sustainability

MDPI

Article

Comparison of Response Scales as Measures of Indoor Environmental Perception in Combined Thermal and Acoustic Conditions

Wonyoung Yang [1], Hyeun Jun Moon [2] and Jin Yong Jeon [1,*]

[1] Department of Architectural Engineering, Hanyang University, Seoul 04763, Korea
[2] Department of Architectural Engineering, Dankook University, Yongin 16890, Korea
* Correspondence: jyjeon@hanyang.ac.kr; Tel.: +82-2-2220-1795

Received: 7 May 2019; Accepted: 16 July 2019; Published: 22 July 2019

check for updates

Abstract: Response scales are widely used to assess the personal experience of sensation and perception in built environments, and have a great impact on the quality of the responses. The purpose of this study was to investigate the effects of response scales on human sensation and perception in moderate indoor environments. Four different response scales were compared under three room temperatures (19.0 °C, 24.5 °C, and 30.0 °C) and five acoustic stimuli (ambient noise, 42 and 61 dBA × water sounds and traffic noise): a bipolar seven-point scale according to ISO 10551:1995, a unipolar 11-point scale according to ISO/TS 15666:2003, these two scales combined for each sensory comfort assessment, and a bipolar visual analogue scale. The degree of relative differentiation based on indoor physical factors made no significant difference across the four response scales. Therefore, the effects of physical factors on human response could be assessed by using any of the four scales tested in this study, with a statistical significance at $p < 0.05$ in moderate environments. The choice of response scale would depend not only on the type of physical stimulus but also on the question of sensation or perception. The reliability of each response scale was different according to the subjective attributes. The bipolar visual analogue scale was subjectively preferred by the respondents.

Keywords: response scales; visual analogue scale; seven-point scale; 11-point scale; respondent preference; thermal comfort; acoustic comfort; indoor environmental comfort; sensation; perception

1. Introduction

1.1. Background and Objectives

It is common in indoor environmental quality research to collect data using subjective questionnaires. Capturing true responses regarding occupants' comfort in and acceptance of an indoor environment is crucial when evaluating a building's performance. The impact of the questionnaire design characteristics on the quality of the responses should be minimized to achieve true responses [1]. Determining the characteristics of the response scale is often the most important decision in ensuring good measurement properties [2–4]. DeCastellarnau [1], in her latest review article, classified 23 different characteristics of response scales and provided their effects on data quality. Ten characteristics out of the 23 have been found to affect data quality: the scales' evaluative dimensions, the types of scales, scale length, verbal labels, number of fixed reference points, order of numerical labels, correspondence between numerical and verbal labels, scale illustrative format, scale layout display, and the label visual separation.

The study of response scales is considered fundamental research in every discipline. With regards to indoor environmental perception, the effects of response scales used for combined thermal and acoustic conditions have not been studied yet. In fact, indoor environmental factors such as thermal comfort, indoor air quality (IAQ), lighting, and acoustics have been studied independently. Therefore, subjective response scales also have been developed for discipline-specific purposes. However, in combined thermal and acoustic conditions, which are more realistic indoor environmental setups, will each discipline-specific response scale still be the optimal response scale? This study focuses on the effects of subjective response scales for combined thermal and acoustic conditions rather than the combined effects of thermal and acoustic conditions to answer this fundamental research question.

In the present study, our objective is to investigate effects of response scales for young adults with regard to indoor environmental sensation and perception in combined thermal and acoustic conditions. Four different response scales were compared in this study: a bipolar seven-point scale according to ISO 10551:1995 [5], a unipolar 11-point scale according to ISO/TS 15666:2003, [6] combined scale with seven-point and 11-point scales for each environment, and a bipolar visual analogue scale. Similarities and differences among the four response scales will be discussed for combined thermal and acoustic conditions in moderate indoor environments. The performances of and preference for the response scales were also investigated in this study.

1.2. Literature Review on Response Scales in Indoor Environmental Perception

To investigate effects of response scales on indoor environmental sensation and perception in combined thermal and acoustic conditions, the response scales used in previous studies have been reviewed. Response scales for thermal comfort and for noise perception have been independently developed and standardized in each area of research.

For thermal comfort assessments, a bipolar seven-point scale (Figure 1) has been used according to ISO 10551:1995 [5] based on Fanger's model [7]. Because of the nature of neutrality in thermal sensation and perception, a bipolar scale was adopted to evaluate subjective responses in thermal conditions. By contrast, either a unipolar five-point verbal scale or a unipolar 11-point numeric scale has been used for subjective acoustic assessments according to ISO/TS 15666:2003 [6]. As acoustic assessment was developed with noise, which is a negative sound to be controlled, a unipolar scale was introduced. However, to date, there has been no explicit academic consensus on response scales for assessing human sensation and perception in combined thermal and acoustic conditions. EN 15251 [8] specifies indoor environmental input parameters for the design and assessment of the energy performance of buildings. These parameters address the indoor air quality (IAQ), thermal environment, lighting, and acoustics. It adopted the seven-point thermal sensation scale and the acceptable percentage of votes for thermal environment and air quality classification, but no direct subjective assessment methods for acoustics and lighting were recommended.

In previous studies regarding the combined effects of indoor environmental sensation and perception in thermal and acoustic environments, visual analogue scales (VAS) were often used to assess thermal and acoustic sensation and perception [9–15]. Visual analogue scales are known to present better metrical features than category scales [16] because they allow participants to make free subjective assessments [17]. Thus, VAS can have many observational points. Yang and Moon [18] used a unipolar 11-point scale instead of a bipolar VAS for subjective assessments in combined thermal, acoustic, and lighting conditions. Most of these assessments used identical scales for both thermal and acoustic assessments. In indoor environmental quality field studies dealing with thermal, acoustic, and illuminous conditions and indoor air quality at a same time, various types of scales have been used for subjective assessment. Interval scales of 4–7 points [19–35] have been frequently used for combined environmental assessments. Dichotomous scales [21,22,36–38], 11-point scales [39], 13-point scales [39], and VASs [20,36] were also used for indoor environmental assessments. Most of them used identical scales throughout the assessment regardless of the type of environment [23–35,37,38], but some used different scales for each environmental factor [19–22,36,39]. Table 1 summarizes the response scales used for subjective assessment in previous studies dealing with combined environmental factors. Response scales were used as research methods in these previous studies, even though the appropriateness of the response scales for use in the studies was not fully investigated.

Table 1. Response scales in combined environmental research.

Study		Population (Category)	Sample Size	Scale			Survey Tool
				Type	Polarity	Length	
Effects of indoor environmental sensation and perception in thermal and acoustic conditions							
Nagano & Horikoshi 2001 [9]	Lab	Age 19–37	29 men	VAS	Unipolar	0 to 100	Paper
Pellerin & Candas 2004 [10]	Lab	Mean age 23.1 to 24.1	18 (W9/M9)	VAS			
		(Sensation, Preference)			Bipolar	0 to 100	Not identified
		(Comfort)			Unipolar	−50 to 50	
Witterseh et al. 2004 [11]	Lab	Age 18–29	30 (W14/M16)	VAS	Bipolar	0 to 100	Not identified
Nagano & Horikoshi 2005 [12]	Lab	Age 19–26	22 men	VAS	Unipolar	0 to 100	Paper
Tiller et al. 2010 [13]	Lab		30 (W16/M14)	VAS	Unipolar	0 to 100	Computer
Yang et al. 2017 [14]	Lab	Mean age 25.5 to 25.8	26 (W13/M13)	VAS	Bipolar	−5.0 to 5.0	Paper
Yang and Moon 2018 [15]	Lab	Age 19–27	24 (W12/M12)	VAS	Bipolar	−5.0 to 5.0	Paper
Yang and Moon 2019 [18]	Lab	Mean age 21.3 to 23.3	60 (W30/M30)	11-point numerical scale with five verbal labels	Unipolar	0 to 10	Tablet
Indoor environmental quality							
Humphreys 2005 [19]	Field	26 office buildings, France, Greece, Portugal, Sweden, UK	4655 responses				Not identified
		(TH, AV, H, L N)		5-point verbal scale	Bipolar	0 to 2	
		(IAQ)		7-point verbal scale	Bipolar	0 to 2	
		(Overall comfort)		6-point verbal scale	Bipolar	0 to 5	
Wong et al. 2008 [36]	Field	Typical AC offices, Hong Kong	293				
		(Thermal/IAQ)		Dichotomous scale	Acceptable/not acceptable	Yes (1), No (0)	
		(Visual/Acoustic)		VAS			
Andersen et al. 2009 [20]	Field	Dwellings, Denmark	933 summer 636 winter				Paper via mail
		(L, IAQ)		VAS	Bipolar	0 to 100	
		(A)		VAS	Bipolar	−50 to 50	
		(TH)		7-point interval scale	Bipolar	−3 to 3	

Table 1. *Cont.*

Study		Population (Category)	Sample Size	Scale			Survey Tool
				Type	Polarity	Length	
			Indoor environmental quality				
Li et al. 2013 [21]	Field	Traditional Chinese buildings vs. rural buildings	139 Tulou 97 normal rural				Not identified
		(Satisfaction/Dissatisfaction)		Dichotomous scale	Unipolar	1 to 0, 0 to −1	
		(Sensation TH)		7-point numerical scale	Bipolar	−3 to 3	
		(Sensation V, L)		5-point numerical scale	Bipolar	−2 to 2	
		(Sensation IAQ, A)		5-point numerical scale	Unipolar	0 to 4	
Ricciardi and Buratti 2018 [39]	Field	7 university classrooms, Italy					Not identified
		(TH)	331	13-point scale	Bipolar	−3 to 3	
		(L, A)	597	11-point numerical scale	Unipolar	0 to 10	
Mui et al. 2018 [22]	Field	Small residential units, Hong Kong	52	Dichotomous scale	Acceptable /unacceptable		Not identified
		(TH)		7-point numerical-verbal scale	Bipolar	−3 to 3	
		(IAQ)		5-point verbal scale	Bipolar		
		(A, L)		Point award		0 to 100	
Paul and Taylor 2008 [23]	Field	1 green building, 1 conventional building, Australia	40 green 53 conventional	7-point numerical scale	Bipolar	1 to 7	Paper
Lai et al. 2009 [37]	Field	32 residential apartments, Hong Kong	125 (W82/M43)	Dichotomous scale	Acceptable/not acceptable	Yes (1), No (0)	Interview
Bluyssen et al. 2011 [24]	Field	59 office buildings, Germany, Switzerland, Italy, Finland, Denmark, Portugal, The Netherlands, UK	5732	7-point scale	Bipolar	1 to 7	Paper via mail
Huang et al. 2012 [38]	Lab	Mean age 22, Chana	120 (W60/M60)	Dichotomous scale	Unipolar	−1 and 0 0 and 1	Not identified
Frontczak et al. 2012 [25]	Field	CBE POE data 351 office buildings, USA	52980	7-point ordered scale	Bipolar	−3 to 3	Web-based
Hwang and Kim 2013 [26]	Field	1 office building, Korea	2744 (5 times)	7-point	Bipolar	−3 to 3	Web-based

