Taylor & Francis
Taylor & Francis Group

Architectural Engineering and Design Management

ISSN: 1745-2007 (Print) 1752-7589 (Online) Journal homepage: www.tandfonline.com/journals/taem20

Deconstructing the constructs in multi-domain investigations of perceived indoor-environmental quality

Irene Martínez-Muñoz, Ardeshir Mahdavi, Christiane Berger, Jian Pan & Isabel Mino-Rodriguez

To cite this article: Irene Martínez-Muñoz, Ardeshir Mahdavi, Christiane Berger, Jian Pan & Isabel Mino-Rodriguez (05 Jul 2025): Deconstructing the constructs in multi-domain investigations of perceived indoor-environmental quality, Architectural Engineering and Design Management, DOI: 10.1080/17452007.2025.2527422

To link to this article: https://doi.org/10.1080/17452007.2025.2527422

ARCHITECTURAL ENGINEERING AND DESIGN MANAGEMENT
https://doi.org/10.1080/17452007.2025.2527422

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Deconstructing the constructs in multi-domain investigations of perceived indoor-environmental quality

Irene Martínez-Muñoz[a], Ardeshir Mahdavi[a], Christiane Berger[b], Jian Pan[c] and Isabel Mino-Rodriguez[d]

[a]Faculty of Civil Engineering Sciences, Institute of Building Physics, Services, and Construction, TU Graz, Graz, Austria; [b]Department of Architecture, Design and Media Technology, Aalborg University, Aalborg, Denmark; [c]Healthy Living Spaces lab, Institute for Occupational, Social and Environmental Medicine, Medical Faculty, RWTH Aachen University, Aachen, Germany; [d]Building Science and Technology, Institute for Building Design and Technology, Karlsruhe Institute of Technology, Karlsruhe, Germany

ABSTRACT
This paper reviews the application of constructs and scales in multi-domain studies of perceived indoor-environmental quality (IEQ) in office buildings. Constructs represent hypothetical variables used to capture participants' experiences and evaluations in the course of experimental studies. The review documents the variety of the applied constructs and scales involving thermal, visual, acoustic, and indoor air quality domains in 39 past multi-domain occupant-centric studies. Thereby, a central point of query pertains to the question if the reviewed papers routinely include validation studies regarding the deployed constructs or provide corresponding references to such validation studies. The results of the review reveal certain inconsistencies in the use of construct labels and scales. Moreover, the results point to considerable shortcomings with regard to inclusion of validation tests or related references. As such, the paper highlights the need for a more systematic approach towards the specification and validation of constructs in investigations of subjective human responses to various dimensions of indoor-environmental exposure.

ARTICLE HISTORY
Received 17 March 2025
Accepted 22 June 2025

KEYWORDS
Indoor environment; multi-domain research; constructs; scales; perceived IEQ quality; validation

Introduction

Recently, there has been a growing interest in investigating the combined effects of multiple indoor-environmental stimuli from different domains on building occupants' perception, performance, and behaviour. This interest has resulted in a considerable number of research publications (Candas & Dufour, 2005; Torresin, Pernigotto, Cappelletti, & Gasparella, 2018). However, recent reviews of these publications have pointed out certain methodological shortcomings in past multi-domain IEQ research (Chinazzo et al., 2022; Mahdavi et al., 2020; Mahdavi & Berger, 2023; Pan et al., 2024). Specifically, it has been suggested that some of these shortcomings may be related to issues regarding the selection and deployment of constructs (Mahdavi & Berger, 2023). It is essential to follow up on these, given the critical importance of IEQ research and its quality for the generation of related domain knowledge and formulation of related standards.

CONTACT Irene Martínez-Muñoz irene.martinezmunoz@tugraz.at Faculty of Civil Engineering Sciences, Institute of Building Physics, Services, and Construction, TU Graz, Lessingstraße 25/III, 8010 Graz, Austria

Multi-domain IEQ studies and their limitations

The bulk of past research on IEQ treats its various dimensions (e.g. thermal, visual, auditory, olfactory) in isolation (Galasiu & Veitch, 2006; Gujar & Deshmukh, 2023; Tweed, Dixon, Hinton, & Bickerstaff, 2014; Vardaxis & Bard, 2018). However, occupants in buildings experience these dimensions simultaneously and are hence subjected to the combined influence of the respective indoor-environmental stimuli. To properly understand occupants' perception of and reaction to indoor-environmental conditions, the combined effects of the multifaceted IEQ dimensions need to be taken into consideration. Having recognized this necessity, interest in investigating such combined effects has recently increased, resulting in numerous research publications (see, for instance, Baniya, Tetri, Virtanen, & Halonen, 2016; Geng, Ji, Lin, & Zhu, 2017; Sun, Wu, & Wu, 2020; Witterseh, Wyon, & Clausen, 2004).

Given the relevance of these studies to both fundamental scientific inquiry and practical applications, they have been subjected to content and quality reviews (Chinazzo et al., 2022; Mahdavi et al., 2020; Schweiker et al., 2020a). These reviews suggest that, notwithstanding the valuable contributions that many of the past cross-domain and multi-domain IEQ research studies have delivered, they do involve certain methodological limitations pertaining to matters of research designs, experimental procedures, and statistical analyses. One specific aspect of these limitations constitutes the main concern of the present contribution, namely issues regarding the selection and deployment of constructs in general, and the choice of respective numeric scales and point-based semantic queries in particular (Mahdavi & Berger, 2023). For instance, numerous previous multi-domain studies were suggested to have 'mainly focused on the investigation of subjective perceptual responses, most commonly through numeric scales […] to capture test participants' responses regarding perception, comfort, satisfaction, and preference. At times, a different number of points and different labels were used, even though the same assessment category was involved. This, as well as the inconsistent use of dimensions in analogue scales, disables the comparison of results from different studies and poses a problem for conducting large-scale meta-analyses'(Chinazzo et al., 2022).

