm (e.g. decision tree) on a partitioned part of the available dataset, and then test the performance of the trained model onto the remaining unseen part of the dataset (Kotsiantis 2007). Since the nature of the data (e.g. its size, number of variables, expected outcome, etc.) plays a significant role on learning outcomes for different algorithms (Sarker 2021), we wanted to know which of the three selected models would yield the highest classification accuracy.

We independently compared the four pairs of images and disregarded the use of one single model considering all views. This was to avoid comparing views where differences in preferences had not been affirmed, but were likely minimal in some cases. Classification of these views would yield poor measures of accuracy, which would be aggregated with higher accuracy measures from views with different content and preferences, leading to an unreliable measure of overall performance. This applies especially to the models' capacity to predict images used in the null condition, whereby we anticipated low measures of performance when classifying two identical and equally preferable views.

We partitioned the data into training and test datasets using a 7:3 allocation ratio (Dobbin and Simon 2011). The training dataset was used to determine optimal parameter settings for each model (Boser et al. 1992). A grid-search approach (Hsu et al. 2010) across a wide range of relevant parameters was used. Repeated k-fold ($k = 5$) cross-validation was performed during the training process (Raschka 2020). To minimize overfitting, the decision tree model was pruned (Mingers 1986; Bohanec and Bratko 1994). The performance of each model was verified using the receiving operating characteristic (ROC) curve (Metz 1978; Hanley and McNeil 1982), accuracy, and F-measure (Sokolova et al. 2006).

The association between the five environmental information criteria with measures of psychological restoration and affect were evaluated using heatmaps. We also used k-means clustering (Hartigan and Wong 1979) to examine the relationship between the average (median) values in each matrix. This is an unsupervised technique that can be used to provide further insights once data has been summarized or reduced into a smaller subset (Ding and He 2004). To determine the optimal number of clusters, we used both the elbow and silhouette approaches (Yuan and Yang 2019).

3. Results

3.1. Stage one

Figure 3 presents the results in stage 1. For pairs one (a), two (b), and three (c), the differences in preference are both statistically ($p \leq 0.005$) and practically ($r \geq 0.20$) significant. The effect sizes for pairs one and three are consistently "large" while for pair two, these differences were "small". Nature and distant content are generally preferred over urban and nearby content. Interestingly, nature nearby had higher ratings of visual content than nature distant (Appendix A). This helped solidify the notion that visual content alone does not always determine general preferences. For pair four, the differences are neither statistically or practically significant. The statistical results supported the intended role served by the null condition, revealing no prevailing preferences across different views containing similar visual content. Preferences across the two (i.e. first and second) sessions did not change when participants viewed pairs one and three, but varied slightly for pairs two and four. This may have been due to indecision felt when evaluating equally preferable views.

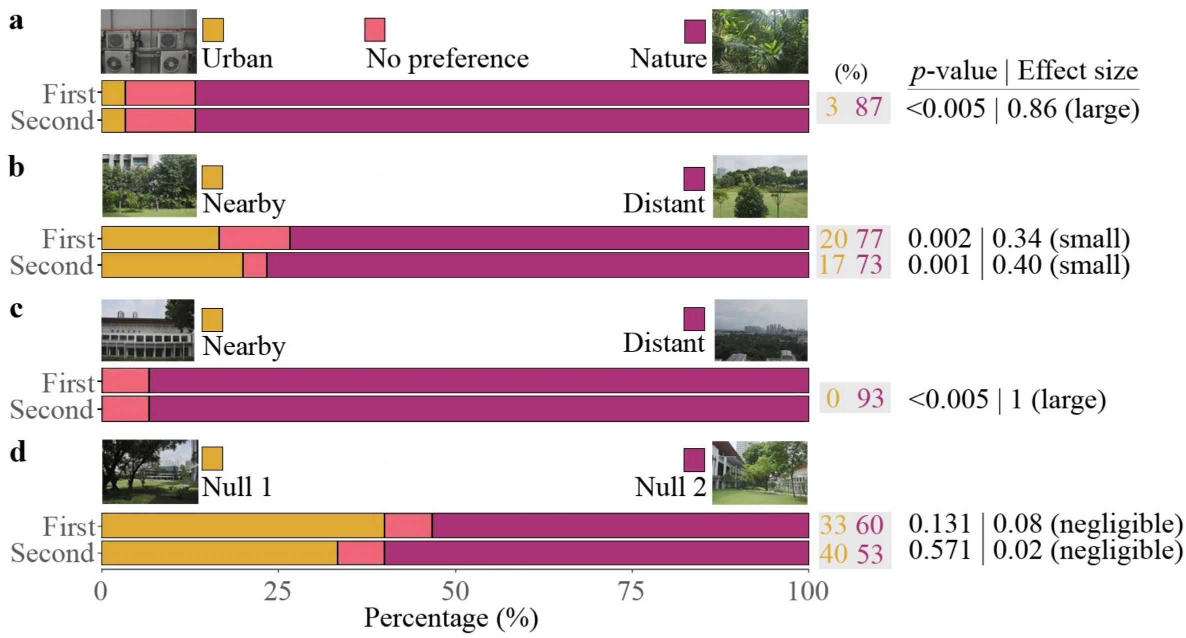

Fig. 3. Percentage plots showing the results of the first and second blocks when evaluating the side-by-side comparisons for: (a) pair one (urban vs. nature), (b) pair two (nearby nature vs. distant nature), (c) pair three (nearby urban vs. distant urban), and (d) pair four (null 1 vs. null 2). Note 1: the percentages show final preference made to each window view image (excluding "no preference" criterion). Note 2: unless percentages were the same for each session, the results of the χ^2 test (p-value and effect size (r)) are shown

3.2. Stage two

 https://doi.org/10.1080/15502724.2022.2077753

The performance for each machine learning model is shown in Table 3. Summary statistics presenting the average (median) and extent to which ratings varied for the views used in stage two can be found in Appendix C. Decision tree consistently outperforms random forest (i.e., more views are correctly classified). Compared to neural networks, decision tree has relatively similar performance measures. Since decision trees are generally easier to interpret than neural networks, this model was used in further analyses. The neural network structures for all three pairs are shown in Appendix D. When preferences favored one view over the other (i.e. pairs one and three – as shown in Fig. 3), classification accuracy ranged from 83% to 90%. For a small difference in preference (i.e. pair two), the model performance was at its lowest (67%). This seems to indicate that when one view is more preferable over its counterpart, the decision tree model has less difficulty classifying the pair of images using the environmental information criteria.

When comparing the exact same view (i.e., the null condition (pair four)), prediction accuracy across all indicators revealed poor measures of performance. The ROC values across the three models approximated 0.50, indicating that neither model could discriminate one view from the other. This is overtly facilitated by the almost indistinguishable ratings (Appendix C) given to the environmental criteria used to describe the differences between two identical views. Although Table 2 shows that demographic features were relatively similar across the two surveys, these results indicate there were no reasons to suspect that inter-individual differences or the presence of an ordering effect (i.e., high or low preference views in either survey influencing subsequent evaluations given to the next view) were present across, or in either survey. Since the null conditions fulfilled their intended purpose, further analyses pertaining to these views have not been reported.