Table 1. *Cont.*

Study		Population (Category)	Sample Size	Scale			Survey Tool
				Type	Polarity	Length	
			Indoor environmental quality				
Fassio et al. 2014 [27]	Field	1 university classroom, Italy	17	4-point verbal scale	Bipolar	0 to 3	Paper
Liang et al. 2014 [28]	Field	3 green buildings, 2 conventional buildings, Taiwan	134 green 99 conventional	7-point verbal scale	Unipolar	0 to 100	Paper
Woo 2014 [29]	Field	4 green buildings, Korea	114	5-point verbal scale	Bipolar	1 to 5	
Pei et al. 2015 [30]	Field	10 green buildings, China	+1000	6-point scale	Unipolar	−1 and 0 0 and 1	Paper
Ravindu et al. 2015 [31]	Field	1 LEED certified factory, Sri Lanka	70	5-point scale	Unipolar	1 to 5	Not identified
Martellotta et al. 2016 [32]	Field	1 hypermarket, Italy	120	7-point verbal scale	Unipolar	1 to 7	Not identified
Xue et al. 2016 [33]	Field	5 40-story residential buildings, Hong Kong	482	5-point verbal scale	Unipolar	1 to 5	Paper via mail
Karmann et al. 2017 [34]	Field	34 all-air buildings 26 radiant buildings, USA	2247 all-air 1645 radiant	7-point verbal scale	Unipolar	Verbal analysis	Web-based
Choi and Moon 2017 [35]	Field	9 university buildings, 5 commercial buildings, USA	411	7-point numerical scale	Bipolar	3-point: negative, neutral, positive 2-point: negative, positive	Paper

Note: A (Acoustics), AC (Air conditioned), AV (Air movement), H(Humidity), IAQ (Indoor Air Quality), L (Lighting), N (Noise), TH (Thermal condition), V(Ventilation).

The limited information on response scales for combined environmental assessments is a challenge to the success of subjective building performance. To the best of the authors' knowledge, no studies on an optimized response format for combined environmental assessments have been reported to date, although the choice of response format should be an explicit step in the process of constructing a questionnaire, and different response formats may be appropriate for different constructs [40].

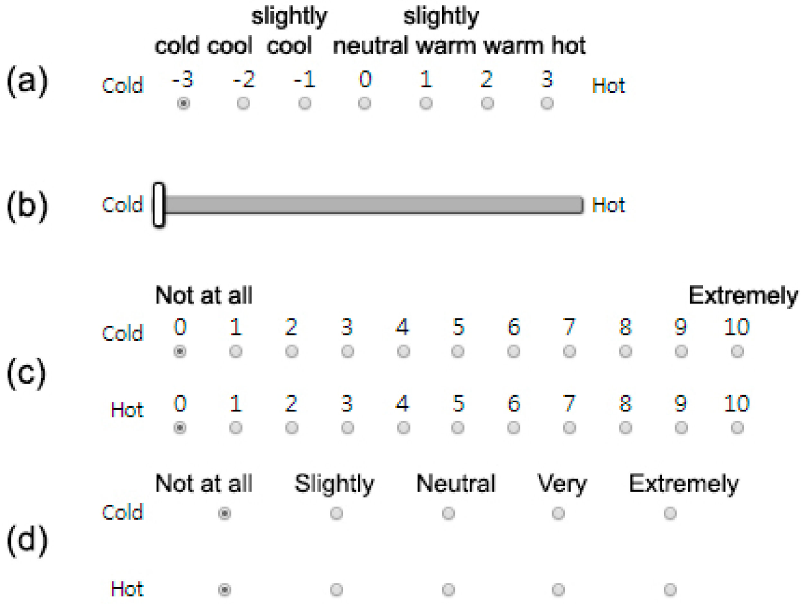

Figure 1. Examples of the response scales: (**a**) Bipolar seven-point scale, (**b**) bipolar VAS, (**c**) unipolar 11-point numeric scale, and (**d**) unipolar five-point verbal scale.

2. Methods

2.1. Respondents

Fifty university students (23 men and 27 women) participated in three experimental sessions each and received compensation for their participation. Informed consent was obtained from each participant. No noise hearing tests were performed, but potential participants who had hearing impairments were excluded. Participants were asked to wear a clothing ensemble of nearly 0.75 clo according to ASHRAE Standard 55-2004 [41]. The clothing ensemble consisted of a thick pair of straight trousers (0.24 clo), a long-sleeved flannel shirt (0.34 clo), a pair of socks (0.02 clo), underwear (0.04 clo), a T-shirt (0.08 clo), and slippers (0.03 clo). Table 2 lists the respondents' physical characteristics.

Table 2. Physical characteristics of respondents.

Gender		Age (years)	Height (cm)	Weight (kg)	BMI (kg/m^2)
Women	Mean (S.D.)	21.8 (1.9)	163.0 (6.0)	56.2 (5.6)	21.2 (2.0)
Men	Mean (S.D.)	23.2 (1.9)	173.7 (6.7)	69.4 (9.1)	23.0 (2.2)

Feedback on indoor environmental sensitivity was requested for temperature, humidity, indoor air quality, lighting, and acoustics using a bipolar seven-point scale in an additional survey after completion of the all sessions.

2.2. Test Laboratory and Experimental Conditions

The experiment was conducted in a test laboratory (4.0 m × 5.0 m × 2.4 m) furnished as a small classroom. The test laboratory in this study was built for indoor environmental research. The air temperature and humidity of the indoor environmental chamber were controlled using variable refrigerant flow systems, humidifiers, dehumidifiers, and ventilation systems located in the indoor and outdoor chambers. The ventilation system was in operation during the experiment. The local air velocity was measured to be less than 0.1 m/s.

The mean illuminance levels along the desk surface during the experiments were 995.0 × (Konica Minolta T-10A: Tokyo, Japan). The color temperature of the lamp was 6500 K according to the specification sheets provided by the manufacturer. A loudspeaker system (Turbosound Milan M10: Patridge Green, UK) was used as a sound source and was located on the rear side to minimize the spatial sensitivity of the sound sources. The reverberant time of the test laboratory was measured as 0.3 s at 500 Hz for octave bands (01dB dB4: Lyon, France). The ambient noise level in the laboratory was 38 dBA (01dB Solo: Lyon, France) when the thermal systems were operated.

Three room-temperature levels (19.0 °C, 24.5 °C, and 30.0 °C), corresponding to cool, neutral, and warm sensations, were chosen [41]. For each room air temperature, a constant relative humidity of 40% was set. Thermal data, measured using a digital thermometer (Autonics THD-W: Busan, Korea) installed on the wall inside the chamber, were validated through comparisons with data obtained using digital thermometers (SATO SK-L200THIIα: Tokyo, Japan) placed at the positions of the four participants. The target conditions for the air temperature and relative humidity are listed in Table 3. The variations are within the range of the just noticeable difference (JND) of temperature and relative humidity [42,43].

Table 3. Thermal conditions and variations.

Target Temperature Relative Humidity	Measured Temp Mean (S.D.)	Measure RH Mean (S.D.)	Sensation ASHRAE 55-2013
19.0 °C, 40%	19.5 °C (0.4)	43.3% (2.0)	Cool
24.5 °C, 40%	24.7 °C (0.3)	40.8% (1.4)	Neutral
30.0 °C, 40%	29.8 °C (0.2)	38.9% (1.2)	Warm

Four different sounds (traffic and water × 42 and 61 dBA) were played through a loudspeaker, considering measured daytime median noise exposure levels [44]. Water sounds were acquired from an open website [45], and traffic noises were recorded in the living room of a residential building. The levels of the sound sources were adjusted with an audio controller. Sound level differences across the participants' positions were measured at ± 0.3 dBA. Figure 2 shows the octave band frequency spectra of the sound sources, including ambient noise in the chamber.

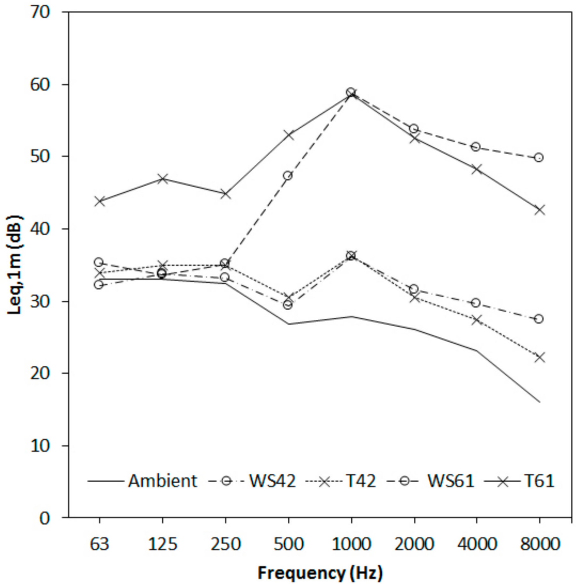

Figure 2. Frequency spectra of sound sources.

2.3. Response Scales and Semantic Adjectives

Four different response scales were compared in this study: bipolar seven-point scale, unipolar 11-point scale, combined scale with seven-point and 11-point scales for each environment, and a visual analogue scale. The 10 semantic adjectives used throughout the tests were soft vs. loud, acoustically uncomfortable vs. comfortable, cold vs. loud, thermally uncomfortable vs. comfortable, and uncomfortable vs. comfortable (overall indoor environment). All subjective assessments were conducted using a web-based tablet interface with a finger touch.

A bipolar seven-point numerical scale (bipolar7) with end-only labels was introduced based on ISO 10551:1995 [5], which was developed for thermal sensation assessment. Radio buttons were used to create discrete rating scales from −3 to 3. Respondents immediately saw the response options that were available and could choose between them when the radio buttons were used for the survey [46].

A unipolar 11-point numerical scale (unipolar11) with end- and midpoint labels was adopted based on ISO/TS 15666:2003, [6] which was developed for socioacoustic surveys. It is assumed that a 0-to-10 scale would be more readily understood and manipulated than shorter ones. Most people are familiar with base-10 numeric systems through currency and other familiar counted materials. Radio buttons were also used to create 11 discrete scales from 0 to 10. Three verbal labels of “Not at All,” “Neutral,” and “Extremely” were placed at the top of “0,” “5,” and “10.” The number of questions doubled over the bipolar scales because the unipolar scale could only evaluate to a degree of one attribute.

A combined scale (combined) with both the bipolar seven-point and the unipolar 11-point scales was also introduced. The bipolar seven-point scale was used for thermal attributes and overall comfort assessment, and the 11-point scale was used for acoustic assessment. The default marker was always placed at the very left end of the interval scale.

A bipolar visual analogue scale (bipolar VAS) was introduced in the study. The questionnaire contents were identical to those of the bipolar seven-point scale except for the response format. VAS is acknowledged in the medical sector because a measurement with continuous scales is sensitive [47]. VAS consists of a plain, mostly horizontal line, and mostly verbal end labels. Respondents give a rating by placing a mark on the line. The length of the line is 100 mm, which corresponds to a VAS score between 0 and 100. In this study, a numerical value of −5.0 to 5.0 was assigned to the responses for the statistical analyses. A slider was placed at the left end in the default setting as an indicator of the rating mark. However, respondents were required not to drag but to click the slider to avoid the potential technical problems caused by dragging the slider with fingers.