Previous investigations have underlined the importance of proper selection of constructs in specific fields, such as perceived indoor air quality (Torriani, Torresin, Lara-Ibeas, Albatici, & Babich, 2024). Specifically, they have brought the attention to the considerable variance and the lack of consistency in the use of scales, which makes the comparison of findings from different studies difficult. Given the critical observations in past research and given the considerable importance of IEQ research for generating domain knowledge and formulating standards, it is essential to address the expressed concerns. Specifically, the nature, scope, and robustness of deployed constructs and scales are crucial for the quality and reliability of research in this area. These constructs include both domain-specific ones (i.e. pertaining to individual IEQ domains) and those addressing the overall assessment of total IEQ. As such, the findings of this paper have the potential to improve both scientific IEQ research and practical IEQ standardization efforts.

Objective

The main objective of this paper is to explore if and to which extent the constructs in multi-domain studies have been subjected to systematic validation procedures. To this end, the kinds of constructs and respective scales in 39 past multi-domain occupant-centric studies on indoor-environmental quality (IEQ) are considered.

In the pertinent literature, the term 'construct' has been used in different fields and defined in different ways. In this contribution, which is focused on the IEQ field, it denotes hypothetical variables used in experimental research to capture participants' experience of various aspects of IEQ (e.g. perception of comfort). The assumption is that what constructs are assumed to represent cannot be directly observed or measured but is derived from participants' semantic reports and choices.

Theoretical framework

About measurements, constructs, and scales

Measuring occupants' perception and evaluation of IEQ conditions is a non-trivial endeavour, independent of the types and number of the domains involved. Reflecting the challenges in this endeavour may shed light on a key factor behind the aforementioned shortcomings. This factor can be best explained considering a broadly defined distinction between two kinds of measurements: those that are representational and those with more pragmatic character (Fayers & Hand, 2002). Whereas measurements in physical sciences are tendentially representational, those in human and social sciences can be characterized as rather pragmatic. Physical representational measurements are conducted based on the assumption that the object of the measurement is a real entity 'out there'. As such, different measurements of that entity are expected to consistently deliver the same result, assuming adequate measurement setup and properly calibrated equipment (Chang, 2004; de Grijs, 2011; Keithley, 1999). However, in psychology and social sciences, measurements are suggested to follow a pragmatic approach, where the target of measurement cannot be claimed to be a unique, objective, or concrete entity existing 'out there'. Rather, what a construct is purported to measure is indeed constructed, to a certain extent, in the process of measurement itself. Classical examples of physical representational measurements are those pertaining to length, area, weight, and volume of objects. Numbers used to express the results of these kinds of physical measurements are assumed to correspond to objectively real properties of physical entities. On the other hand, commonly cited examples of pragmatic measurements include economic indicators such as inflation rate and several stock market indices (Chen, 2024), as well as health-related indicators such as the APGAR score (Apgar, Holaday, James, Weisbrot, & Berrien, 1958), and various pain and wellbeing scores (Fayers & Machin, 2002; Stinson, Kavanagh, Yamada, Gill, & Stevens, 2006). Although pragmatic measures cannot be suggested to represent objective (observer-independent) realities, they can be applied in multiple practical scenarios, providing useful means for purposes such as communication, standardization, organization, and decision-making.

Measurement processes in the IEQ domain frequently involve both representational and pragmatic ingredients. For instance, a Likert-type scale (Carifio & Perla, 2007; Derrick & White, 2017) may be used in an experiment to ask participants about their perception of the loudness of a specific acoustical setting. The order of loudness level selected by participants in that case can be suggested to be the consequence of a real feature of the acoustic field, namely the local magnitude of sound pressure. In other words, the loudness order, as selected by participants from an ordinal scale, commonly correlates – albeit not necessarily in a linear or otherwise simple mode – with the order of numbers associated with objectively measured sound pressure levels. However, the choice of the loudness scale (the specific numbers associated with those steps) is clearly a pragmatic one.

In this context, it is important to note that different types of scales can be used when measuring occupants' perception and evaluation of IEQ. Some of the most commonly used in the field are Likert-type scales, Visual Analogue scales (VAS), and Numerical rating scales (Yang, Moon, & Jeon, 2019; Yang & Jeon, 2021). Moreover, these types of scales can also vary in structure (e.g. unipolar versus bipolar), data type (e.g. categorical versus numerical), the number of scale steps (e.g. 7-point versus 5-point), and the inclusion or exclusion of a 'neutral' response option. Several characteristics of the scales, including type, extent, label, numerical order, and the presence of a neutral alternative influence the quality of the obtained data (DeCastellarnau, 2018). Using different types of scale can also lead to different responses, for example, using continuous and categorical scales can produce significantly different outcomes for the same construct (Favero, Luparelli, & Carlucci, 2023). Deciding the type of scale to use when designing a questionnaire is a complex endeavour, as it can directly influence both the quality of the data and the diversity of responses. Likewise, the exact wording used to capture peoples' experience of IEQ is key concern in composing questionnaires and conducting interviews.

Other aspects requiring careful considerations are the specifics of the target population (e.g. age, gender, cultural background) and the language used (Schweiker et al., 2020b). The comprehension and interpretation of the meaning of the constructs and scales by participants can decisively depend on these aspects. The need for adapting and customizing scales to the specifics of the target populations has been demonstrated in recent studies. For instance, Aparicio-Ruiz, Barbadilla-Martín, Guadix, & Muñuzuri, 2021, customized the scales they used in their study with color-coded elements to improve comprehension among young population.