Table 3. Performance of the decision tree, random forest, and neural network models showing the ROC values with its associated lower and upper 95 % confidence intervals, accuracy, and *F*-score. This can be used to determine how well the models performed when they were applied to the test dataset containing image pairs one, two, and three

Model	**Images**	**ROC**			**Accuracy**	***F*-score**
		Value	*Lower*	*Upper*		
Decision tree	Pair One	0.90	0.85	0.96	0.90	0.90
	Pair Two	0.67	0.58	0.75	0.67	0.69
	Pair Three	0.83	0.76	0.89	0.83	0.82
	Pair Four	0.53	0.44	0.62	0.53	0.53
Random forest	Pair One	0.87	0.81	0.93	0.87	0.87
	Pair Two	0.61	0.53	0.70	0.62	0.62
	Pair Three	0.81	0.74	0.88	0.81	0.82
	Pair Four	0.46	0.38	0.55	0.46	0.46
Neural networks	Pair One	0.90	0.85	0.96	0.85	0.86
	Pair Two	0.70	0.61	0.79	0.63	0.67
	Pair Three	0.89	0.83	0.95	0.85	0.85
	Pair Four	0.51	0.41	0.61	0.50	0.47

Figure 4 plots the pruned decision trees. Since nature in pair one and distant urban in pair three contained more greenery than their counterparts, the root node classifies a majority of the data by the criterion “connected to nature” (herein, connect nature). This criterion single-

 https://doi.org/10.1080/15502724.2022.2077753

handedly categorized both views in pair one. For pairs one and three, there is a higher probability that the terminal node corresponds to a high preference view (i.e. nature (pair one), and distance urban (pair three)) when “connect nature” was rated higher than its accompanying image pair. However, the decision rule for the same criterion varied across the two decision trees: for pair one this refers to “not at all” and pair three to “moderately”. This could be attributed to the presence of nature seen in both views in pair three, whereby nearby urban also contained greenery, albeit to a lesser extent than distant urban, while this feature is completely absent in urban and is saturated in pair one’s counterpart. Beneath the root node, pair three requires a secondary criterion to further classify the views. According to terminal node 3, a relatively large majority of data had a high probability of being classified as nearby urban when ratings for “location” were less than “completely”. This could be explained by the content distance. When considering there were fewer visual cues due the nearby proximity of content, it becomes more difficult to determine the location seen in view.

Pair two shows a more elaborate decision tree. Unlike the other decision trees, “location” is the root node. When rated lower (<moderately), there is a higher probability that the view contained nearby nature. This supports our inferences derived from pair three. When the ratings to this criterion and “weather” were high (i.e. ≥ moderately and completely, respectively), there is a higher probability that the view is distant nature. The relevance of “weather” is quite overt when considering the latter view contains more visible sky. Decision nodes containing “time” are also used in the classification process. Generally, there is a higher probability that the view is nearby nature when “time” was rated slightly higher than its counterpart. Its relevance within the classification process is somewhat unclear, but may be driven by psychological factors rather than purely visual reasons.

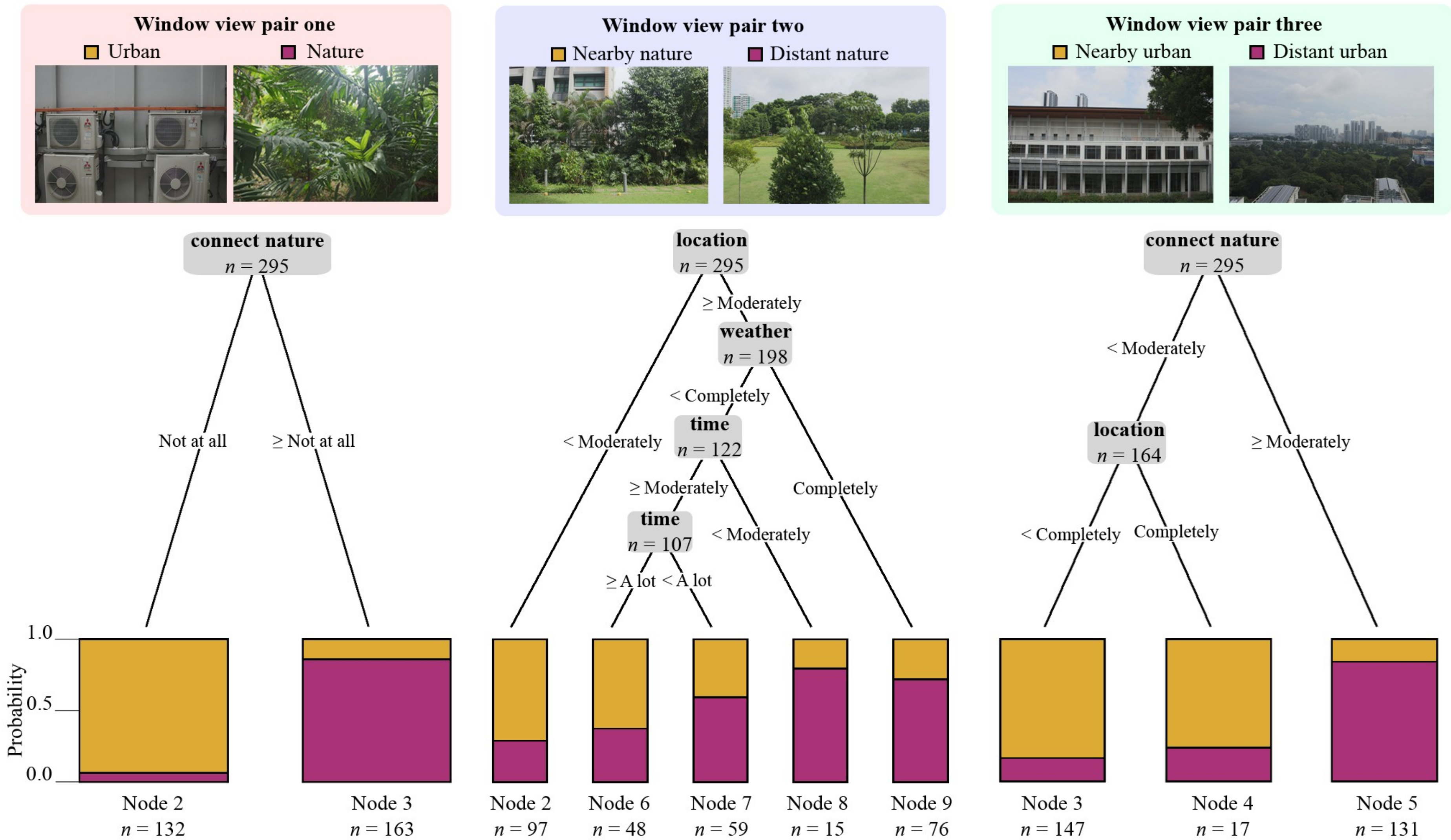

Fig. 4. Pruned decision trees used on the training dataset for window view image pairs one, two, and three. This shows the root (upper-most) node containing the entire training dataset, branches that contain rules based on the ratings given to the environmental information criteria and filter toward decision nodes, and terminal nodes showing probability plots classifying the window view image pairs

https://doi.org/10.1080/15502724.2022.2077753

Figure 5 shows the variable importance as determined by the *z*-scores from the decision trees. Although the height of each node on its classification tree (Fig. 4) also represents the relative importance for each environmental information criterion, where the root node contains the largest weighted importance, each tree had been pruned to reduce its complexity, making it difficult to discern this information. This process changed the hierarchical importance for some criteria and removed others from the classification tree. Even though some criteria did not appear on a classification tree (Fig. 4), Fig. 5 clearly shows they were still considered important. Figure 4a exemplifies this observation, showing only one criterion ("connect nature"), but Fig. 5a demonstrates that three additional criteria were consider important, albeit their weighted importance was much lower.