2.4. Experimental Design and Procedure

A factorial within-subject design with repeated measures was employed with three independent variables: response scale (bipolar VAS, bipolar7, unipolar11, and combined), room temperature (19.0 °C, 24.5 °C, and 30.0 °C), and noise (ambient, WS42, T42, WS61, and T61).

The questionnaire consisted of ten semantic differential adjectives to survey acoustic sensation (soft vs. loud), perception (uncomfortable vs. comfortable), thermal sensation (cold vs. hot), perception (uncomfortable vs. comfortable), and overall indoor environmental perception (uncomfortable vs. comfortable). The semantic attribute “comfortable” was placed at the right end, and the semantic attribute “uncomfortable” was positioned at the left end for bipolar scales. For the unipolar scale questionnaire, the sematic attributes at the right and left ends for the bipolar scale questionnaire were asked one by one.

Participants were required to attend all three sessions performed in the test laboratory. During each session, a quick demonstration regarding how to use the tablet was provided to the participants, but the experimental conditions were not mentioned to avoid potential experimental bias. A maximum of six participants simultaneously assessed the combined environmental conditions. The response data provided by the participants were automatically saved on a server.

All sessions were initially conducted at a room temperature of 19.0 °C, followed by 30 °C, and thereafter at 24.5 °C. To ensure an equivalent room air temperature and mean radiant temperature, each thermal condition was set at least 15 h prior to the test. The clothing insulation of each subject was visually inspected as it should not significantly deviate from the 0.75-clo requirement prior to testing.

In each 60-min-long session, a 20-min adaptation period was implemented at the beginning of the session for thermal adaptation, as shown in Figure 3. The participants were sedentary during the adaptation period. Each sound stimulus was presented for 50 s, and a response time was provided until all participants in the test group submitted their responses. The ambient sounds for the four different response scales were assessed at the beginning and at the end. The four sound sources combined with the four response scales were randomly presented in each test session, and their replicas were also presented in random order to validate the reliability of the response scales.

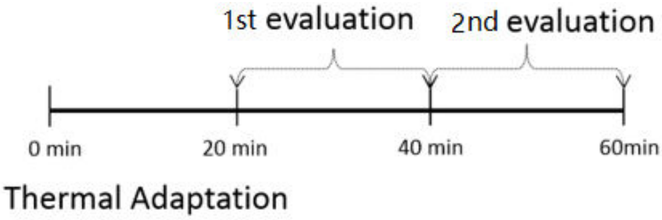

Figure 3. Experimental procedure for each session.

After completion of the three different thermal sessions, respondents' preferences for the scales were measured by a quick survey. Respondents' feedback with regard to indoor environmental sensitivity was also obtained for temperature, humidity, indoor air quality, lighting, and acoustics. The current standards of the rating scales for each condition were not mentioned to avoid potential experimental bias.

2.5. Statistics

The statistical analyses were performed using two different approaches: original and normalized data analyses. The original data from the respondents were used to analyze the correlation performance for reliability. Fisher's Z transformation was applied to compare the correlation coefficients of repeated measures on each response scale. The original data were also applied to a factorial analysis of variance (ANOVA) to validate sensitivity to differentiate the effects of temperature and sound on each response scale. ANOVA is a powerful statistical test and was used in this case, although normality cannot be guaranteed for subjective ratings [48,49].

A repeated-measures ANOVA was also used to test the scale factor for the two repeated measures to confirm the effects of the repeated measures and response scales at the same time. The original data were converted to unipolar 0.0-to-10.0 scales to perform an ANOVA on the four response scales having different numerical ranges. If a response value was greater than 0, it was treated as a right-end semantic attribute, and if a response value was less than 0, it was treated as a left-end semantic attribute. Then, the weighting factors were applied to normalize the response values from 0.0 to 10.0. Three corrections (Greenhouse–Geisser, Huynh–Feldt, and lower-bound) for violations of sphericity were used to test the sphericity. The Mauchly sphericity test needs more than three repeated measures, but only two measures were performed in this study. An epsilon (ε) value of 1 was found for the three corrections across all subjective attributes, which indicates that the condition of sphericity was exactly met. A Bonferroni post hoc test was applied.

3. Results

3.1. Response Times

The response time for the questionnaire was observed based on the submission time as listed in Table 4. The bipolar seven-point scale had the shortest response time, followed by the bipolar and combined scales. However, the response times for the bipolar7 and bipolar VAS were not significantly difference. The unipolar 11-point scale had the longest response time among the four scales tested in this study because the number of questions with the same content was twice that of the bipolar scales.

Table 4. Means and Bonferroni's post hoc test results for response time by response scale (means with different letters are significantly different. $p < 0.05$; A > B > C in each row).

	Bipolar VAS		**Bipolar7**		**Unipolar11**		**Combined**	
1st (s)	45.2	C	44.4	C	64.4	A	57.9	B
2nd (s)	40.2	B	38.7	B	56.9	A	54.6	A
Average (s)	42.7		41.5		61.1		55.7	

3.2. Correlation Coefficients for Repeated Measures

Correlations were assessed for the pair of first and second measures on each response scale. The bipolar VAS had higher correlation strength than the bipolar seven-point scale using Fisher's Z transformation ($p < 0.05$) with regard to the loudness sensation, thermal sensation, and thermal comfort attributes in Table 5. The unipolar 11-point scale was better than the bipolar seven-point scale with regard to acoustic attributes. The reliabilities of softness for both the unipolar 11-point and combined scales were lower than those of loudness.

Table 5. Pearson's correlation coefficients between repeated measures ($p < 0.0005$) and Fisher's Z transformation ($p < 0.05$) results (coefficients that do not share a letter are significantly different; A > B > C > D in each column).

		N	**Sensation**				**Acoustic**		**Thermal**		**Indoor Environmental**	
			Soft (L)–Loud (R)		**Cold (L)–Hot (R)**		**Discomfort (L)–Comfort (R)**		**Discomfort (L)–Comfort (R)**		**Discomfort (L)–Comfort (R)**	
Bipolar VAS		750	0.860	B	0.870	A	0.775	A	0.813	A	0.772	B
Bipolar7		750	0.820	C	0.843	B	0.764	AB	0.771	B	0.775	B
Unipolar11	L	750	0.766	D	0.866	A	0.757	AB	0.727	C	0.747	B
	R	750	0.877	AB	0.845	B	0.750	AB	0.734	C	0.692	C
Combined	L	750	0.778	D	0.850	AB	0.735	B	0.746	BC	0.922	A
	R	750	0.895	A			0.780	A				

Spearman's correlation coefficients were calculated between each pair of response scales, and are listed in Tables 6 and 7. The correlation coefficients between bipolar7 and the combined scales, and between unipolar11 and the combined scales, which have identical response scales, were the highest for each sensation attribute. The combined scale consisted of the unipolar11 scale for acoustics and the bipolar7 scale for thermal conditions. Contrastingly, for the comfort attributes, the correlation coefficients between the bipolar7 and bipolar VAS scales were the greatest. In general, for all response scales, the correlation coefficients for the comfort attributes were relatively higher and more stable than those for the sensation attributes.

Figures 4 and 5 show scatter plots for the response scales. The mean values of the two repeated measures were used for the analyses. For the unipolar scales, the right-end semantic attributes, which were "loud," "hot," and "comfortable," were used to draw the scatter plots.

Table 6. Spearman's correlation coefficients between response scales for acoustic and thermal sensations ($p < 0.0005$).

		Acoustic Sensation					**Thermal Sensation**				
		Bipolar VAS	**Bipolar 7**	**Unipolar 11**		**Combined**		**Bipolar VAS**	**Bipolar 7**	**Unipolar 11**	
		Soft–Loud	**Soft–Loud**	**Soft**	**Loud**	**Soft**		**Cold–Hot**	**Cold–Hot**	**Cold**	**Hot**
Bipolar7	Soft–Loud	0.905					Cold–Hot	0.913			
Unipolar11	Soft	−0.813	−0.826				Cold	−0.793	−0.777		
	Loud	0.883	0.869	−0.765			Hot	0.785	0.773	−0.480	
Combined	Soft	−0.784	−0.797	0.891	−0.715		Cold–Hot	0.907	0.927	−0.782	0.779
	Loud	0.875	0.869	−0.744	0.932	−0.727					

Table 7. Spearman's correlation coefficients between response scale for acoustic, thermal, and indoor environmental comfort ($p < 0.0005$, D: discomfort, C: comfort).

		Acoustic Comfort					**Thermal Comfort**					**Indoor Environmental Comfort**			
		Bipolar VAS	**Bipolar 7**	**Unipolar 11**		**Combined**		**Bipolar VAS**	**Bipolar 7**	**Unipolar 11**			**Bipolar VAS**	**Bipolar 7**	**Unipolar 11**	
		D–C	**D–C**	**D**	**C**	**D**		**D–C**	**D–C**	**D**	**C**		**D–C**	**D–C**	**D**	**C**
Bipolar7	D–C	0.882					D–C	0.893				D–C	0.892			
Unipolar11	D	−0.785	−0.804				D	−0.808	−0.797			D	−0.785	−0.767		
	C	0.841	0.845	−0.801			C	0.865	0.846	−0.836		C	0.840	0.810	−0.797	
Combined	D	−0.800	−0.779	0.868	−0.768		D–C	0.883	0.886	−0.785	0.834	D–C	0.853	0.876	−0.758	0.802
	C	0.833	0.838	−0.758	0.877	−0.800										