About constructs validation

Formal tests in science need to meet the pertinent validity criteria. Specifically, applied constructs must be shown to measure the intended target. Given the variety of research designs in different fields, criteria for and methods of validations may differ. As mentioned before, conceiving and conducting an adequate form of validation is especially challenging for pragmatic measurements, given the frequently underdetermined ontological status of the entity or the phenomenon that is to be measured. Facing with previously unexplored questions and novel constructs, researchers might seek ad hoc means of validation (Mahdavi & Eissa, 2002; Mahdavi & Gurtekin, 2001, 2002, 2004). But it would be more prudent to apply systematic validation methods and procedures commonly used in psychological and social disciplines (Cronbach & Meehl, 1955; Flake, Pek, & Hehman, 2017; Heale & Twycross, 2015; Lawshe, 1975, 1985; Lievens, 1998; Phye, Saklofske, Andrews, & Janzen, 2001; Smith, 2005), where measurements and respective constructs are typically pragmatic.

Validity-related terminology is not entirely consistent across literature. Nonetheless, certain types of measurement validity assessment and demonstration have been frequently proposed, including construct validity, content validity, face validity, and criterion validity. These refer to the following questions respectively: Does the test measure what is meant to be measured? Can the test be shown to be representative of the measurement target? Is the appearance of the test consistent with its objective? Do test results correspond correctly to what they were meant to measure?

The construct validity refers to the requirement that the designated construct should measure what it is meant to measure (e.g. perceived thermal comfort), rather than something else (e.g. thermal sensation). Construct validity has been suggested to be of two main types: convergent validity and discriminant validity. The former type focuses on the correspondence between what the designated constructs measure and what similar constructs measure, while the latter type aims to ensure that what the designated constructs yield does not correlate – or negatively correlates – with unrelated or opposing constructs in the same domain. Different methods could be deployed to assess construct validity, including pilot studies based on smaller samples and statistical analyses that check for convergent and discriminant validity or examine the potential of the construct in obtaining reproducible prediction of experimental outcomes.

Approach

Illustrative instances of multi-domain studies

To assemble a list of recent multi-domain studies for the purpose of the present investigation, a database of publications underlying a previous systematic review of experimental multi-domain IEQ studies (Chinazzo et al., 2022). This review study focused in turn on previous investigations of multi-domain research in the IEQ field (Schweiker et al., 2020a; Torresin et al., 2018; Wu, Sun, & Wu, 2020). Thereby, the key criteria for the selection of papers was based on three main considerations, namely (i) the study set-up (dependent variables, independent variables, research hypothesis, setting and exposure features, and experimental design quality), (ii) the study process and analysis (data collection and processing, participants, and statistical analysis), and (iii) the study outcomes (reported results and discussion).

The application of these criteria resulted in a total number of 128 papers included in the database (Chinazzo et al., 2022). The present contribution applied the same criteria to cover the relevant studies published after the aforementioned review. In other words, the same systematic search procedure was applied to update the database, resulting in the inclusion of 15 further multi-domain studies, thus including 143 items. Given the present study's specific focus (occupants' perception of office environments' multi-domain stimuli), a sub-set of these items had to be identified. To this end, two further filtering criteria (f_i, f_{ii}) were applied. The first criterion (f_i) identified those studies that addressed participants' perception of indoor-environmental parameters, either as the sole concern of the study or in combination with performance assessments or physiological monitoring, resulting in 105 items. The second criterion (f_{ii}) further identified those studies that involved office environments or laboratory settings simulating office environments. The paper selection process (see the diagram in Figure 1) resulted in 39 studies that met the filtering criteria and constitute the main source of analysis for the present treatment (Baniya et al., 2016; Berglund & Cain, 1989; Brambilla et al., 2020; Chinazzo et al., 2018a, 2018b, 2020, 2021; Crosby & Rysanek, 2021; Du et al., 2022; Fang et al., 1998, 2004; Garretón et al., 2016; Gauthier et al., 2015; Geng et al., 2017; Gu et al., 2023; Haldi & Robinson, 2010; Hettiarachchi & Emmanuel, 2017; Huang et al., 2012; Jia et al., 2023; Kim & de Dear, 2012; Knez & Hygge, 2002; Ko et al., 2020; Lan et al., 2011; Laurentin et al., 2000; Luo et al., 2023a, 2023b; Martins Gnecco et al., 2024; Melikov & Kaczmarczyk, 2012; Nakamura & Oki, 2000; Ncube & Riffat, 2012; Peng et al., 2023; Salamone et al., 2020; Seyedrezaei et al., 2023; Sun et al., 2020; Tang et al., 2020; Varjo et al., 2015; Witterseh et al., 2004; Wong et al., 2008; Wu et al., 2020). Table A in the Appendix provides an overview of these publications, including an identification code, the reference, the full title, the publication year, and the domains included in the study.

Elements of the research assessment matrix

To gain an overview of the relevant content of the selected studies and their use of constructs, several pieces of information were collected. The resulting summary matrix entailed the following items:

(i) the article code (as per Table A in the Appendix),
(ii) the label of the construct(s) entailed in the article (e.g. 'thermal sensation'),

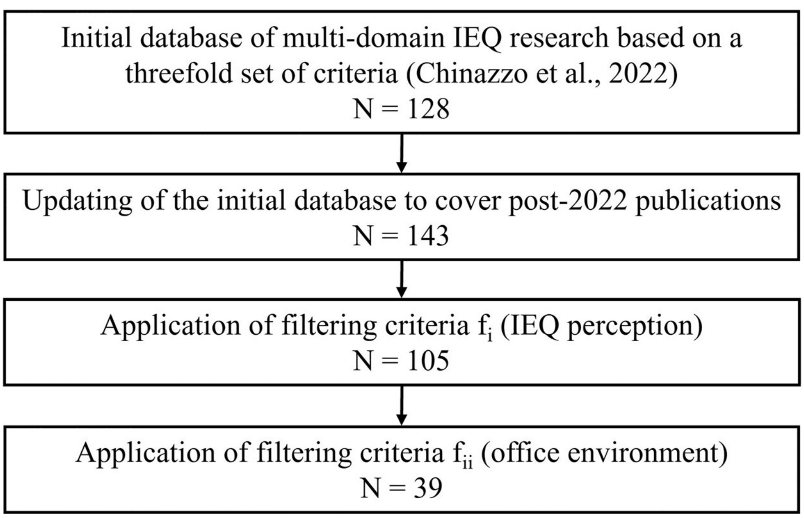

Figure 1. Flow diagram of the literature selection process.