Variables denoted as unimportant depend on the views being compared. None of the environmental information criteria are deemed unimportant in pair two (b), while "time" in pair one (a), and "movement" in pair three (c) were. Despite its importance in pairs one and three, "connect nature" was unimportant when classifying views in pair two. This could be due to the abundance of greenery seen in both views (i.e. nature cancels out). The lack of visible paths (e.g., sidewalks) and roads across views in pair three explained why "movement" was unimportant, while the absence of dynamic cues usually present in the ground (e.g., people) and sky (e.g., weather) layers may explain the unimportance of "time" in pair one.

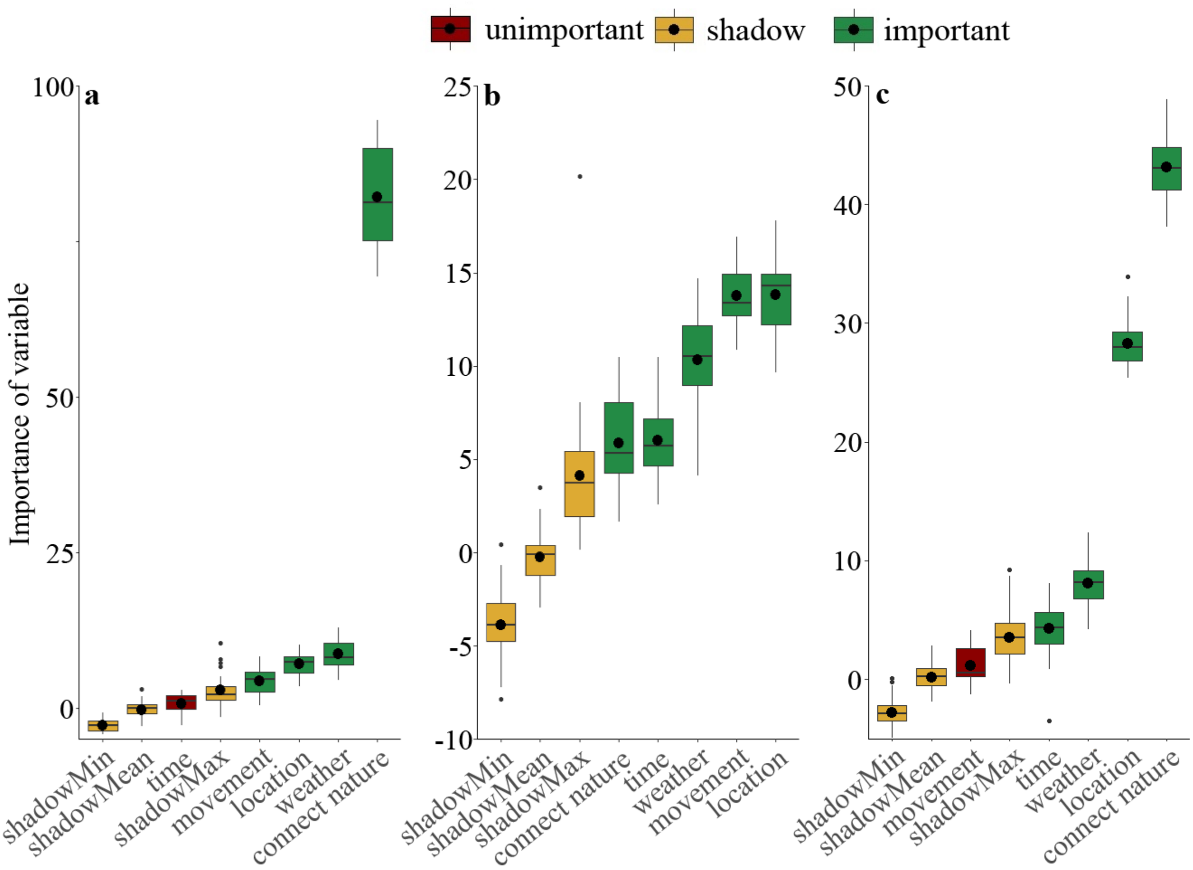

Fig. 5. Importance of each variable (expressed as the *z*-score) from each decision tree model used to classify: (a) pair one, (b) pair two, and (c) pair three. The results of the analysis were used to determine which variables were important and unimportant when they were used to classify the window view images. Note: shadow features (i.e. shadowMin, shadowMean, and shadowMax)

 https://doi.org/10.1080/15502724.2022.2077753

are randomly generated values created from the data. Real features (i.e. environmental criteria) are compared to shadow features to gauge their relative importance

Figure 6 presents a heatmap containing the median ratings for the three view pairs. The cluster analysis revealed two groups of views. There are two views (i.e. nearby urban and urban) in one cluster, and four (i.e. distant urban, nearby nature, distant nature, and nature) in the other. Although the 15 variables clustered into two groups, this was due to how questions were phrased to ensure the same scale semantics could be used (i.e., positive responses given to some variables (e.g., strained, busy, etc.), were inversely related to the magnitude rating on the scale. While for other variables (e.g., focus, content, etc.), they were linearly related). The exception to this being "movement", which generally wasn't rated highly on the 5-point scale for most views. As elucidated for Fig. 5, low evaluations for this criterion might reflect the lack of visual cues (e.g. sidewalks and roads) inherent within the views.