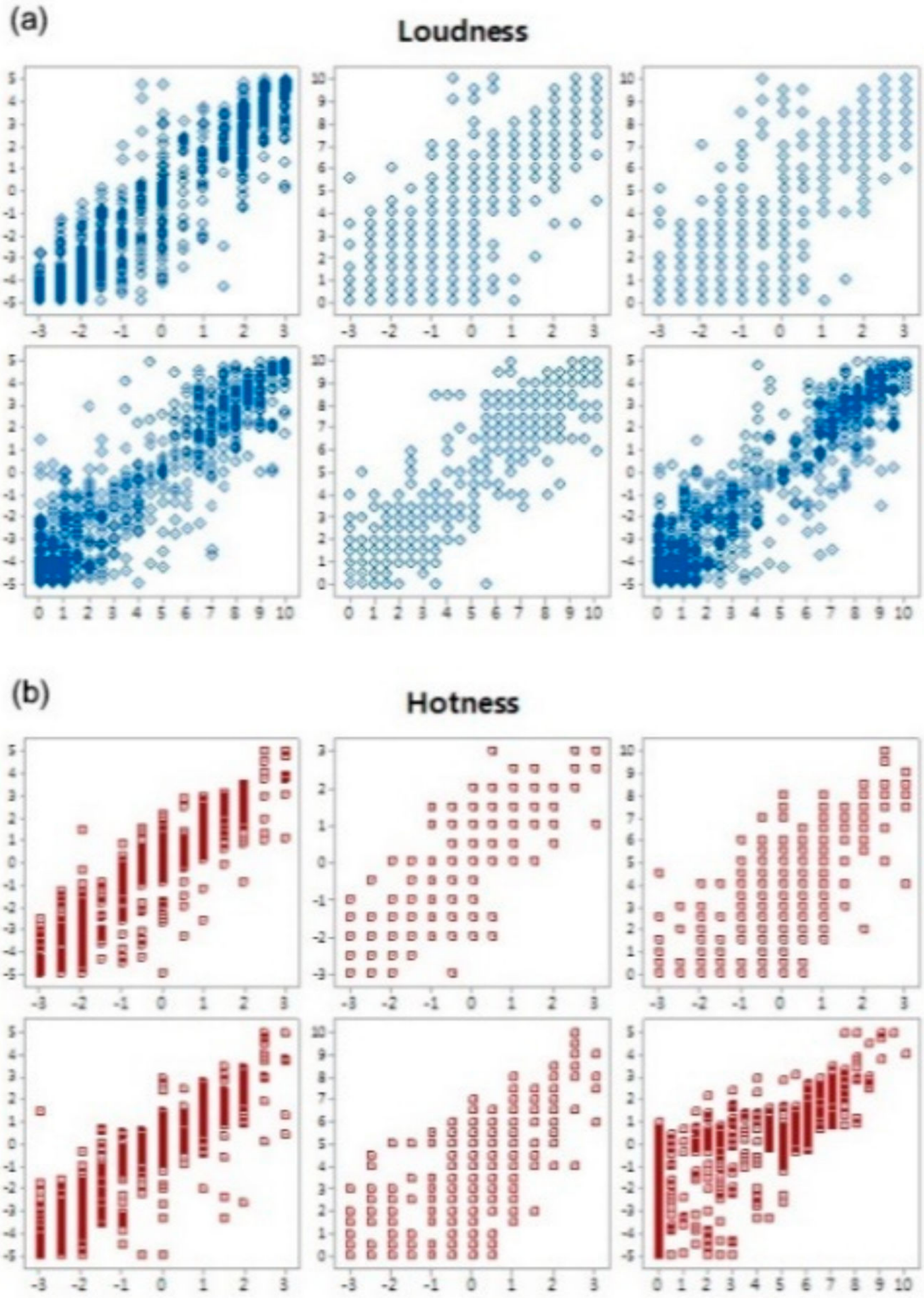

Figure 4. Scatter plots for (**a**) loudness and (**b**) hotness (top left: bipolar7 × bipolar VAS; top center: bipolar7 × combined; top right: bipolar7 × unipolar11; bottom left: combined × bipolar VAS; bottom center: combined × unipolar11; and bottom right: unipolar11 × bipolar VAS).

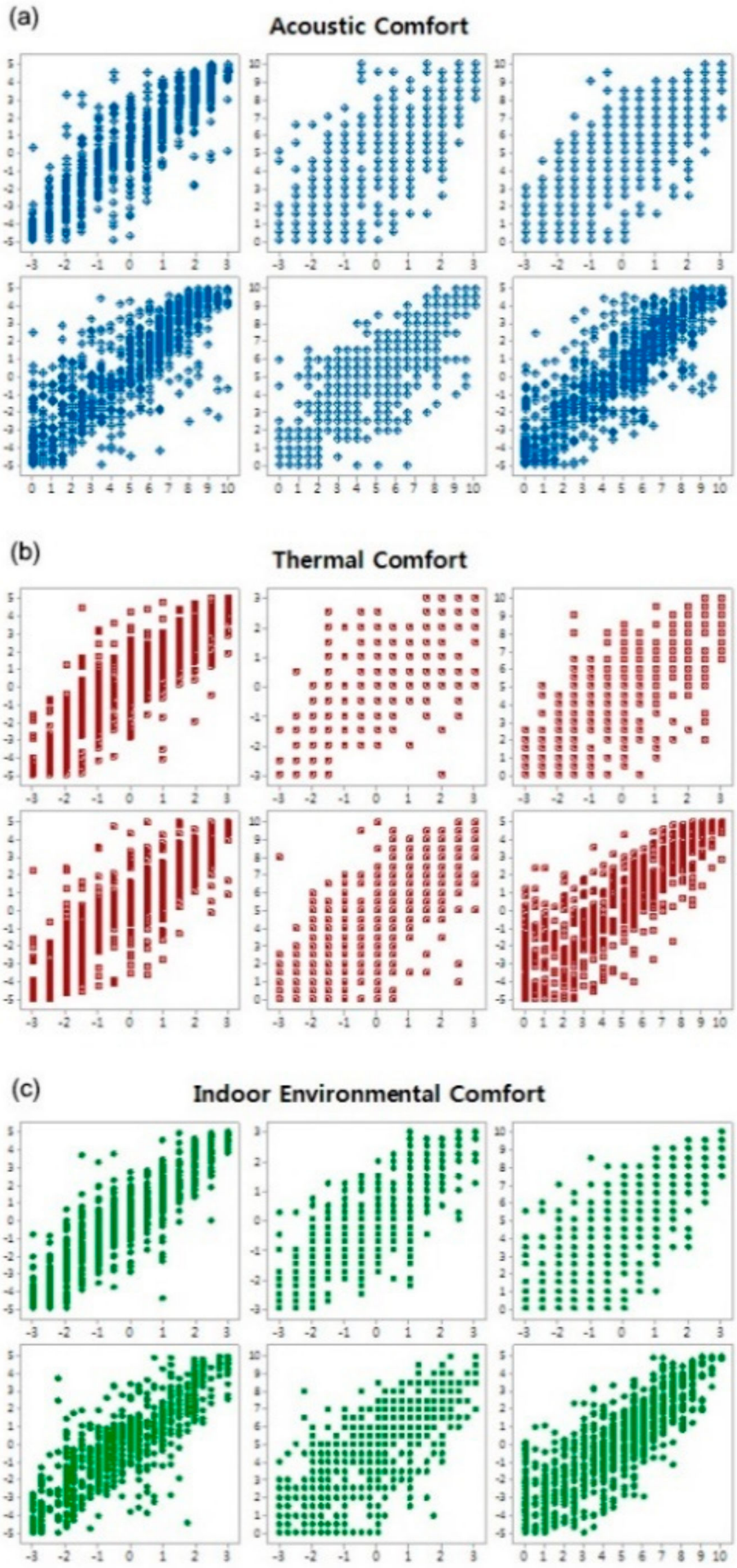

Figure 5. Scatter plots for (**a**) acoustic comfort, (**b**) thermal comfort, and (**c**) indoor environmental comfort (top left: bipolar7 × bipolar VAS; top center: bipolar7 × combined; top right: bipolar7 × unipolar11; bottom left: combined × bipolar VAS; bottom center: combined × unipolar11; and bottom right: unipolar11 × bipolar VAS).

3.3. Effects of Repeat and Response Scales

Table 8 lists the significance levels and effect sizes of the repeated-measures ANOVA results for normalized subjective responses. The effects of the response scales were found for softness, coldness, hotness, acoustic comfort, thermal discomfort, and indoor environmental comfort. This is shown in Figure 6. The Bonferroni post hoc test results for the scale effect are listed in Table 9. In general, the mean values from the bipolar VAS and unipolar11 were greater than the mean values from bipolar7.

Table 8. Results of significance level ($p < 0.05$) and effect size (η2) of repeated-measures ANOVA using normalized data (D: discomfort, C: comfort).

			Acoustic Sensation		Thermal Sensation		Acoustic		Thermal		Indoor Environmental	
			Soft	Loud	Cold	Hot	D	C	D	C	D	C
Within Subjects	Repeat	p	0.307	0.872	0.000	0.008	0.390	0.081	0.000	0.024	0.000	0.083
		η^2	0.000	0.000	0.007	0.004	0.000	0.001	0.012	0.003	0.017	0.000
	Repeat x Scale	p	0.304	0.139	0.092	0.819	0.013	0.508	0.849	0.029	0.333	0.868
		η^2	0.002	0.003	0.003	0.001	0.005	0.001	0.000	0.005	0.002	0.000
	Scale	p	0.000	0.927	0.003	0.000	0.830	0.003	0.010	0.811	0.076	0.000
		η^2	0.013	0.000	0.007	0.017	0.000	0.006	0.007	0.001	0.004	0.018
Between Subjects	Sound	p	0.000	0.000	0.295	0.574	0.000	0.000	0.095	0.554	0.000	0.000
		η^2	0.096	0.374	0.003	0.002	0.265	0.172	0.005	0.002	0.166	0.086
	Temperature	p	0.666	0.199	0.000	0.000	0.102	0.000	0.000	0.000	0.000	0.000
		η^2	0.000	0.002	0.446	0.292	0.002	0.011	0.133	0.071	0.034	0.035
	Sound x Scale	p	0.182	0.081	0.997	0.930	0.413	0.654	0.936	0.980	0.101	0.530
		η^2	0.007	0.009	0.001	0.003	0.006	0.004	0.003	0.002	0.010	0.006
	Temp. x Scale	p	0.943	0.986	0.000	0.000	0.802	0.433	0.000	0.001	0.000	0.054
		η^2	0.001	0.000	0.018	0.015	0.001	0.003	0.017	0.013	0.015	0.007

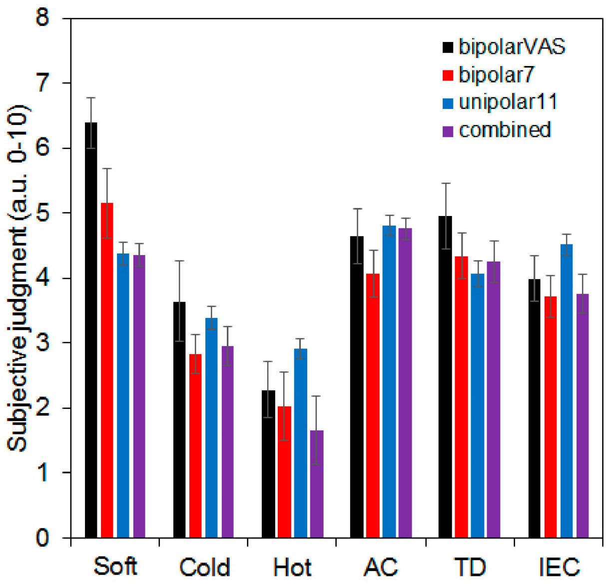

Figure 6. Normalized mean subjective judgment with 95% confidence intervals according to response scale (black: bipolar VAS; red: bipolar; blue: unipolar11; purple: combined).