(iii) the question devised to obtain participants' choice of the value for the construct (e.g. 'How would you rate your thermal sensation at this moment?'),
(iv) the type of scale used for participants to specify the value of the construct (e.g. categorical seven-point scale),
(v) the scale details (e.g. thermal sensation scale with steps from − 3 (cold) to +3 (hot)),
(vi) the main independent variables (IV) in the study that were controlled during the experiments (e.g. air temperature, relative humidity, CO_2 concentration, sound pressure level). Further monitored variables measured for control purposes are included if they were directly relevant to the respective construct,
(vii) the type of provided references with potential relevance to the construct validation issue. This is denoted through the letters 'S' (standards), 'P' (papers), 'B' (books), 'O' (others), and 'N/A' (in case no reference is provided).

The information captured in this matrix was structured in terms of five separate tables (see Tables B to F in the Appendix). Four tables (Tables B to E) summarize the collated information according to the constructs for the four domains (thermal, visual, acoustic and IAQ), whereas the fifth table (Table F) focuses on constructs meant to capture participants' overall evaluation of indoor-environmental exposure. In these tables, the use of quotation marks indicates the exact formulation of the items as per the original studies. Terms included without quotation marks suggest the authors' best assumptions as to the original studies' intentions given the information they provide. Some cells in these tables include explanatory comments and additional information marked with an asterisk (*). The ellipsis (…) in column 5 refers to the existence of other categories present in the scales that are not explicitly specified in the original papers. The numbers in the square brackets in Tables B to F represent papers with direct relevance to the provided information. It is also important to mention that in these tables constructs from different papers that can be considered to be for the most part identical (i.e. having the same label and scale specifics) are included in the same row.

Results and discussion

Overview of the constructs and associated questions, scales, and references

As noted earlier, 39 papers, which involved investigations of the perception and evaluation of cross-domain IEQ aspects, were selected for the present study. Thereby, 28 papers investigated two domains, 5 papers three domains, and 6 papers four domains (see Table A in the Appendix). The most frequent combination was thermal and visual (18 of the reviewed papers), followed by thermal and air quality (7 of the reviewed papers), and the combination of all four domains (6 of the reviewed papers). Interestingly, a recent assessment of IEQ experts regarding the relative importance of IEQ domains (Mahdavi, Mino-Rodriguez, Berger, Martínez-Muñoz, & Wagner, 2024) suggested that the thermal, indoor air quality, and visual dimensions of indoor-environmental exposure are, in this order, the most important ones. This may explain, at least to a certain extent, the frequency distribution of domain combinations in multi-domain studies.

In addition to the above IEQ-centric aspects, 14 of the reviewed papers also included specific factors regarding occupants and buildings, thus involving constructs that aimed at assessing participants' emotional states, their physical symptoms or their perception of ergonomic comfort. These aspects are not the focus of the present study and are not discussed in detail. However, they are mentioned to reflect the broader scope adopted by some multi-domain studies. As described in section 3.2., the essential information extracted from the reviewed studies was structured in terms of five tables (Tables B to F in the Appendix).

Once the necessary information was collected and processed, certain observations and inferences could be made. First, a comprehensive overview of the review findings was generated, including the total number of constructs and their classification by their respective domains. Organizing constructs based on their labels facilitated, among others, the derivation of their frequency of use in the reviewed studies. The observations also provided the ground to explore the consistency of the constructs and their scales as well as the manner of their communication. Likewise, the studies' transparency was investigated regarding the applied scales and the questions that were used in experiments to elicit participants' responses. Finally, the issue of construct validation was addressed in detail. The collated information on the reviewed studies was used to determine if they included explicit evidence of the validity of the constructs applied or if, alternatively, they included references to applicable validation studies in other sources (e.g. standards, research papers).

The variety and application frequency of constructs

The information captured in Tables B to F (see Appendix) warrants certain general observations regarding the variety and application frequency of constructs in different categories. In the 39 reviewed studies, a total of 246 distinguishable cases of construct usage were documented across all five categories: thermal (81), visual (87), acoustic (31), IAQ (30), and overall (17). The 'construct label' reflects the variety of terms that have been used to elucidate participants' subjective responses to various aspects of the exposure circumstances, including sensation, acceptability, satisfaction, or comfort, among others. Note that in these tables, constructs sharing the same label were considered distinguishable if their values were not elicited using identical scales. As these tables entail a large amount of information on the identified constructs, it may be useful to look at certain aspects more concisely. One such question relates to the application frequency of the constructs in the set of reviewed studies. To this end, Table 1 provides an overview of the most frequently used constructs in each of the five categories. Note that sometimes slightly different labels were used for labelling essentially the same constructs (e.g. visual sensation, brightness sensation, and lighting sensation were used to label practically the same construct).

Challenges in communicating the meaning of constructs

Based on the information included in the reviewed papers, it is not clear if the semantic nuances in these terms are obvious in the experimental designs, let alone in the minds of the participants. Aside from a certain level of semantic fuzziness associated with labels that are not defined explicitly and in detail, there are even cases of obvious inconsistency in the labelling of constructs. For instance, at times the label 'thermal comfort' is used when the ASHRAE Thermal Sensation 7-point scale (ASHRAE Standard 55-2023) is applied, and hence the use of the term 'thermal sensation' would

Table 1. The most frequently used construct labels for each domain ('Frequency' refers to the number of instances a construct was deployed in the reviewed studies).