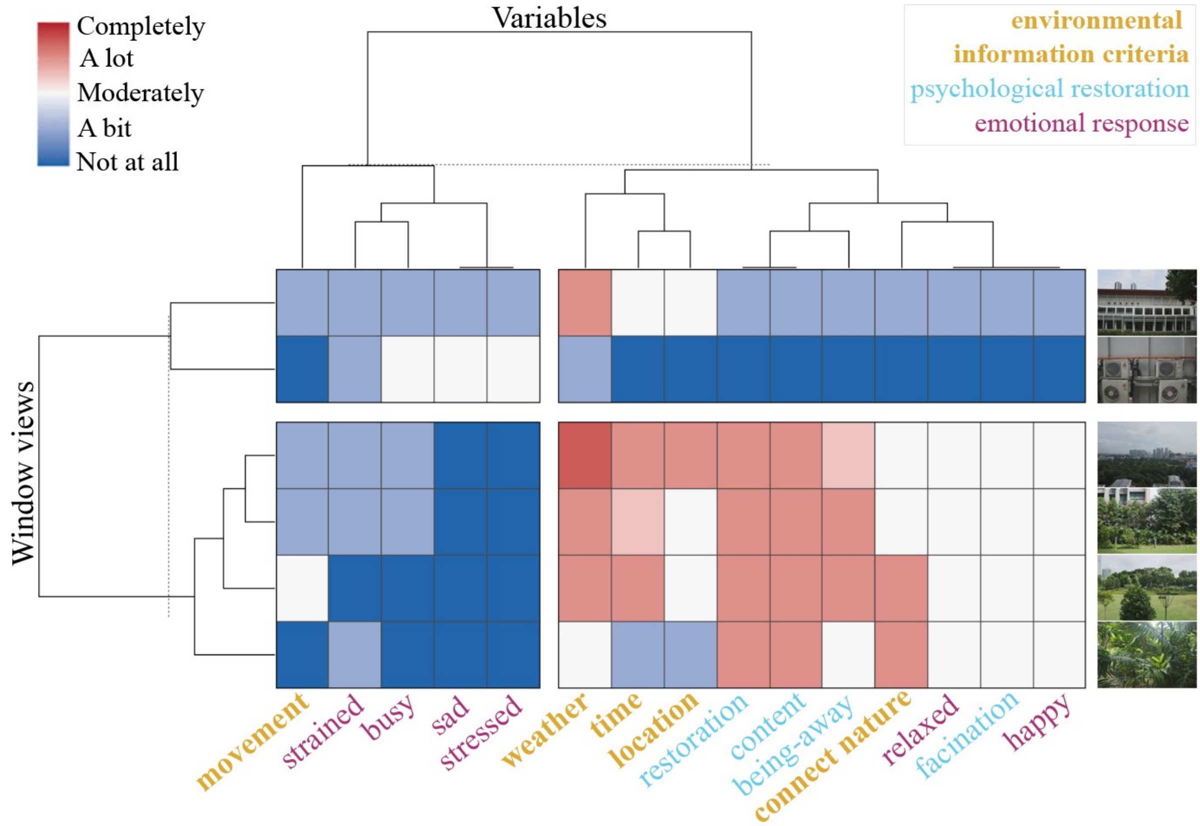

Fig. 6. Heatmap plotting the median value calculated from the 15 different variables given to the six different window view images. Two dendrograms are used to cluster the variables (top) and window views (left) into two groups based on values found in the matrix. The three different colors are used to denote the type of question: environmental information criteria, psychological restoration, or emotional response

Both nearby and distant nature had similar ratings. While distant nature was the more preferred view in pair two, both views can still be considered preferable. This can be explained by distant nature slightly outperforming nearby nature for some criteria ("movement", "time", and "connect nature"), yet for others ("location", and "weather") there were no differences. Interestingly, both these views, and distant urban and nature, all have the same average ratings for content. However, average ratings for other criteria (e.g., location)

 https://doi.org/10.1080/15502724.2022.2077753

and variables (e.g., being-away) varied. Therefore, providing views with high view content ratings may not guarantee success in other aspects (i.e. environmental information criteria and psychological restoration), or accurately reflect overall occupant preferences.

When ratings given to environmental information criteria were higher, positive evaluations (e.g., high restoration and low stress) were given to other variables. While this does not necessarily imply any causal relationships exist, views that score well according to the environmental information criteria are associated with higher psychological restoration and positive affect. For this relationship to occur, high ratings for certain criteria need to be achieved. Nearby urban showed that while it performed well for "weather" and moderately well for both "time" and "location", this did not leverage similar beneficial responses for the human-centered questions. The view of nature generally received high ratings for "connect nature", yet it still elicited somewhat positive responses that were not apparent for nearby urban. Interestingly, we were able to show that the wider impacts of the five environmental information criteria on human-centered responses varied. In other words, designing predominantly for "connect nature" ("nature" in standards) might equate or even outweigh the holistic benefits when trying to design for all five criteria.

4. Discussion

Using the five environmental information criteria (i.e. location, time, weather, nature, and people (movement)) used in daylighting standards (SLL 2014; EN 17037 2018), we tested their ability to classify three pairs of window views with different visual content and preferences. Classification accuracy was determined using three different supervised machine learning models: decision tree, random forest, and neural network. A forth pair containing two identical views of equal preference served as a null condition. When preferences were large (i.e. one view was preferred over the other), classification accuracy ranged from 82% to 90%. For a small difference in preference, the model accuracy was lower (67%). When preferences were indistinguishable, neither model considered could be used to accurately classify views used in the null condition. Feature selection analysis showed that the importance of each variable varied according to the views being compared. For some views, not all the criteria were needed in the classification process. Although this does not understate their general importance, criteria that dominate the classification process across some views (e.g., nature) may play a much smaller role in other situations.

To verify the credibility of the environmental information criteria, we used different machine learning algorithms. Although the application of machine learning algorithms is not necessarily new for occupant survey research, they have proven to produce high prediction accuracy for overall satisfaction ratings given to indoor environmental parameters using the world's largest post-occupancy database (Kent et al. 2021), and showcased the limited ability for existing metrics to predict thermal comfort (Kim et al. 2018).

When comparing the environmental information criteria to other human-centered variables, these appear to influence psychological restoration and positive affect. While we cannot conclude that any causal relationships existed nor can we rule these out, it does emphasis the need to design views using robust recommendations that are able to reinforce human health and well-being, whereby the link between the latter has been firmly established

 https://doi.org/10.1080/15502724.2022.2077753

(Aries et al. 2010; Veitch and Galasiu 2012). Considering that designers can only use these criteria to design for five facets of the window view, our work begins to demonstrate that they could have wider beneficial impacts on building occupants (e.g., reduced feelings of stress and elevated psychological restoration). Another previously unexplored facet warranting further study is the "disconnect" between perceived visual content and view preferences. Views with satisfying visual content (i.e. nearby nature) may not always cater the holistic requirements that underlie preferences. Therefore, occupants may prefer views with less satisfactory content (i.e. distant nature), since they better meet criteria (e.g., "movement" and "time) that influence other important dimensions of the window view.

We would also like to propose the following changes in current daylighting standards (SLL 2014; EN 17037 2018) to better reflect the findings derived in our work. One such alteration addresses the fact that "nature" is not a prerequisite that satisfies the minimum (i.e. "sufficient") assessment threshold. Compared to other criteria, "nature" appears to have a stronger association with psychological restoration and positive affect. In fact, satisfying the requirements for "nature" only may even bring greater benefits than designing for the four remaining criteria. While previous literature (e.g., (Ulrich 1981; Kaplan 1993; Tennessen and Cimprich 1995; Kaplan 2001)) have also highlighted the profound effects of "nature", it is unclear why it is only used to deliver high-end (i.e. "good" and "excellent") views and not to support views below these assessment thresholds.