The effects of repeats were found only in thermal-related attributes, as listed in Tables 8 and 10. Coldness, thermal discomfort, and indoor environmental discomfort were greater at the second measurement than at the first measurement. For thermal sensation, the second sensation was lower than the first sensation for all response scales except for hotness when the unipolar 11-point scale was used. For thermal comfort, the bipolar seven-point scale and the combined scale, which were identical scales, did not differentiate between the first and second measurements. With the bipolar VAS and unipolar 11-point scales, thermal comfort was rated poorer in the second measurement. For indoor environmental comfort, the mean discomfort using the unipolar 11-point scale was greater in the second measurement.

Table 9. Results of Bonferroni's post hoc test for normalized mean subjective judgment according to response scale (means that do not share a letter are significantly different; A > B > C in a column).

	Soft		**Cold**		**Hot**		**AC**		**TD**		**IEC**	
	p	η^2	p	η^2	p	η^2	p	η^2	p	η^2	p	η^2
	0.000	**0.036**	**0.003**	**0.007**	**0.003**	**0.008**	**0.003**	**0.006**	**0.010**	**0.007**	**0.000**	**0.018**
bipolar VAS	6.38	A	3.64	A	2.28	B	4.64	A	4.95	A	3.99	A
bipolar7	5.14	B	2.83	B	2.03	B	4.07	B	4.35	A	3.72	B
unipolar11	4.37	C	3.39	AB	2.92	A	4.81	A	4.07	AB	4.51	A
combined	4.35	C	2.96	AB	1.66	B	4.76	A	4.26	B	3.76	B
	unipolar11		bipolar7		bipolar7		unipolar11		bipolar7		bipolar7	

Table 10. Results of Bonferroni's post hoc test for repeated measures using original data (means that do not share a letter are significantly different; A > B in a row).

		Scale	p	**Repeat**			
				1st		**2nd**	
Acoustic sensation	Soft–Loud	BipolarVAS	0.321	−0.843	A	−0.669	A
		Bipolar7	0.874	−0.248	A	−0.231	A
	Soft	Unipolar11	0.506	4.315	A	4.429	A
	Loud	Unipolar11	0.907	4.008	A	3.987	A
	Soft	Combined	0.645	4.389	A	4.311	A
	Loud	Combined	0.340	3.845	A	4.019	A
Thermal sensation	Cold–Hot	BipolarVAS	0.017	−0.403	A	−0.698	B
		Bipolar7	0.002	−0.209	A	−0.456	B
	Cold	Unipolar11	0.044	3.208	B	3.563	A
	Hot	Unipolar11	0.051	3.064	A	2.768	A
	Cold–Hot	Combined	0.018	−0.223	A	−0.421	B
Acoustic comfort	D–C	BipolarVAS	0.976	0.406	A	0.402	A
		Bipolar7	0.187	0.099	A	−0.029	A
	D	Unipolar11	0.805	3.877	A	3.917	A
	C	Unipolar11	0.973	4.813	A	4.808	A
	D	Combined	0.303	3.727	A	3.893	A
	C	Combined	0.713	4.787	A	4.729	A
Thermal comfort	D–C	BipolarVAS	0.005	0.408	A	−0.021	B
		Bipolar7	0.170	0.068	A	−0.065	A
	D	Unipolar11	0.002	3.820	B	4.316	A
	C	Unipolar11	0.010	4.779	A	4.359	B
	D–C	Combined	0.058	0.076	A	−0.109	A
Indoor Environmental Comfort	D–C	BipolarVAS	0.100	0.148	A	−0.079	A
		Bipolar7	0.379	−0.132	A	−0.211	A
	D	Unipolar11	0.035	4.328	B	4.641	A
	C	Unipolar11	0.131	4.623	A	4.401	A
	D–C	Combined	0.246	−0.065	A	−0.168	A

3.4. Effects of Temperature and Sound

The effects of sound and temperature on acoustic, thermal, and indoor environmental comfort/discomfort are compared in Figure 7 using the normalized data. Although the relative values for each observational point were different with statistical significance, the overall trends according to temperature or sound type and level were similar.

The statistical comparisons using the original data were also analyzed in Table 11. No specific response scale with a high degree of differentiation for temperature or sound type and level was found for comfort/discomfort. The effects of temperature on thermal attributes, the effects of sound on acoustic attributes, and the cross-modal effects of temperature and sound on comfort attributes according to the four response scales used in this study were identical with statistical significance ($p < 0.0005$). Even though the acoustic comfort and discomfort of 42 dBA traffic noise were clearly distinguished from that of ambient noise by the four response scales, the loudness of the 42-dBA traffic noise was not differentiated from that of the ambient noise by any response scales. The softness of the 42 dBA traffic noise was distinguished from that of ambient noise by the unipolar 11-point scale only.

For the effects of temperature on thermal attributes, all response scales were able to differentiate thermal perception at each temperature. All of the response scales could also distinguish the difference in acoustic perception between 19.0 °C and 24.5 °C or 19.0 °C and 30.0 °C for the effects of temperature on acoustic attributes. For the effects of sound on acoustic attributes, all of the response scales could differentiate ambient sound, any sounds of 41 dBA, water sound of 61 dBA, and traffic noise of 61 dBA. For the effects of sound on thermal attributes, the thermal perception of 61 dBA traffic noise was different from the thermal perception of ambient noise for all response scales.

The effects of sound on indoor environmental attributes were the same for all response scales. The effects of temperature on indoor environmental attributes were the same for the bipolar seven-point, the unipolar 11-point, and the combined scales. The bipolar VAS did not differentiate indoor environmental comfort between 24.5 °C and 30.0 °C.

There was no statistically significant difference between water sounds and 42 dBA traffic noise in acoustic comfort and indoor environmental comfort. However, at 61 dBA, water sounds always showed positive aspects compared to traffic noise in acoustic comfort and indoor environmental comfort perception. Between 19.0 °C, 24.5 °C, and 30.0 °C, 19 °C was the least preferable temperature for all subjective attributes.

3.5. Respondents' Survey

For acoustic conditions, 68% of respondents preferred the bipolar VAS, and for thermal conditions, 52% of respondents preferred the bipolar VAS. For combined conditions, 48% of respondents voted the bipolar VAS as the preferred scale, and the combined scale was ranked as the second preference of the respondents, as shown in Figure 8. Figure 9 shows reasons why a particular rating scale was voted for. Most respondents stated that the bipolar VAS allowed them to express their feelings adequately.

The mean indoor environmental sensitivities that were subjectively voted for were 1 ± 0.5, which means the respondents were slightly sensitive to acoustics, lighting, temperature, humidity, and indoor air quality. The respondents in this study were not particularly demanding with regards to thermal or acoustic conditions (Figure 10).

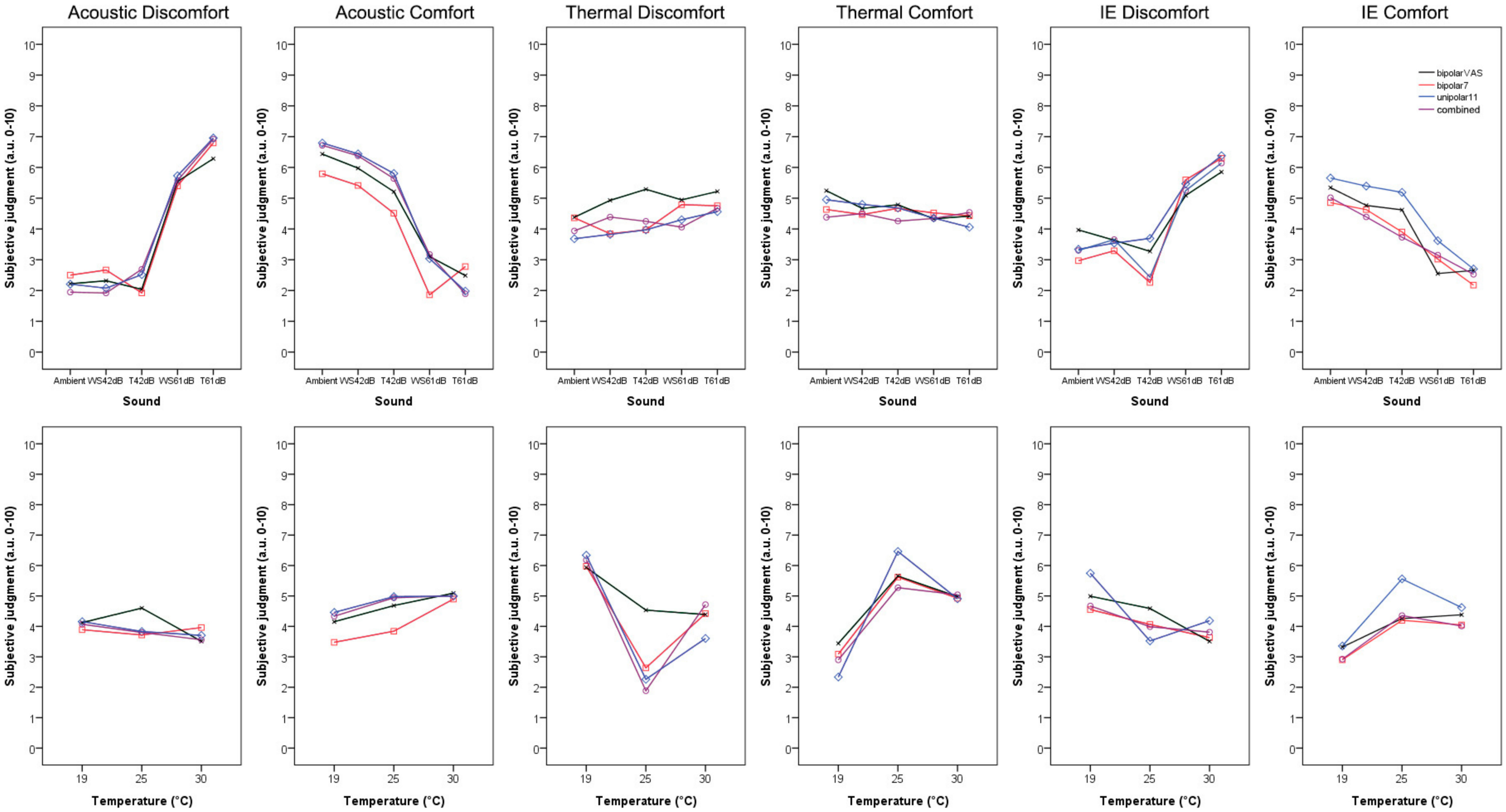

Figure 7. Normalized mean subjective judgment according to sound and temperature (black: bipolar vas; red: bipolar7; blue: unipolar11; and purple: combined).

Table 11. Results of Bonferroni's post hoc test for the original mean subjective judgment according to sound and temperature (means that do not share a letter are significantly different in a cell. A > B > C > D in a column; cross-modal effects are shown in bold).