Domain	Frequency	Construct label
Thermal	22	Thermal sensation
	15	Thermal comfort
	10	Thermal satisfaction
Visual	12	Visual/Lighting comfort
	11	Visual (brightness/lighting) sensation
	7	Visual/Lighting satisfaction
Acoustic	6	Acoustic satisfaction
	4	Acoustic comfort
	4	Noise/Acoustic sensation
IAQ	7	Air quality acceptability
	5	Odour intensity/Olfactory sensation
	3	Perceived air freshness
Overall	7	Overall satisfaction
	2	Overall comfort
	2	Indoor environment acceptability

have been the appropriate (Sun et al., 2020; Wu, Wu, Sun, & Liu, 2020). These instances reconfirm findings of previous reviews (Chinazzo et al., 2022), pointing to the inconsistent use of terms such as 'comfort' or 'sensation', which renders it difficult to fully trust and generalize the purported findings. Additionally, it is a well-known fact that different labels can lead to different outcomes (Fuchs, Becker, Schakib-Ekbatan, & Schweiker, 2018).

Participants' understanding of what a construct denotes depends highly on the nature of the questions asked to obtain their responses. Hence, to assess the reliability and validity of the findings of studies related to participants' evaluation of indoor-environmental exposure conditions, information on the exact wording of such questions is crucial. This issue is further examined in the following section of the paper, based on the information obtained from the studies and summarized in Tables B to F. A closely related issue concerns the language in which constructs and scales are expressed. As it was briefly alluded to in section 2.1. of the paper, given the diversity of linguistic and cultural contexts both in past and future studies, it is important to consider that language can both accommodate and constrain the comprehension and interpretation of constructs' meaning and intention by participants. Most of the reviewed studies did not explicitly mention the language (or languages) in which the respective of their questionnaires were formulated. This underlines the necessity of systematically providing this essential information in all future studies to provide a better basis for the assessment of their results.

Questions asked to obtain responses from the participants

Notably, not all reviewed papers provide information about the questions used to elicit participants' subjective responses regarding the constructs and the selection of the options in the respective scales. The second column in Tables B to F (see Appendix) lists the questions asked in those papers that contain this information. The Tables also show that questions may differ even if they are posed for the constructs with identical labels and scales. Note that individual papers differ in how they disclose construct-related questions (some papers may have several constructs and disclose the formulation of the construct-related questions for all of them, whereas others may provide that information for only a subset of such questions).

Nonetheless, an overall descriptive statistical impression of the related circumstances can be offered as follows (Figure 2): Of the 81 instances where thermal constructs were used, only 31% disclosed the wording of the questions. This varies between domains and is 48% in case

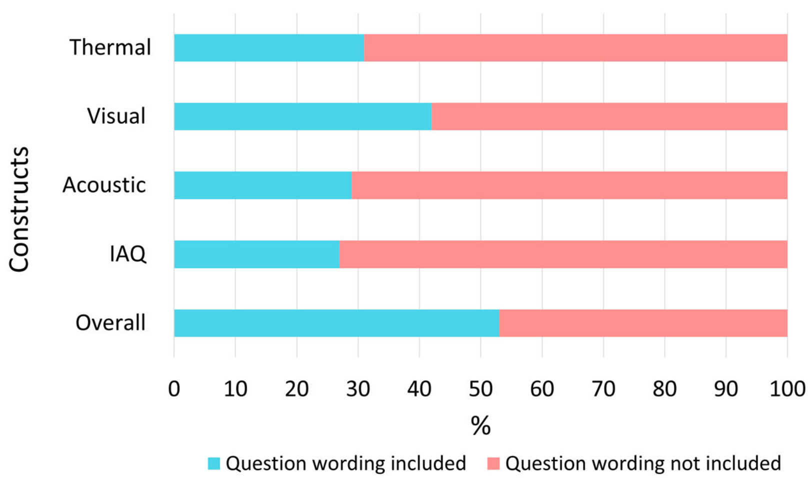

Figure 2. Fraction of construct use cases that either included or did not include the wording of the questions put to the participants to obtain their responses.

of the 87 visual constructs, 29% in case of the 31 acoustic constructs, 27% in case of the 30 IAQ constructs, and 53% in case of the 17 constructs belonging to the overall category. This means that of the 246 distinguishable construct use cases, only 38% provided the wording of the questions set to the participants. The absence of comprehensive documentation and disclosure of such questions in this considerable number of cases must be seen as problematic. It represents an obstacle for an objective appraisal of the studies' findings and their generalization. The wording of the questions asked to the participants can arguably influence the responses received, which highlights the importance of transparency in communicating them together with the studies' findings.

From constructs to scales

Another important consideration is that constructs with identical or highly similar labels do not always use the same scales for obtaining participants' responses. Hence, as described before, in Tables B to F (see Appendix), constructs with the same labels were considered distinguishable and were not included in the same row if their respective scales differed. This is because the responses obtained when using a certain construct label can be arguably very different if different scales are used (see the discussion in section 2.2. above). It is thus useful to organize the information in these tables in a manner that provides a clearer view of those cases where different scales accompanied the same construct label. The previous grouping by construct labels in Table 1 can also be used to group the different scales used, offering a clearer overview of some of these different scales applied for the same construct label. This information (provided in Table 2) suggests that anywhere from 2 to 12 different scales were used for the same construct label. A clear reason for such variety in the scale constitution could not be identified. Hence, whereas the potential necessity for and benefits of such diversity is not obvious, the disadvantages are: Neither upscaling of results, nor systematic meta-analysis of multiple studies are aided by multiplicity of scales that are meant to obtain the values of supposedly identical constructs.

About validation of constructs

The issue of construct validation in the reviewed studies was approached in terms of two questions. The first question (subsection 4.6.1.) explored if the term 'validation' was explicitly mentioned in the studies, and if yes, if it pertained to the constructs used and was accompanied by an actual validation test. The second question (subsection 4.6.2.) explored if the constructs used in the studies were accompanied by any references (to standards, other papers or books, etc.) and if such references indeed provided any evidence of the constructs' validity.