Our proposed changes are outlined in Table 4. Besides "insufficient", we think "nature" should be a requirement in every other threshold that can be used to signify a view that meets and exceeds minimum requirements. This accommodates the linkages between "nature", psychological restoration and positive affect, which were more pronounced compared to the other four criteria. Beyond this, designers may select any of the other four criteria and for each threshold increase, an additional criterion is required to meet that level of view quality. The freedom over criteria selection may also help designers select a more appropriate combination of attributes, which better suit the available outdoor content. For example, a view of an outdoor public green-space may guarantee nature, location and people, but not necessarily time or weather. Therefore, the previous version (Table 1b) would denote the view as insufficient, but would meet sufficient assessment criteria in our new table.

Table 4. A revised version of the environmental information criteria

<table>
<tr><th colspan="5">Revised environmental information criteria</th></tr>
<tr><th>Criteria</th><th>Insufficient</th><th>Sufficient (Minimum)</th><th>Good (Medium)</th><th>Excellent (High)</th></tr>
<tr><td>Nature</td><td rowspan="5">Only one available</td><td>□</td><td>□</td><td>□</td></tr>
<tr><td>Location</td><td rowspan="4">Two needed + nature</td><td rowspan="4">Three needed + nature</td><td>□</td></tr>
<tr><td>Time</td><td>□</td></tr>
<tr><td>Weather</td><td>□</td></tr>
<tr><td>People</td><td>□</td></tr>
</table>

Despite both standards (i.e. (SLL 2014; EN 17037 2018)) being conceived in Europe, we believe that their current recommendations, and revisions we have proposed in Table 4, can be applied globally for window view design. Studies in Asia showed that distant views were generally more satisfactory than nearby counterparts (Kent and Schiavon 2020), while both LEED v4.1 (USGBC 2020) and WELL v2 (IWBS 2020) pilot schemes advocate similar recommendations found in the SLL LG10 and EN 17037 (e.g. views of sky, movement, features seen from a certain distance). A framework by Ko *et al.* (Ko et al. 2022) reviewed many international standards and scientific literature, finding some consensus across current recommendations for window view design.

To our findings come some limitations. To test recommended features in daylighting standards (SLL 2014; EN 17037 2018), a limited range of views were used. Beyond these more prominent features, it is unlikely that the criteria in Table 1a accurately characterize the minutiae of every view. One feature missing, but can influence view preferences is privacy (Veitch et al. 2012), which is more prevalent in windows located on ground floors (SLL 2014). Secondly, our approach utilized images to represent the view. Although these are widely used (Tuaycharoen and Tregenza 2005; Brooks et al. 2017; Masoudinejad and Hartig 2018; Kent and Schiavon 2020), the context in which they are applied can undermine their reliability. In Fig. 6, we found that the criterion "movement" was not rated highly across all six views. Inherently, these views are not situated near features (e.g., busy roads) that would generate a lot of movement. Nonetheless, most participants would have been unaware of this fact and may have had difficulty evaluating static stimuli, while simultaneously imagining scenarios where dynamic features could have been in them. As part of future research endeavors, window views depicting extreme variations (e.g. busy highway vs. countryside road) for each environmental criterion could be used. Views that systematically exhibit extreme differences for one or more environmental criteria not only elucidates their ability to dictate view preferences, but may also reveal inherent drawbacks for some criteria (e.g. while some movement is conducive for high view quality, too much movement may cause distraction and would be detrimental for window view design).

5. Conclusions

Using subjective assessments collected from surveys and machine learning algorithms, we evaluated the previously untested environmental information criteria found in daylighting standards. The results showed that these criteria could provide relatively accurate predictions when used to classify preferences across different window view images. The main conclusions we can draw from our work are as follows:

- We recommend the use of the five environmental information criteria
- Promoting these criteria may help produce windows views that are able to support occupant health and wellbeing
- Designing for "nature" in views has a much larger influence on psychological restoration and positive affect than other recommended criteria, and therefore should be used as a minimum requirement. Although we anticipated this result, further work might be needed to substantiate whether nature always outweighs other design criteria across a larger sample of views with more diverse or unique content

Considering the importance of daylighting standards that are used to assess the quality of window view, future work should be carried out to derive and verify the recommendations we promote that can be catered toward the visual needs of building occupants.

Appendix A: (Dis)satisfaction ratings given to visual content.