	Acoustic Sensation						**Thermal Sensation**					**Acoustic Comfort**						**Thermal Comfort**					**Indoor Environmental**				
	Soft		**Soft–Loud**		**Loud**		**Cold**	**Cold–Hot**			**Hot**	**D**		**D–C**		**C**		**D**	**D–C**		**C**	**D–C**	**D**	**D–C**		**C**	**D–C**
	Unipolar11	**Combined (Unipolar11)**	**Bipolar VAS**	**Bipolar7**	**Unipolar11**	**Combined (Unipolar11)**	**Combined (Unipolar11)**	**Bipolar VAS**	**Bipolar7**	**Unipolar11**	**Combined (Unipolar11)**	**Unipolar11**	**Combined (Unipolar11)**	**Bipolar VAS**	**Bipolar7**	**Unipolar11**	**Combined (Unipolar11)**	**Unipolar11**	**Bipolar VAS**	**Bipolar7**	**Unipolar11**	**Combined (bipolar7)**	**Unipolar11**	**Bipolar VAS**	**Bipolar7**	**Unipolar11**	**Combined (bipolar7)**
Ambient	A	A	C	C	B	C	A	A	A	A	A	C	D	A	A	A	A	**B**	**A**	**A**	**A**	**A**	C	A	A	A	A
WS42dB	A	A	C	C	B	C	A	A	A	A	A	C	D	AB	AB	AB	A	**AB**	**AB**	**AB**	**AB**	**AB**	C	A	A	A	A
T42dB	A	B	C	C	B	C	A	A	A	A	A	C	C	B	B	B	B	**AB**	**A**	**AB**	**AB**	**AB**	C	A	A	A	A
WS61dB	B	C	B	B	A	B	A	A	A	A	A	B	B	C	C	C	C	**AB**	**AB**	**AB**	**AB**	**AB**	B	B	B	B	B
T61dB	C	D	A	A	A	A	A	A	A	A	A	A	A	D	D	D	D	**A**	**B**	**B**	**B**	**B**	A	C	C	C	C
19.0 °C	A	A	A	A	A	A	A	C	C	C	C	A	**A**	**B**	**B**	**B**	**B**	A	C	C	C	C	A	B	C	C	C
24.5 °C	A	A	A	A	A	A	B	B	B	B	B	A	**AB**	**A**	**AB**	**A**	**A**	C	A	A	A	A	C	A	A	A	A
30.0 °C	A	A	A	A	A	A	C	A	A	A	A	A	**B**	**A**	**A**	**A**	**A**	B	B	B	B	B	B	A	B	B	B

Note: D = discomfort, C = comfort.

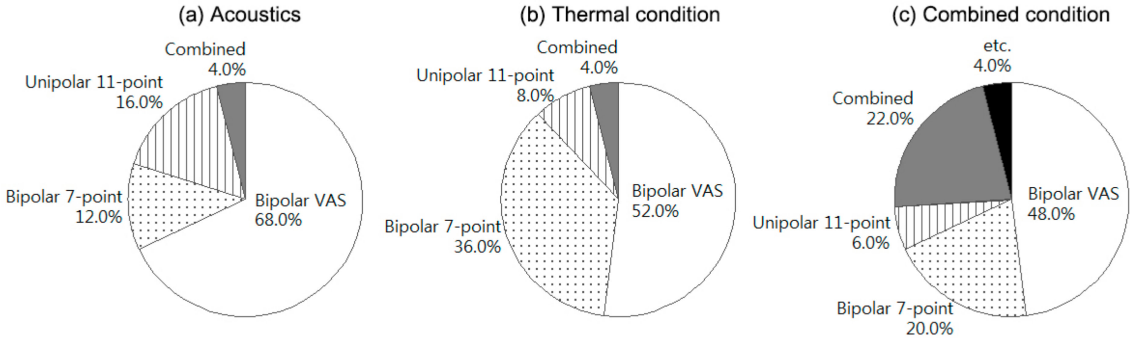

Figure 8. Respondents' preferred response scales.

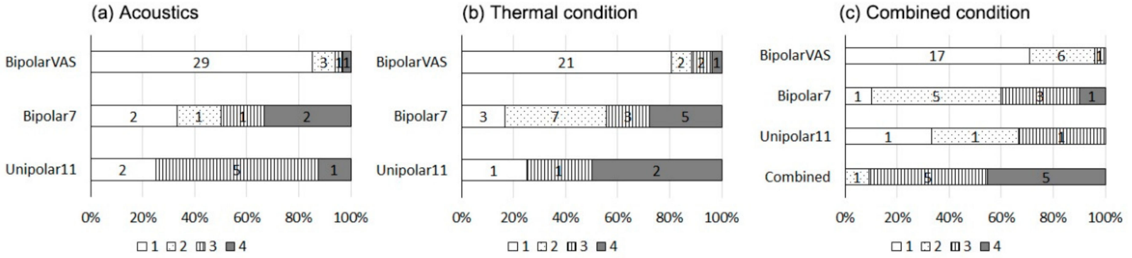

Figure 9. Reasons why a scale was voted for 1: "allowed you to express your feeling adequately"; 2: "ease of use"; 3: "allowed you to rate in detail"; 4: etc. (The number in the bars is the frequency).

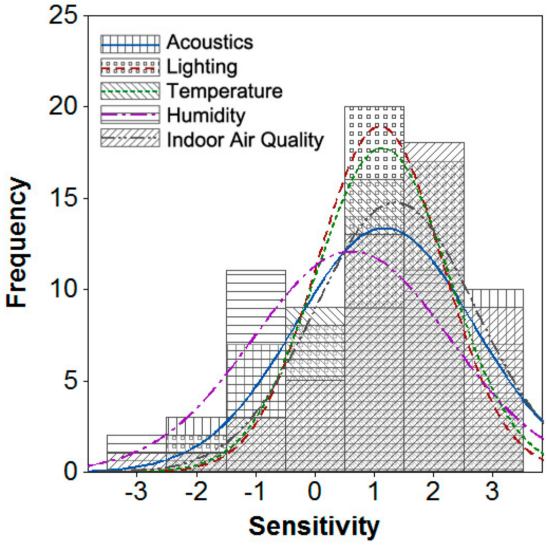

Figure 10. Sensitivity votes for indoor environmental factors (−3: very insensitive; −2: somewhat insensitive; −1: slightly insensitive; 0: neither insensitive nor sensitive; 1: slightly sensitive; 2: somewhat sensitive; and 3: very sensitive).

4. Discussion

4.1. Reliability (Duplicate Sample Analysis)

Reliability, defined as the ability of the measuring instrument to give the same value over repeated measurements, has long been a criterion in any quantitative sensory method. The most reliable response scale was the unipolar 11-point scale for loudness as listed in Table 5. For thermal sensation, the unipolar 11-point scale was also reliable, as was the bipolar VAS. For acoustic comfort attributes, both the bipolar VAS and the unipolar 11-point scale ranked in a top reliability group. The most reliable response scale for thermal comfort was the bipolar VAS. For indoor environmental comfort attributes, the combined scale, which is the bipolar seven-point scale, was the most reliable scale. However, the correlation coefficients of the bipolar seven-point scale and the combined scale, which are identical scales, were significantly different for indoor environmental comfort attributes. Thus, the bipolar seven-point scale may not be a reliable measure to assess overall indoor environmental comfort.

The bipolar VAS seemed relatively reliable over repeated measures for all subjective attributes tested in this study, even though it did not always have the highest reliability. Rausch and Zehetleitner [47] also reported that Cronbach's alphas for internal consistency were larger for the VAS than for the discrete scale at four out of six levels of coherence in their conscious experience of motion. However, no clear evidence that the bipolar VAS has the highest reliability has been found yet.

In previous studies in other disciplines, no statistical reliability difference was observed. Lawless et al. [50] found, in a study on food quality, that the labeled magnitude scale, nine-point, and 11-point scales had no obvious advantage in reliability over another when using Fisher's Z transformation with $p < 0.05$. The reliability coefficients of the seven-point, 11-point, and VAS reported by Lewis and Erdinc [51] exceeded 0.8. The researchers also reported that the reliability coefficients of the full and discretized VAS and discrete scale were greater than 0.8.

A prerequisite for the reliability is to maintain homogeneous physical (thermal and acoustical) conditions in the test laboratory during all the test sessions. The actual room air temperature varied by 0.2–0.4 °C depending on the target temperature as listed in Table 3, owing to the systematic characteristics of the variable refrigerant flow systems. However, the temperature variations could not explain the lowered thermal sensation values, because the temperature variations have been continuously repeated to maintain the target temperature automatically. Changes in thermal sensation might result from a thermal adaptation time or its sensitive characteristics. Furthermore, 20 min may not be sufficient for thermal adaptation. Even though it was sufficient for thermal adaption, thermal sensation might not be constant for 60 minutes.

4.2. Sensitivity (Degree of Differentiation by Indoor Physical Factors)

The sensitivity of a measurement was defined by Scott and Huskisson [52] as the number of units of change for a given force applied. The degree of differentiation for the indoor physical factors was not statistically different for any response scales as listed in Table 11. Thus, the sensitivity of the response scales used in this study exhibited no statistical difference. Any response scale can be used to test the combined effects of thermal and acoustic conditions on indoor environmental human perception relatively. This is consistent with previous studies in various research fields [47,50,51,53–57].

Bolognese et al. [53] found that 100-mm VAS and five-point Likert responses were highly correlated and yielded similar precision for discriminating treatments in osteoarthritis patients. Couper et al. [54] compared a VAS with variations of 20-point scales for a web-based social survey, and found no evidence for the differences between the VAS and alternative approaches. Van Schaik and Ling [55] found in a web-based rating scale comparison study that the VAS and seven-point Likert scale resulted in essentially the same psychometric properties of scales. Davey et al. [56] suggested that either a five-point Likert response scale or a 100-mm VAS could adequately measure anxiety. Lawless et al. [50] compared a labeled affective magnitude scale, an 11-point category scale, and a nine-point category scale for assessing food likes and dislikes in a survey. They found that all three scales performed equally well, with no one scale showing a consistent superiority over another.

Lee et al. [57] reported conclusive results in their thermal sensation assessment, which compared nine-point categorical scales and VAS. The researchers could not assert that this is an optimal scale for the measurement of perceived thermal sensation at this time. Rausch and Zehetleitner [47] reported that both VAS and discrete scales were reliable measures of the subjective perception of the global motion experience. Lewis and Erdinc [51] concluded in the context of subjective usability research that they could not find any particular measurement advantage associated with the use of bipolar VAS, bipolar seven-point, unipolar 11-point, or combined scales.

Thus far, the sensitivities of the response scales in combined thermal and acoustic conditions was discussed that sensitivity of all response scales is not statistically different. However, sensitivity in thermal comfort differentiation seems to differ in the response scales. Since the changes in thermal sensation due to repeated measures occurred with all the response scales in Table 10, we could assume that the changes in thermal sensation were true. If changes in thermal sensation can cause changes in thermal comfort, no changes in thermal comfort using the bipolar seven-point scales including the combined scale imply lower sensitivity of bipolar seven-point scales in thermal comfort differentiation according to thermal sensation. The lack of a seven-point thermal sensation scale has been reported in many cases [58–61]. A 13-value thermal comfort scale provided a more accurate evaluation of thermal sensation. [62] A recent study has reported that sensitivity in subjective acoustic differentiation using the unipolar 11-point numerical scale was greater than that obtained using the bipolar VAS [43]. However, they did not verify what caused the difference between the two scales. For acoustics sensation and comfort, no obvious advantage of one response scale over another was found in this study.