Do studies directly entail constructs validation tests?

Categorically speaking, none of the 39 reviewed papers offered an explicit formal validation test of the constructs they had used in the experiments they entailed. However, for the sake of thoroughness, it may be prudent to qualify this observation by sharing a few related details about cases where the term validation is used. In fact, four papers in the list of reviewed papers do mention the term 'validation' in the sense of systematic construct validation. These are briefly mentioned in the following.

Wong, Mui, & Hui, 2008, mention a kind of internal validation of one of their constructs, which pertained to the acceptability of the perceived indoor environment (as influenced by thermal environment, IAQ, noise level, and illumination), and was assessed using a dichotomous assessment scale. The authors state that the validity of the participants' responses was confirmed in that 'each respondent had to use a semantic differential evaluation scale for the subjective assessment of the first two aspects (i.e. thermal environment, IAQ), and a visual analogue assessment scale for the

Table 2. Overview of different scales in each domain used to obtain participants' responses regarding the same construct label ('N' denotes here the number of occurrences of the construct in the reviewed papers).

Domain	Construct label	Scales	N
Thermal	Thermal sensation	Categorical (7 points) scale from cold (−3) to hot (+3)	20
		Continuous (9 points) from very hot (1) to very cold (9)	1
		Continuous semantic scale from extremely cold to extremely hot	1
	Thermal acceptability	2 categorical scales from clearly acceptable (−1) to just acceptable (0) and from just unacceptable (0) to clearly unacceptable (+1)	1
		Dichotomous continuous scale from clearly acceptable (+1) to clearly unacceptable (−1)	1
		Numerical visual analogue scale from unacceptable (0) to acceptable (100)	1
		Categorical (7 points) scale from very unacceptable (−3) to very acceptable (3)	1
		Continuous (7 points) scale from very uncomfortable (−3) to very comfortable (+3) *no neutral value	1
		Binary scale from acceptable to unacceptable	2
		Dichotomous assessment scale, 1 = yes, acceptable; 0 = no, not acceptable	1
	Thermal satisfaction	Numeric (3 points) scale from dissatisfied (−1) to satisfied (+1)	2
		2 continuous scales from dissatisfied (−1) to just dissatisfied (0) and from just satisfied (0) to satisfied (+1)	1
		Likert (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	1
		Visual analogue (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	1
		Categorical (7 points) scale from strongly dissatisfied (−3) to strongly satisfied (+3)	1
		Likert (7 points) scale from very dissatisfied (1) to very satisfied (7)	1
		Likert (7 points) scale from very satisfied (1) to very dissatisfied (7)	1
		Categorical (10 points) scale from very dissatisfied (1-2) to very satisfied (9-10)	1
		Numerical (6 points) scale from very dissatisfied (−2) to very satisfied (2)	1
	Thermal comfort	Continuous (9 points) scale from very comfortable (1) to very uncomfortable (9)	1
		Categorical (5 points) scale from very comfortable (5) to very uncomfortable (1)	2
		Categorical (5 points) scale from comfortable (0) to extremely uncomfortable (−4)	1
		Visual analogue (6 points) scale from very comfortable to very uncomfortable	2
		Categorical (7 points) scale from very uncomfortable (−3) to very comfortable (3)	1
		Categorical (5 points) scale from very uncomfortable to very comfortable	1
		Likert scale (7 points) scale from very uncomfortable (−3) to very comfortable (+3)	1
		Numerical (6 points) scale from very uncomfortable (1) to very comfortable (6)	1
		Continuous (7 points) scale from very uncomfortable (−3) to very comfortable (+3) *no neutral value	1
		Bipolar (7 points) scale from very dissatisfied (−3) to very dissatisfied (+3)	1
		Numeric (7 points) scale from cold (−3) to hot (+3)	2
		Categorical (7 points) scale from cold (1) to hot (7)	1
	Thermal pleasantness	Continuous scale from very unpleasant to very pleasant	1
		Semantic (continuous) scale from extremely unpleasant to extremely pleasant	1
	Thermal preference	Visual analogue (6 points) scale from much warmer to cooler	1
		Categorical (7 points) scale from much cooler (−3) to much warmer (+3)	1
		Semantic (continuous) scale from much colder to much hotter	1
		Categorical (7 points) scale from cooler (−3) to hotter (+3)	1
Visual	Visual/brightness/	Categorical (7 points) scale from very low (1) to very high (7)	2
		Categorical (10 points) scale from extremely dim (1-2) to extremely bright (9-10)	1
		Categorical (7 points) scale from very dark (−3) to very bright (+3)	3
		Categorical (7 points) scale from too dark (−3) to too bright (+3)	1
		Semantic (continuous) scale from extremely low to extremely high	1
		Semantic (continuous) scale from extremely dark to extremely light	1
		Categorical (7 points) scale from very high to very low	2
	Light/illumination acceptability	Continuous (5 points) scale from not acceptable (1) to acceptable (5)	1
		Dichotomous assessment scale, 1 = yes, acceptable; 0 = no, not acceptable	1
		Categorical (7 points) scale from very unacceptable (−3) to very acceptable (3)	1
	Visual/lighting satisfaction	2 continuous scales from dissatisfied (−1) to just dissatisfied (0) and from just satisfied (0) to satisfied (+1)	1
		Numeric (3 points) scale from dissatisfied (−1) to satisfied (+1)	2
		Bipolar/Likert (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	2
		Categorical (7 points) scale from strongly dissatisfied (−3) to strongly satisfied (3)	1
		Numeric (6 points) scale from very dissatisfied (−2) to very satisfied (2)	1
	Visual/lighting comfort	Numeric (7 points) scale from dark (−3) to extremely bright (3)	2
		Bipolar (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	1
		Categorical (5 points) scale from comfortable (0) to extremely uncomfortable (−4)	1
		Categorical (6 points) scale from very comfortable to very uncomfortable	2

(*Continued*)

Table 2. Continued.