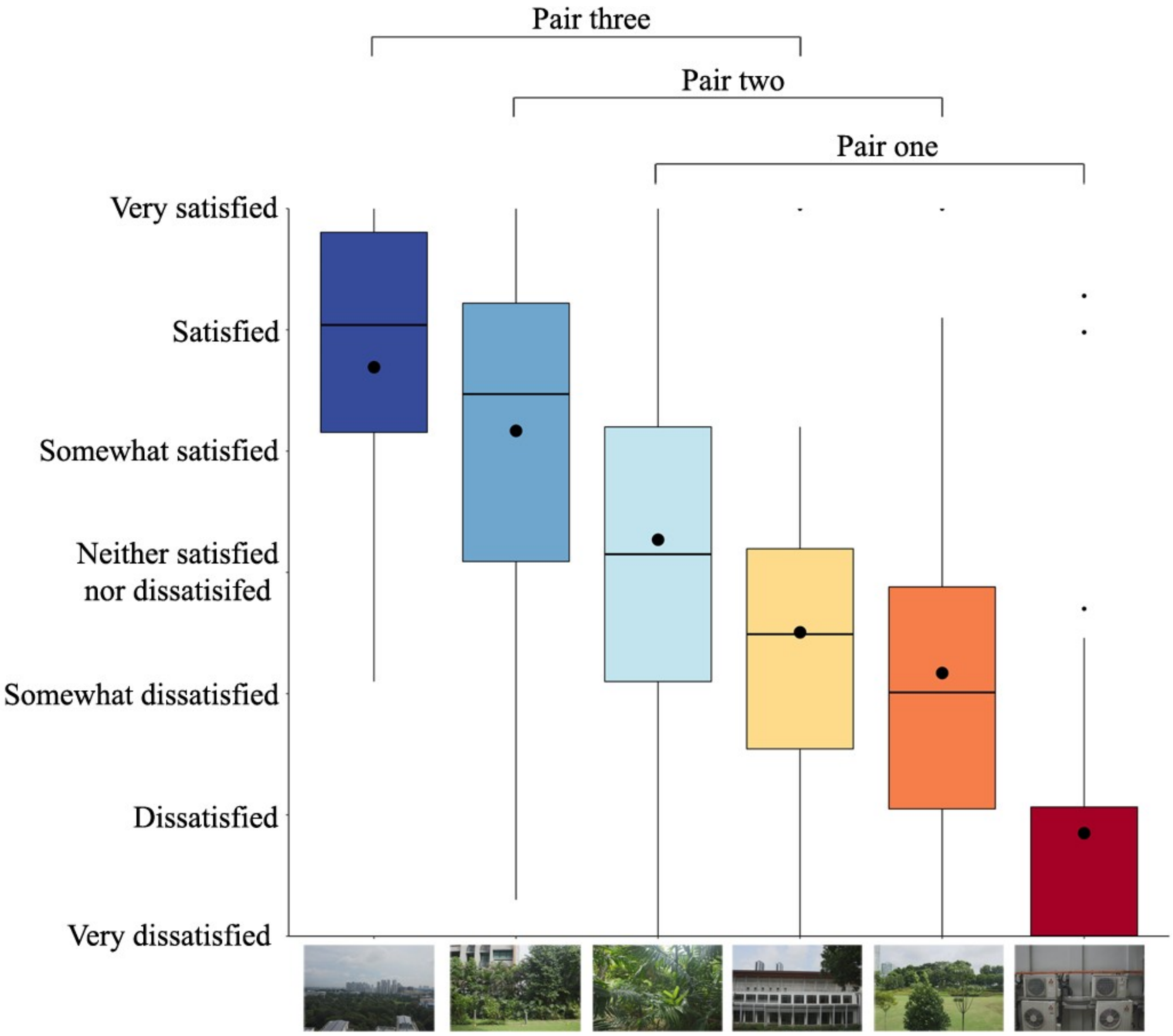

Fig. A1. Boxplots that present subjective (dis)satisfaction ratings given to the variable "visual content" for each of the six different window view images. Observers (*n*= 30) gave a single rating to each window view on a continuous scale containing semantic labels ranging from "Very dissatisfied" to "Very satisfied". Note: the larger circle inside the boxes represents the mean value. Source: the graph was reproduced from the data collecting in our previous work (Kent and Schiavon 2020)

 https://doi.org/10.1080/15502724.2022.2077753

Appendix B: Copy of survey questions used in part two.

When considering the content and looking at the window view, I think that

	Not at all	A bit	Moderately	A lot	Completely
This view is fascinating. My attention is drawn to many interesting things	○	○	○	○	○
I like the content that I can see in this view	○	○	○	○	○
I would be able to rest and recover my ability to focus if I sat and looked at this view	○	○	○	○	○
Looking at this view gives me a break from my day-to-day routine and helps me to relax my focus on getting things done	○	○	○	○	○

When looking at the window view, I am able to determine

	Not at all	A bit	Moderately	A lot	Completely
The general location of the building (e.g. nearby or in a city-centre, sub-urban area, forest, etc.)	○	○	○	○	○
Movement and potential changes that may be occurring outside (e.g. traffic conditions, people walking, etc.)	○	○	○	○	○
Information that can be used to approximate the outside weather condition (e.g. raining, sunny, etc.)	○	○	○	○	○
The approximate time of day (e.g. early morning, noon, late afternoon, etc.) without looking at a clock	○	○	○	○	○

When considering the context and looking at the window view, I feel

	Not at all	A bit	Moderately	A lot	Completely
Busy	○	○	○	○	○
Stressed	○	○	○	○	○
Strained for time	○	○	○	○	○
Relaxed	○	○	○	○	○
Sad	○	○	○	○	○
Happy	○	○	○	○	○
Connected to nature	○	○	○	○	○

Fig. B1. Image showing the 15 survey questions used in part two. Each question contained standardized semantics labels, ranging from “Not at all” to “Completely”. Note: the figure shows the default ordering of the questions that appeared the survey platform. Because the question order was randomized for each participant, it is unlikely that they would have appeared in this sequence

Appendix C. Summary statistics for images presented to participants in stage two showing the average (median), inter-quartile range (IEQ), minimum, and maximum values. Note: numerical values correspond to semantic labels anchored onto the 5-point unipolar scales. Not at all= 0; A bit= 1; Moderately= 2; A lot= 3; Completely= 4

Pair	View	Criteria	Median	IEQ	Minimum	Maximum
One	Urban	Location	0	1	0	4
		Time	0	1	0	3
		Weather	1	2	0	4
		Nature	0	0	0	4
		Movement	0	0	0	4
	Nature	Location	1	2	0	4
		Time	1	1	0	4
		Weather	2	2	0	4
		Nature	3	2	0	4
		Movement	0	1	0	4
Two	Nearby nature	Location	2	1	0	4
		Time	2.5	1	0	4
		Weather	3	1	0	4
		Nature	2	1	0	4
		Movement	1	1	0	4
	Distant nature	Location	2	1	0	4
		Time	3	2	0	4
		Weather	3	1	1	4
		Nature	3	1	0	4
		Movement	2	2	0	4
Three	Nearby urban	Location	2	1	0	4
		Time	2	2	0	4
		Weather	3	1	0	4
		Nature	1	1	0	4
		Movement	1	2	0	4
	Distant urban	Location	3	2	0	4
		Time	3	2	0	4
		Weather	3.5	1	0	4
		Nature	2	1	0	4
		Movement	1	1	0	4
Four	Null 1	Location	3	1	0	4
		Time	3	2	0	4
		Weather	4	1	0	4
		Nature	2	2	0	4
		Movement	3	1	0	4
	Null 2	Location	3	1.25	0	4
		Time	3	2	0	4
		Weather	3	1	0	4
		Nature	2	1	0	4
		Movement	3	1.75	0	4

Appendix D: Structure of neural network plot for image pairs one, two, and three used in stage two.