4.3. Validity in Use for Indoor Environmental Sensation and Perception Assessment

To the best of our knowledge, this is the first study to compare the visual analogue scale, bipolar seven-point scale, and unipolar 11-point scale to assess indoor environmental sensation and perception in combined thermal and acoustic conditions.

For the sensation level, no cross-modal effects were observed for any response scales. At the perception level, effects of sound sources on thermal comfort and the effects of temperature on acoustic comfort were found for all response scales used in this study. The cross-modal effects of thermal and acoustic conditions on subjective judgment depended on the level of environmental sensation and perception. This is consistent with Yang and Moon [15] in their recent study using a bipolar visual analogue scale.

For the sensation level, the existing standard response scales [5,6] for each environmental factor were highly correlated, as listed in Table 6. The highest correlation was shown with the unipolar 11-point scale and the combined scale for softness and loudness. For coldness and hotness, the bipolar seven-point scale and the combined scale were highly correlated. They were basically the same scale, so the results were expected. However, at the perception level, the highest correlation was observed for the visual analogue scale and the bipolar seven-point scale for all comfort/discomfort attributes as listed in Table 7. No differences in response scales existed for thermal and acoustic comfort.

Sensation is a mental process resulting from the immediate external stimulation of a sense organ, and perception is the awareness of elements of the environment through physical sensation [63]. Each sensation may need a unique evaluation method based on the sensory organs. However, at the perception level, characteristics of sensory organs might be neutralized through a more complicated mental process. Schweiker et al. [64] commented that a seven-point thermal sensation scale is suitable for describing a one-dimensional relationship between the physical parameters of indoor environments and subjective thermal sensation; however, human thermal comfort is not merely physiological; it is also a psychological multidimensional conceptualization. We have not found studies regarding the impact of scale on sensation and perception. Thus, more research is required to confirm its relevance.

4.4. Unipolar and Bipolar

ISO 10551:1995 [5] notes that a bipolar scale is useful for taking thermal perception into account and is more sensitive than the unipolar scale in the region of thermal conditions near thermal neutrality. By contrast, ISO/TS 15666:2003 [6] recommends a unipolar scale for socioacoustic surveys because reactions to noise are overwhelmingly either negative or neutral. Furthermore, acoustic intensity sensation has no neutrality between softness and loudness. If the nonexistence of cold or heat sensation is thermal neutral, the nonexistence of soft and loud sensation is the nonexistence of an acoustic stimulus. This is a critical difference between thermal and acoustic environments, which affects the selection of scale polarity in subjective assessments. However, comfort attributes do not have polarity differences according to sensory characteristics.

In the present study, both bipolar and unipolar scales were used to assess subjective thermal and acoustic attributes, although this was not a direct polarity comparison with an identical scale type and length. The unipolar scale had the highest reliability for soft and loud sensations, and the bipolar scale had the highest reliability for cold and hot sensations. The unipolar scale had the lowest reliabilities in the thermal sensation assessment. Furthermore, the unipolar values for the bipolar left end, which were soft and uncomfortable attributes, had relatively lower reliabilities than any other paired comparisons. Thus, subjective attributes for unipolar scales should be carefully chosen.

On the other hand, bipolar scales have a disadvantage in that their reliability is somewhat lower on bipolar scales than unipolar scales. [65] The impact of scale polarity on data quality has not been fully investigated yet.

4.5. Respondent Preferences

The visual analogue scale was preferred in acoustic, thermal, and combined conditions in this study with young collegians because it allowed the study participants to express their feelings adequately. Research information about the respondents' preferences with regard to rating scales for combined environmental conditions is scant. In medicine or psychology, respondent preference has been investigated for reasons specific to the field of research For pain scale studies in medicine, socioeconomic educational factors were important in choosing a scale, and in this respect, VAS was found not to be a priority of scale preference [66–68]. Preston and Colman [69] reported respondent preference results in depth. Longer scales tended to receive more favorable ratings on "allowed you to express your feelings adequately." For "ease of use," scales with six, seven, and 10 response categories were the most preferred, and scales with 11 and 101 response categories were rated as least easy to use.

Respondent preference could be a factor determining the selection of a rating scale, considering a positive association between the users' performance and their subjectively expressed preferences was obvious in the meta-analysis [70]. However, more research needs to be done on how respondent preference factors apply to selecting rating scales.

4.6. Limitations

First, the sample was non-random, comprising only young, educated participants. However, the non-randomized sample was justified by the study purpose. Second, the present findings can be attributed to indoor environmental configurations. As the observational points of the thermal and acoustic physical factors increase, the performance and preference of the response scales might change.

5. Conclusions

Four response scales were compared to assess subjective indoor environmental sensation and perception in combined thermal and acoustic conditions with university students in their early twenties. The results revealed that the bipolar visual analogue scale was subjectively preferred in all conditions. On the other hand, the performances of the response scales were not significantly different for young university students.

The four response scales exhibited no significant differences in sensitivity, the degree of relative differentiation based on indoor physical factors. Thus, the effects of physical factors on human response could be assessed using any of the four scales tested in this study with a statistical significance at $p < 0.05$ in moderate environments. However, the normalized means were dependent on the response scales for each subjective attribute. Furthermore, the reliability of each response scale was different according to the subjective attributes. The existing standard scales (bipolar seven-point scale for thermal assessment and unipolar 11-point scale for acoustic assessment) were the most reliable for the sensation assessment of each stimulus. For comfort attributes including acoustic, thermal, and overall comfort, the bipolar VAS and the bipolar seven-point scale had the highest correlation coefficients. The bipolar seven-point scale was not sensitive for thermal comfort differentiation, even though it was the most reliable for thermal sensation assessment. A choice of response scale could depend not only on the type of physical stimulus but also on the question of sensation or perception, which is the process of sensing or interpreting.

Author Contributions: Conceptualization, W.Y.; Funding acquisition, J.Y.J.; Resources, H.J.M.; Writing—original draft, W.Y.

Funding: This work was supported by the Basic Science Research Program of the National Research Foundation (NRF) (grant number 2018R1D1A1B07048157) funded by the Ministry of Education, Republic of Korea. This work was also supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20172010000580).

Acknowledgments: We gratefully acknowledge the contributions of all participants who volunteered in the experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. DeCastellarnau, A. A classification of response scale characteristics that affect data quality: A literature review. *Qual. Quant.* **2018**, *52*, 1523–1559. [CrossRef] [PubMed]
2. Andrews, F.M. Construct validity and error components of survey measures: A structural modeling approach. *Public Opin. Q.* **1984**, *48*, 409–442. [CrossRef]
3. Rodgers, W.L.; Andrews, F.M.; Regula Herzog, A. Quality of survey measures: A structural modeling approach. *J. Off. Stat.* **1992**, *8*, 251.
4. Krosnick, J.A.; Fabrigar, L.R. Designing Rating Scales for Effective Measurement in Surveys. In *Survey Measurement and Process Quality*; Lyberg, L., Biemer, P., Collins, M., Leeuw, E.D., Dippo, C., Schwarz, N., Trewin, D., Eds.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 1997.
5. Standardization IOF. *Ergonomics of the Thermal Environment—Assessment of the Influence of the Thermal Environment Using Subjective Judgement Scales*; ISO 10551: 1995; ISO: Geneva, Switzerland, 1995.
6. Standardization IOF. *Acoustics—Assessment of Noise Annoyance by Means of Social and Socio-Acoustic Surveys*; ISO/TS 15666: 2003; ISO: Geneva, Switzerland, 2003.
7. Fanger, P.O. *Thermal Comfort. Analysis and Applications in Environmental Engineering*; Danish Technical Press: Copenhagen, Denmark, 1970; p. 244.
8. Standardization CECF. *Indoor Environmental Input Parameters for Design Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics*; EN 15251: 2007; European Committee for Standardization: Brussels, Belgium, 2007.
9. Nagano, K.; Horikoshi, T. New index of combined effect of temperature and noise on human comfort: summer experiments on hot ambient temperature and traffic noise. *Arch. Complex Environ. Stud.* **2001**, *13*.
10. Pellerin, N.; Candas, V. Effects of steady-state noise and temperature conditions on environmental perception and acceptability. *Indoor Air* **2004**, *14*, 129–136. [CrossRef]
11. Witterseh, T.; Wyon, D.P.; Clausen, G. The effects of moderate heat stress and open-plan office noise distraction on SBS symptoms and on the performance of office work. *Indoor Air* **2004**, *14*, 30–40. [CrossRef]
12. Nagano, K.; Horikoshi, T. New comfort index during combined conditions of moderate low ambient temperature and traffic noise. *Energy Build.* **2005**, *37*, 287–294. [CrossRef]
13. Tiller, D.K.; Wang, L.M.; Musser, A.; Radik, M. Combined effects of noise and temperature on human comfort and performance. *ASHRAE Trans.* **2010**, *116*, 522–540.
14. Yang, W.; Moon, H.J.; Kim, M.-J. Combined effects of short-term noise exposure and hygrothermal conditions on indoor environmental perceptions. *Indoor Built Environ.* **2018**, *27*, 1119–1133. [CrossRef]
15. Yang, W.; Moon, H.J. Cross-modal effects of noise and thermal conditions on indoor environmental perception and speech recognition. *Appl. Acoust.* **2018**, *141*, 1–8. [CrossRef]
16. Reips, U.-D.; Funke, F. Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator. *J. Behav. Res. Methods* **2008**, *40*, 699–704. [CrossRef]
17. Bond, A.; Lader, M. The use of analogue scales in rating subjective feelings. *Br. J. Med. Psychol.* **1974**, *47*, 211–218. [CrossRef]
18. Yang, W.; Moon, H.J. Combined effects of acoustic, thermal, and illumination conditions on the comfort of discrete senses and overall indoor environment. *Build. Environ.* **2019**, *148*, 623–633. [CrossRef]