Domain	Construct label	Scales	N
		Categorical (7 points) scale from uncomfortable (1) to comfortable (7)	2
		Categorical (7 points) scale from very uncomfortable (−3) to very comfortable (3)	1
		Numeric (6 points) scale from very uncomfortable (1) to very comfortable (6)	1
		Categorical (5 points) scale from very comfortable (5) to very uncomfortable (1)	2
	Visual/lighting pleasantness	Categorical (7 points) scale from unpleasant (1) to pleasant (7)	2
		Semantic (continuous) scale from extremely unpleasant to extremely pleasant	2
		Rating (10 points) scale from not at all pleasant (1) to very pleasant (10)	1
	Visual/brightness preference	Categorical (7 points) scale from much brighter (1) to much dimmer (7)	3
		Semantic (continuous) scale from much lower to much higher	1
		Semantic (continuous) scale from much darker to much brighter	1
Acoustic	Noise/acoustic sensation	Categorical (10 points) scale	1
		Continuous (5 points) scale from very quiet (1) to very noisy (5)	1
		Visual analogue scale from too quiet (0) to too noisy (100)	1
		Likert scale (5 points) from very quiet (1) to very loud (5)	1
	Noise acceptability	Dichotomous continuous scale from clearly acceptable (1) to clearly unacceptable (−1)	1
		Dichotomous assessment scale, 1 = yes, acceptable; 0 = no, not acceptable	1
		Continuous (5 points) scale from not acceptable (1) to acceptable (5)	1
	Acoustic satisfaction	2 continuous scales from dissatisfied (−1) to just dissatisfied (0) and from just satisfied (0) to satisfied (+1)	1
		Numeric (3 points) scale from dissatisfied (−1) to satisfied (+1)	2
		Categorical (7 points) scale from strongly dissatisfied (−3) to strongly satisfied (3)	1
		Bipolar (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	1
		Numeric (6 points) scale from very dissatisfied (−2) to very satisfied (2)	1
	Acoustic comfort	Numeric (7 points) scale from annoying (−3) to extremely quiet (3)	2
		Scale (5 points) from very uncomfortable to slightly comfortable	1
		Numeric (6 points) scale from very uncomfortable (1) to very comfortable (6)	1
IAQ	Odour/intensity, olfactory sensation	Continuous (6 points) scale from no odour (0) to overwhelming odour (5)	1
		Categorical (6 points) scale from no odour (0) to overwhelming odour (5)	1
		Continuous (6 points) scale from no odour (0) to overpowering odour (5)	2
		Categorical (7 points) scale from unacceptable to excellent	1
	Perceived air freshness	Likert (7 points) scale from totally agree (7) to totally disagree (1)	1
		Continuous scale from air fresh (0) to air stuffy (100)	1
		Continuous scale from fresh to stale	1
	Air quality acceptability	Acceptability (continuous) scale from clearly acceptable (+1) to clearly unacceptable (−1)	2
		2 acceptability (continuous) scales from clearly acceptable (+1) to just acceptable (0) and from just acceptable (0) to clearly unacceptable (−1)	2
		Continuous (5 points) scale from not acceptable (1) to acceptable (5)	1
		Dichotomous assessment scale, 1 = yes, acceptable; 0 = no, not acceptable	1
		Dichotomous assessment scale, acceptable/ not acceptable	1
	Air quality satisfaction	Categorical (7 points) scale from strongly dissatisfied (−3) to strongly satisfied (+3)	1
		Bipolar (7 points) scale form very dissatisfied (−3) to very satisfied (+3)	1
Overall	Indoor environmental acceptability	Dichotomous assessment scale, yes/ no	1
		Continuous dichotomous scale from clearly acceptable (1) to clearly unacceptable (−1)	1
	Overall comfort	Categorical (5 points) scale from very uncomfortable to very comfortable	1
		Numerical (6 points) scale from very uncomfortable (1) to very comfortable (6)	1
	Overall satisfaction	Numeric (3 points) scale from dissatisfied (−1) to satisfied (+1)	2
		Categorical (7 points) scale from strongly dissatisfied (−3) to strongly satisfied (+3)	1
		Bipolar (7 points) scale from very dissatisfied (−3) to very satisfied (+3)	1
		Categorical (10 points) scale from very dissatisfied (1-2) to very satisfied (9-10)	1
		Numerical (6 points) scale from very dissatisfied (−2) to very satisfied (2)	1
		2 continuous scales from dissatisfied (−1) to just dissatisfied (0) and from just satisfied (0) to satisfied (+1)	1

evaluation of the aural and visual comfort.' However, this does not establish a genuine case of construct validation.

In case of a paper by Chinazzo, Wienold, & Andersen, 2020, which includes nine visual constructs and one thermal construct, one construct was suggested to have gone through an internal validation. Since they obtained similar questionnaire responses in case of two of the visual constructs ('visual comfort' and 'lighting comfort'), they discarded one of them. Likewise, in case of a second paper by Chinazzo, Wienold, & Andersen, 2021, which included 13 visual constructs, one construct was suggested to have been validated based on similarity of obtained responses to two constructs ('lighting comfort' and 'lighting pleasantness'), one of which was also discarded. Given the limited selection and narrow scope, these cases do not constitute instances of systematic construct validation.

Finally, in a paper by Varjo et al., 2015, the values for the parameter 'Cronbach's alpha' for acoustic constructs, which is a measure of internal consistency, are mentioned. However, this does not constitute a conclusive case of construct validation.

The kinds of references provided in relation to constructs

As stated before, the reviewed studies did not directly, explicitly, and conclusively treat the issue of construct validation. We thus examined if the authors provided references in connection to the constructs they use in their studies. To this end, we considered all references associated with the introduced constructs. The analysis indicates that, of the 39 reviewed papers, 31 mentioned a reference with connection to at least one of the constructs used. This means that 8 papers did not provide any reference concerning any of their constructs. Of the 246 construct use cases emerged from the review, 147 are accompanied by references, whereas 99 (40%) have been used without any mention or reference to their origin.