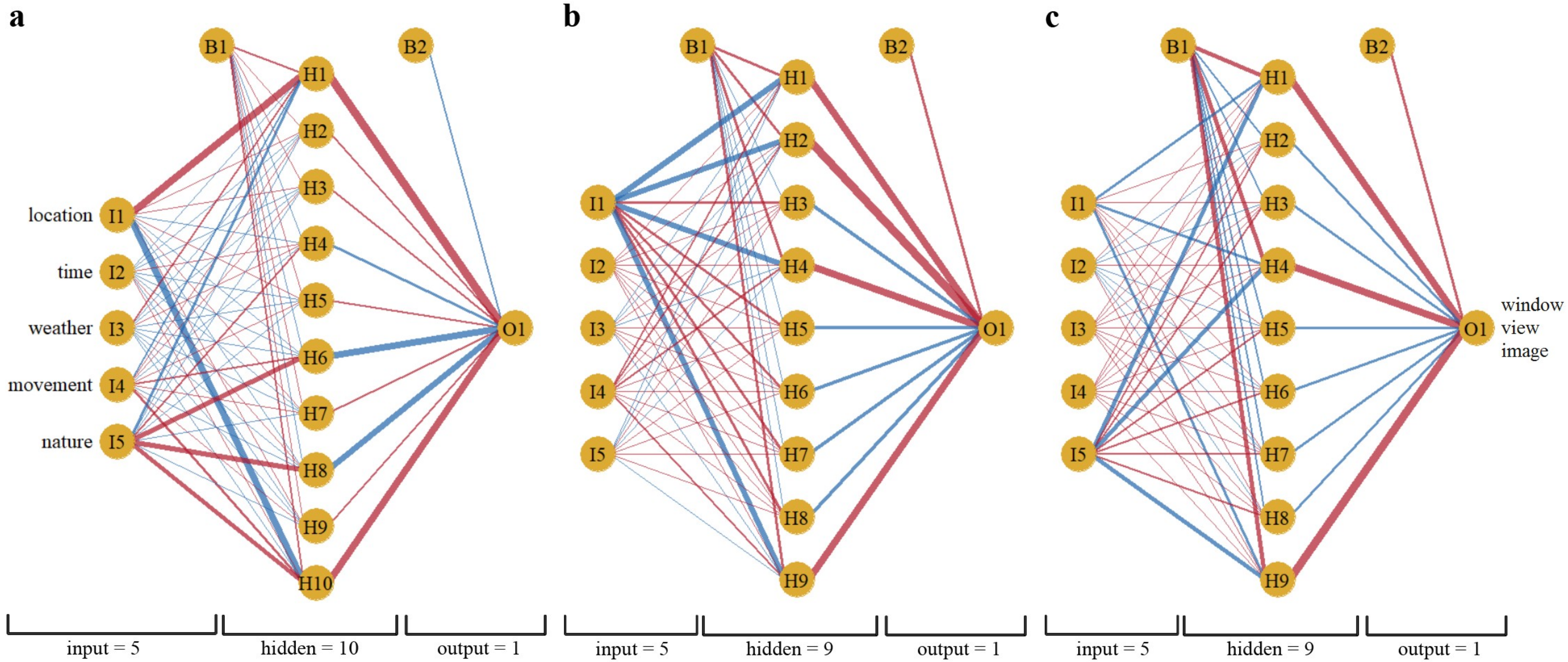

Fig. D1. Neural network plots showing neurons (yellow circles) representing the input (I), hidden (H), and output (O) layers when used to classify: (a) pair one, (b) pair two, and (c) pair three. In both the hidden and output layers, the associated bias (B) is also shown. The neurons are connected together by the weight of the input on the outcome (lines). The color and thickness of the lines is used to determine the sign (red is positive, and blue is negative), and relative importance of each input (i.e. environmental information criteria) on the output (window view image)

 https://doi.org/10.1080/15502724.2022.2077753

Acknowledgements

This study was supported by the Center for the Built Environment (CBE), at the University of California, Berkeley; and the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program.

Funding

The authors report no funding.

Disclosure statement

The authors have no financial interests to declare.

References

Aries M, Veitch J, Newsham G. 2010. Windows, view, and office characteristics predict physical and psychological discomfort. J Environ Psychol. 30(4):533–541. https://doi.org/10.1016/j.jenvp.2009.12.004

Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers E-J, Berk R, Bollen KA, Brembs B, Brown L, Camerer C, et al. 2018. Redefine statistical significance. Nat Hum Behav. 2(1):6–10. https://doi.org/10.1038/s41562-017-0189-z

Bohanec M, Bratko I. 1994. Trading accuracy for simplicity in decision trees. Mach Learn. 15(3):223–250. https://doi.org/10.1023/A:1022685808937

Boser BE, Guyon IM, Vapnik VN. 1992. A training algorithm for optimal margin classifiers. In: Proc Fifth Annu Workshop Comput Learn Theory [Internet]. New York, NY, USA: Association for Computing Machinery; [accessed 2020 Nov 23]; p. 144–152. https://doi.org/10.1145/130385.130401

Breiman L. 2001. Random Forests. Mach Learn. 45(1):5–32. https://doi.org/10.1023/A:1010933404324

Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. Classification and regression trees [Internet]. 1st Edition. Boca Raton: Routledge. https://doi.org/10.1201/9781315139470

Brick J, Kalton G. 1996. Handling missing data in survey research. Stat Methods Med Res. 5(3):215–238. https://doi.org/10.1177/096228029600500302

Brooks AM, Ottley KM, Arbuthnott KD, Sevigny P. 2017. Nature-related mood effects: Season and type of nature contact. J Environ Psychol. 54:91–102. https://doi.org/10.1016/j.jenvp.2017.10.004

Butler DL, Biner PM. 1989. Effects of setting on window preferences and factors associated with those preferences. Environ Behav. 21(1):17–31. https://doi.org/10.1177/0013916589211002

Bzdok D, Altman N, Krzywinski M. 2018. Statistics versus machine learning. Nat Methods. 15:233–234.

Cheung T, Schiavon S, Parkinson T, Li P, Brager G. 2019. Analysis of the accuracy on PMV – PPD model using the ASHRAE Global Thermal Comfort Database II. Build Environ. 153:205–217. https://doi.org/10.1016/j.buildenv.2019.01.055

Davydenko M, Peetz J. 2017. Time grows on trees: The effect of nature settings on time perception. J Environ Psychol. 54:20–26. https://doi.org/10.1016/j.jenvp.2017.09.003

Ding C, He X. 2004. K-means clustering via principal component analysis. In: Twenty-First Int Conf Mach Learn - ICML 04 [Internet]. Banff, Alberta, Canada: ACM Press; [accessed 2021 May 26]; p. 29. https://doi.org/10.1145/1015330.1015408

Dobbin KK, Simon RM. 2011. Optimally splitting cases for training and testing high dimensional classifiers. BMC Med Genomics. 4(1):1–8. https://doi.org/10.1186/1755-8794-4-31

Dogrusoy IT, Tureyen M. 2007. A field study on determination of preferences for windows in office environments. Build Environ. 42(10):3660–3668. https://doi.org/10.1016/j.buildenv.2006.09.010

EN 17037. 2018. Daylight in buildings. Brussels, Belgium: European Committee for Standardization.

Ferguson CJ. 2009. An effect size primer: A guide for clinicians and researchers. Prof Psychol Res Pract. 40(5):532–538. https://doi.org/10.1037/a0015808

Field A, Hole GJ. 2011. How to design and report experiments. London, United Kingdom: SAGE Publications.

Field A, Miles J, Field Z. 2012. Discovering statistics using R. London, United Kingdom: SAGE Publications.

Fotios S. 2019. Using category rating to evaluate the lit environment: Is a meaningful opinion captured? LEUKOS. 15(2–3):127–142. https://doi.org/10.1080/15502724.2018.1500181

Graham LT, Parkinson T, Schiavon S. 2021. Lessons learned from 20 years of CBE's occupant surveys. Build Cities. 2(1):166–184. https://doi.org/10.5334/bc.76

Granzier JJM, Valsecchi M. 2014. Variations in daylight as a contextual cue for estimating season, time of day, and weather conditions. J Vis. 14(1):1–23. https://doi.org/10.1167/14.1.22

Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. 2019. A survey of methods for explaining black box models. ACM Comput Surv. 51(5):1–42. https://doi.org/10.1145/3236009

Hadavi S, Kaplan R, Hunter MCR. 2015. Environmental affordances: A practical approach for design of nearby outdoor settings in urban residential areas. Landsc Urban Plan. 134:19–32. https://doi.org/10.1016/j.landurbplan.2014.10.001

Hanley JA, McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 143(1):29–36. https://doi.org/10.1148/radiology.143.1.7063747

Hartigan JA, Wong MA. 1979. Algorithm AS 136: A K-Means Clustering Algorithm. J R Stat Soc Ser C Appl Stat. 28(1):100–108. https://doi.org/10.2307/2346830

Henderson JC. 2013. Urban parks and green spaces in Singapore. Manag Leis. 18(3):213–225. https://doi.org/10.1080/13606719.2013.796181

Hsu C, Chang C, Lin C. 2010. A practical guide to support vector classification. Department of Computer Science: National Taiwan University.