19. Humphreys, M.A. Quantifying occupant comfort: Are combined indices of the indoor environment practicable? *Build. Res. Inf.* **2005**, *33*, 317–325. [CrossRef]
20. Andersen, R.V.; Toftum, J.; Andersen, K.K.; Olesen, B.W. Survey of occupant behaviour and control of indoor environment in Danish dwellings. *Energy Build.* **2009**, *41*, 11–16. [CrossRef]
21. Li, Q.; You, R.; Chen, C.; Yang, X. A field investigation and comparative study of indoor environmental quality in heritage Chinese rural buildings with thick rammed earth wall. *Energy Build.* **2013**, *62*, 286–293. [CrossRef]
22. Mui, K.W.; Tsang, T.W.; Wong, L.T.; William Yu, Y.P. Evaluation of an indoor environmental quality model for very small residential units. *Indoor Built Environ.* **2018**, *28*, 470–478. [CrossRef]
23. Paul, W.L.; Taylor, P.A. A comparison of occupant comfort and satisfaction between a green building and a conventional building. *Build. Environ.* **2008**, *43*, 1858–1870. [CrossRef]
24. Bluyssen, P.M.; Aries, M.; van Dommelen, P. Comfort of workers in office buildings: The European HOPE project. *Build. Environ.* **2011**, *46*, 280–288. [CrossRef]
25. Frontczak, M.; Schiavon, S.; Goins, J.; Arens, E.; Zhang, H.; Wargocki, P. Quantitative relationships between occupant satisfaction and satisfaction aspects of indoor environmental quality and building design. *Indoor Air* **2012**, *22*, 119–131. [CrossRef]
26. Hwang, T.; Kim, J.T. Assessment of Indoor Environmental Quality in Open-Plan Offices. *Indoor Built Environ.* **2013**, *22*, 139–156. [CrossRef]
27. Fassio, F.; Fanchiotti, A.; Vollaro, R. Linear, Non-Linear and Alternative Algorithms in the Correlation of IEQ Factors with Global Comfort: A Case Study. *Sustainability* **2014**, *6*, 8113–8127. [CrossRef]
28. Liang, H.-H.; Chen, C.-P.; Hwang, R.-L.; Shih, W.-M.; Lo, S.-C.; Liao, H.-Y. Satisfaction of occupants toward indoor environment quality of certified green office buildings in Taiwan. *Build. Environ.* **2014**, *72*, 232–242. [CrossRef]
29. Woo, J. A systematic post-occupancy evaluation in green-rated high-rise office buildings. In Proceedings of the Across: Architectural Research through to Practice: 48th International Conference of the Architectural Science Association 2014, Genoa, Italy, 10–13 December 2014; Madeo, F., Ed.; Genova University Press: Genoa, Italy, 2014.
30. Pei, Z.; Lin, B.; Liu, Y.; Zhu, Y. Comparative study on the indoor environment quality of green office buildings in China with a long-term field measurement and investigation. *Build. Environ.* **2015**, *84*, 80–88. [CrossRef]
31. Ravindu, S.; Rameezdeen, R.; Zuo, J.; Zhou, Z.; Chandratilake, R. Indoor environment quality of green buildings: Case study of an LEED platinum certified factory in a warm humid tropical climate. *Build. Environ.* **2015**, *84*, 105–113. [CrossRef]
32. Martellotta, F.; Simone, A.; Della Crociata, S.; D'Alba, M. Global comfort and indoor environment quality attributes for workers of a hypermarket in Southern Italy. *Build. Environ.* **2016**, *95*, 355–364. [CrossRef]
33. Xue, P.; Mak, C.M.; Ai, Z.T. A structured approach to overall environmental satisfaction in high-rise residential buildings. *Energy Build.* **2016**, *116*, 181–189. [CrossRef]
34. Karmann, C.; Schiavon, S.; Graham, L.T.; Raftery, P.; Bauman, F. Comparing temperature and acoustic satisfaction in 60 radiant and all-air buildings. *Build. Environ.* **2017**, *126*, 431–441. [CrossRef]
35. Choi, J.-H.; Moon, J. Impacts of human and spatial factors on user satisfaction in office environments. *Build. Environ.* **2017**, *114*, 23–35. [CrossRef]
36. Wong, L.T.; Mui, K.W.; Hui, P.S. A multivariate-logistic model for acceptance of indoor environmental quality (IEQ) in offices. *Build. Environ.* **2008**, *43*, 1–6. [CrossRef]
37. Lai, A.C.K.; Mui, K.W.; Wong, L.T.; Law, L.Y. An evaluation model for indoor environmental quality (IEQ) acceptance in residential buildings. *Energy Build.* **2009**, *41*, 930–936. [CrossRef]
38. Huang, L.; Zhu, Y.; Ouyang, Q.; Cao, B. A study on the effects of thermal, luminous, and acoustic environments on indoor environmental comfort in offices. *Build. Environ.* **2012**, *49*, 304–309. [CrossRef]
39. Ricciardi, P.; Buratti, C. Environmental quality of university classrooms: Subjective and objective evaluation of the thermal, acoustic, and lighting comfort conditions. *Build. Environ.* **2018**, *127*, 23–36. [CrossRef]
40. Wetzel, E.; Greiff, S. The World Beyond Rating Scales. *Eur. J. Psychol. Assess.* **2018**, *34*, 1–5. [CrossRef]
41. Standard, A. *Thermal Environmental Conditions for Human Occupancy*; Standard 55-2004; ASHRAE: Atlanta, GA, USA, 2004; pp. 9–11.

42. Jones, L.A.; Berris, M. (Eds.) The psychophysics of temperature perception and thermal-interface design. In Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2002 HAPTICS 2002, Orlando, FL, USA, 24–25 March 2002.
43. Yang, W.; Jeon, J.Y. Performance and preference of response scales for semantic differentials in auditory perception among university students. *Can. Acoust.* **2019**, *47*, 41–49.
44. Kraus, U.; Breitner, S.; Hampel, R.; Wolf, K.; Cyrys, J.; Geruschkat, U.; Gu, J.; Radon, K.; Peters, A.; Schneider, A. Individual daytime noise exposure in different microenvironments. *Environ. Res.* **2015**, *140*, 479–487. [CrossRef]
45. Pitkanen, P. Stream Small Water Flow 01 XY. 2018. Available online: https://www.soundsnap.com (accessed on 17 September 2018).
46. Funke, F. A Web Experiment Showing Negative Effects of Slider Scales Compared to Visual Analogue Scales and Radio Button Scales. *Soc. Sci. Comput. Rev.* **2016**, *34*, 244–254. [CrossRef]
47. Rausch, M.; Zehetleitner, M. A comparison between a visual analogue scale and a four point scale as measures of conscious experience of motion. *Conscious. Cogn.* **2014**, *28*, 126–140. [CrossRef]
48. Budescu, D.V.; Appelbaum, M.I. Variance Stabilizing Transformations and the Power of the F Test. *J. Educ. Stat.* **1981**, *6*, 55–74. [CrossRef]
49. Box, G.E.P. Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification. *Ann. Math. Stat.* **1954**, *25*, 484–498. [CrossRef]
50. Lawless, H.T.; Popper, R.; Kroll, B.J. A comparison of the labeled magnitude (LAM) scale, an 11-point category scale and the traditional 9-point hedonic scale. *Food Qual. Prefer.* **2010**, *21*, 4–12. [CrossRef]
51. Lewis, J.R.; Erdinç, O. User experience rating scales with 7, 11, or 101 points: does it matter? *J. Usability Stud.* **2017**, *12*, 73–91.
52. Scott, J.; Huskisson, E.C. Graphic representation of pain. *Pain* **1976**, *2*, 175–184. [CrossRef]
53. Bolognese, J.A.; Schnitzer, T.J.; Ehrich, E.W. Response relationship of VAS and Likert scales in osteoarthritis efficacy measurement. *Osteoarthr. Cartil.* **2003**, *11*, 499–507. [CrossRef]
54. Couper, M.P.; Tourangeau, R.; Conrad, F.G.; Singer, E. Evaluating the Effectiveness of Visual Analog Scales: A Web Experiment. *Soc. Sci. Comput. Rev.* **2006**, *24*, 227–245. [CrossRef]
55. Schaik, P.V.; Ling, J. Design parameters of rating scales for web sites. *ACM Trans. Comput.-Hum. Interact.* **2007**, *14*, 4. [CrossRef]
56. Davey, H.M.; Barratt, A.L.; Butow, P.N.; Deeks, J.J. A one-item question with a Likert or Visual Analog Scale adequately measured current anxiety. *J. Clin. Epidemiol.* **2007**, *60*, 356–360. [CrossRef] [PubMed]
57. Lee, J.-Y.; Stone, E.A.; Wakabayashi, H.; Tochihara, Y. Issues in combining the categorical and visual analog scale for the assessment of perceived thermal sensation: Methodological and conceptual considerations. *Appl. Ergon.* **2010**, *41*, 282–290. [CrossRef]
58. Singh, M.K.; Mahapatra, S.; Atreya, S.K. Adaptive thermal comfort model for different climatic zones of North-East India. *Appl. Energy* **2011**, *88*, 2420–2428. [CrossRef]
59. Nguyen, A.T.; Singh, M.K.; Reiter, S. An adaptive thermal comfort model for hot humid South-East Asia. *Build. Environ.* **2012**, *56*, 291–300. [CrossRef]
60. Taleghani, M.; Tenpierik, M.; Kurvers, S.; van den Dobbelsteen, A. A review into thermal comfort in buildings. *Renew. Sustain. Energy Rev.* **2013**, *26*, 201–215. [CrossRef]
61. Indraganti, M.; Ooka, R.; Rijal, H.B.; Brager, G.S. Adaptive model of thermal comfort for offices in hot and humid climates of India. *Build. Environ.* **2014**, *74*, 39–53. [CrossRef]
62. Buratti, C.; Palladino, D.; Ricciardi, P. Application of a new 13-value thermal comfort scale to moderate environments. *Appl. Energy* **2016**, *180*, 859–866. [CrossRef]
63. Goldstein, E.B.; Brockmole, J. *Sensation and Perception*; Wadsworth Cengage Learning: Belmont CA, USA, 2016.
64. Schweiker, M.; Fuchs, X.; Becker, S.; Shukuya, M.; Dovjak, M.; Hawighorst, M.; Kolarik, J. Challenging the assumptions for thermal sensation scales. *Build. Res. Inf.* **2017**, *45*, 572–589. [CrossRef]
65. Alwin, D.F. *Margins of Error: A Study of Reliability in Survey Measurement*; John Wiley & Sons: Hoboken, NJ, USA, 2007.

66. Hjermstad, M.J.; Fayers, P.M.; Haugen, D.F.; Caraceni, A.; Hanks, G.W.; Loge, J.H.; Fainsinger, R.; Aass, N.; Kaasa, S. European Palliative Care Research Collaborative. Studies Comparing Numerical Rating Scales, Verbal Rating Scales, and Visual Analogue Scales for Assessment of Pain Intensity in Adults: A Systematic Literature Review. *J. Pain Symptom Manag.* **2011**, *41*, 1073–1093. [CrossRef] [PubMed]
67. Akad, K.; Solmaz, D.; Sari, I.; Onen, F.; Akkoc, N.; Akar, S. Performance of response scales of activity and functional measures of ankylosing spondylitis: numerical rating scale versus visual analog scale. *Rheumatol. Int.* **2013**, *33*, 2617–2623. [CrossRef] [PubMed]
68. Yazici Sayin, Y.; Akyolcu, N. Comparison of Pain Scale Preferences and Pain Intensity According to Pain Scales among Turkish Patients: A Descriptive Study. *Pain Manag. Nurs.* **2014**, *15*, 156–164. [CrossRef] [PubMed]
69. Preston, C.C.; Colman, A.M. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychol.* **2000**, *104*, 1–15. [CrossRef]
70. Nielsen, J.; Levy, J. Measuring usability: preference vs. performance. *J. Commun. ACM* **1994**, *37*, 66–75. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).