IWBS. 2020. WELL building standard V2 pilot. WELL Stand [Internet]. [accessed 2022 Jun 6]. https://v2.wellcertified.com/en/v/light/feature/5

Kaplan R. 1993. The role of nature in the context of the workplace. Landsc Urban Plan. 26(1):193–201. https://doi.org/10.1016/0169-2046(93)90016-7

Kaplan R. 2001. The nature of the view from home: Psychological benefits. Environ Behav. 33(4):507–542. https://doi.org/10.1177/00139160121973115

Kent M, Parkinson T, Kim J, Schiavon S. 2021. A data-driven analysis of occupant workspace dissatisfaction. Build Environ. 205:108270. https://doi.org/10.1016/j.buildenv.2021.108270

Kent M, Schiavon S. 2020. Evaluation of the effect of landscape distance seen in window views on visual satisfaction. Build Environ. 183:107160. https://doi.org/10.1016/j.buildenv.2020.107160

Kim J, Kent M, Kral K, Dogan T. 2022. Seemo: A new tool for early design window view satisfaction evaluation in residential buildings. Build Environ. 214:108909. https://doi.org/10.1016/j.buildenv.2022.108909

Kim J, Schiavon S, Brager G. 2018. Personal comfort models – A new paradigm in thermal comfort for occupant-centric environmental control. Build Environ. 132:114–124. https://doi.org/10.1016/j.buildenv.2018.01.023

Ko WH, Kent MG, Schiavon S, Levitt B, Betti G. 2022. A Window View Quality Assessment Framework. LEUKOS. 18(3):268–293.

Kotsiantis SB. 2007. Supervised machine learning: A review of classification techniques. Informatica. 31(3):249–268.

Krosnick JA, Alwin DF. 1987. An evaluation of a cognitive theory of response-order effects in survey measurement. Public Opin Q. 51(2):201–219.

Li W, Samuelson H. 2020. A new method for visualizing and evaluating views in architectural design. Dev Built Environ. 1:100005. https://doi.org/10.1016/j.dibe.2020.100005

Lindal PJ, Hartig T. 2013. Architectural variation, building height, and the restorative quality of urban residential streetscapes. J Environ Psychol. 33:26–36. https://doi.org/10.1016/j.jenvp.2012.09.003

Lindal PJ, Hartig T. 2015. Effects of urban street vegetation on judgments of restoration likelihood. Urban For Urban Green. 14(2):200–209. https://doi.org/10.1016/j.ufug.2015.02.001

Markus TA. 1967. The function of windows— A reappraisal. Build Sci. 2(2):97–121. https://doi.org/10.1016/0007-3628(67)90012-6

Masoudinejad S, Hartig T. 2018. Window view to the sky as a restorative resource for residents in densely populated cities. Environ Behav. 52(4):401–436. https://doi.org/10.1177/0013916518807274

Mayer FS, Frantz CM, Bruehlman-Senecal E, Dolliver K. 2009. Why is nature beneficial?: The role of connectedness to nature. Environ Behav. 41(5):607–643. https://doi.org/10.1177/0013916508319745

McCulloch WS, Pitts W. 1943. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys. 5(4):115–133. https://doi.org/10.1007/BF02478259

Metz CE. 1978. Basic principles of ROC analysis. Semin Nucl Med. 8(4):283–298. https://doi.org/10.1016/s0001-2998(78)80014-2

Mingers J. 1986. An empirical comparison of pruning methods for decision tree induction. Mach Learn. 4:227–243.

Poulton EC. 1989. Bias in quantifying judgements. Hove and London, United Kingdom: Lawrence Erlbaum Associates.

Raschka S. 2020. Model evaluation, model selection, and algorithm selection in machine learning. ArXiv181112808 Cs Stat.:1–49.

Salesses P, Schechtner K, Hidalgo CA. 2013. The collaborative image of the city: mapping the inequality of urban perception. PLOS ONE. 8(7):e68400. https://doi.org/10.1371/journal.pone.0068400

Sarker IH. 2021. Machine learning: Algorithms, real-world applications and research directions. SN Comput Sci. 2(3):160. https://doi.org/10.1007/s42979-021-00592-x

SLL. 2014. Lighting guide 10: Daylighting - A guide for designers. London, United Kingdom: Chartered Institution of Building Services Engineers (CIBSE).

Sokolova M, Japkowicz N, Szpakowicz S. 2006. Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. In: Sattar A, Kang B, editors. AI 2006 Adv Artif Intell. Berlin, Heidelberg: Springer; p. 1015–1021. https://doi.org/10.1007/11941439_114

Statnikov A, Wang L, Aliferis CF. 2008. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC Bioinformatics. 9(1):319. https://doi.org/10.1186/1471-2105-9-319

Tennessen CM, Cimprich B. 1995. Views to nature: Effects on attention. J Environ Psychol. 15(1):77–85. https://doi.org/10.1016/0272-4944(95)90016-0

Tregenza P, Wilson M. 2013. Daylighting: Architecture and lighting design. Milton Park, United Kingdom: Routledge.

Tuaycharoen N, Tregenza PR. 2005. Discomfort glare from interesting images. Light Res Technol. 37(4):329–341.

Ulrich RS. 1981. Natural versus urban scenes: Some psychophysiological effects. Environ Behav. 13(5):523–556. https://doi.org/10.1177/0013916581135001

USGBC. 2020. LEED v4.1: Building design and construction [Internet]. [accessed 2022 Jun 6]. https://www.usgbc.org/credits/commercial-interiors-retail-commercial-interiors-hospitality-commercial-interiors/v41/eq123

Veitch JA, Christoffersen J, Galasiu AD. 2012. Daylight and view through residential windows: Effects on well-being. In: LDA Mag. Krakow, Poland; p. 1–6.

Veitch JA, Galasiu AD. 2012. The physiological and psychological effects of windows, daylight, and view at home: Review and research agenda: (554552013-001) [Internet]. [accessed 2019 Dec 2]:1–57. https://doi.org/10.1037/e554552013-001

Velarde MD, Fry G, Tveit M. 2007. Health effects of viewing landscapes – Landscape types in environmental psychology. Urban For Urban Green. 6(4):199–212. https://doi.org/10.1016/j.ufug.2007.07.001

Waczynska M, Sokol N, Martyniuk-Peczek J. 2021. Computational and experimental evaluation of view out according to European Standard EN17037. Build Environ. 188:107414. https://doi.org/10.1016/j.buildenv.2020.107414

Warren C, McGraw AP, Van Boven L. 2011. Values and preferences: defining preference construction. Wiley Interdiscip Rev Cogn Sci. 2(2):193–205. https://doi.org/10.1002/wcs.98

Yuan C, Yang H. 2019. Research on k-value selection method of k-means clustering algorithm. J. 2(2):226–235. https://doi.org/10.3390/j2020016