When discussing the constructs within each domain, it is important to distinguish the types and percentages of the provided reference (Figure 3). As previously introduced, this indicates whether the reference belongs to a standard (S), a paper (P), a book (B) or other sources (O). This latest type (O) applied actually to only two papers, where in both cases the constructs in question were extracted from a database. In a few cases, two types of references were provided for the construct (papers and standards, which are denoted with the abbreviation 'S-P'). If a paper does not provide

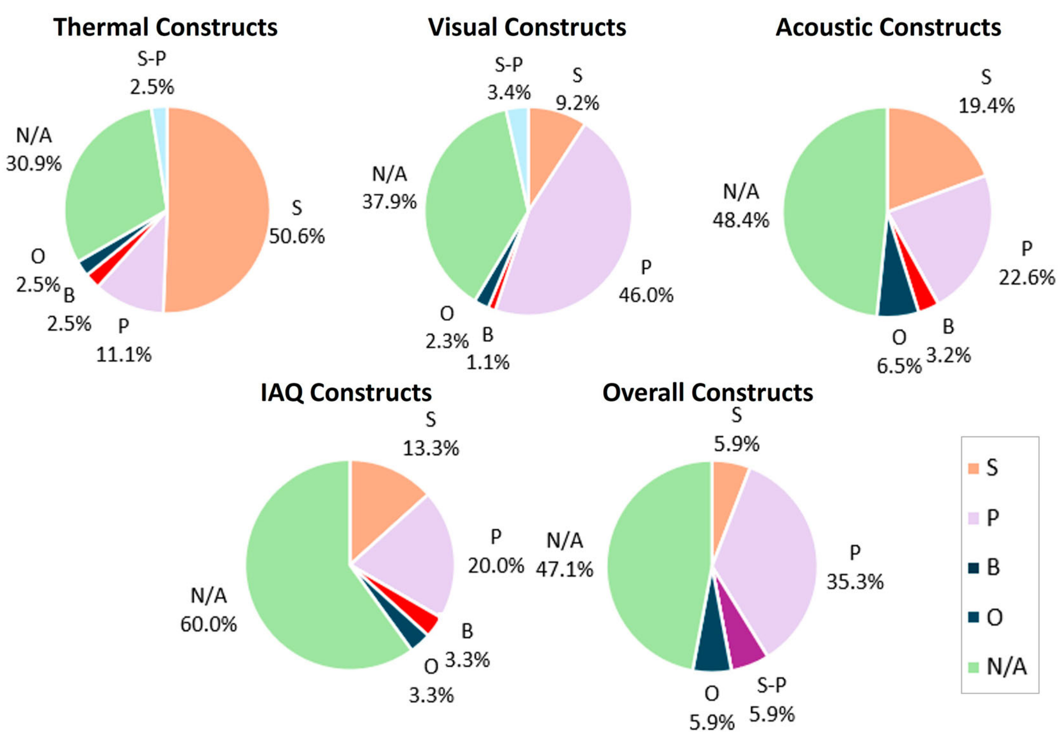

Figure 3. Type of reference provided for the 246 constructs reviewed, grouping in each domain. ('S' = standard, 'P' = paper, 'B' = book, 'O' = other sources, 'N/A' = not applicable (no references provided)).

any kinds of references pertaining to the used constructs, the abbreviation 'N/A' (not applicable) is used. This last circumstance applied to 31%, 38%, 48%, 60%, and 47% of the constructs in the thermal, visual, acoustic, IAQ, and overall domains respectively.

In the thermal domain, over half of the constructs are accompanied by references (Figure 3). This is mainly due to the frequent use of the 'Thermal Sensation' construct in this domain, with ASHRAE (ASHRAE Standard 55-2023) as the main reference provided. In fact, in the thermal domain, 77% of the provided references stem from standards, while in other domains, this percentage is much lower. It should be also noted that referencing a standard does not guarantee the presence of a formal validation. In fact, standards seldom provide evidence for constructs or scales (Berger et al., 2022, 2023a, 2023b; Mahdavi, Cappelletti, & Berger, 2023).

As far as the other domains are concerned, most of the cited references pertain to other papers that used the same or similar constructs and did not necessarily entail detailed and conclusive evidence for the validity of the constructs they applied. In a few cases, the term 'validation' was explicitly mentioned in the provided references. However, also in such cases, this does not involve the introduction of a conclusive formal construct validation test. For instance, Zhang, Huizenga, Arens, & Wang, 2004 is cited in the paper Baniya et al., 2016, in relation to the constructs they used ('Thermal Sensation' and 'Thermal Comfort'). This paper referred to the validation of sensation and comfort models (developed through chamber studies) based on wind tunnel tests, but this does not amount to a relevant formal construct validation test. In case of Wong et al., 2008, the authors refer to a book by Portney & Watkins, 2000, in connection to the constructs they use. This book does indeed cover general issues of validity and reliability of constructs used in health-related questionnaires. However, the link between the treatment of the construct validation issue in the paper and the content of the referenced book is not evident.

Conclusion

In this paper, we assessed the use of constructs and related validity issues in selected studies of perceived multi-domain IEQ research. This work was motivated by concerns, expressed in recent general reviews of related publications, about the quality of respective research designs, particularly regarding selection and deployment of constructs and scales. In psychological research, constructs and their associated scales are generally expected to go through validation tests, documenting their aptitude to reliably measure the intended target of inquiry. Considering this issue from the viewpoint of experimental IEQ research, the main findings of our review can be summarized as follows.

The assessment of the selected papers revealed 246 distinct constructs in five categories (thermal, visual, acoustic, IAQ, and overall). Only 38% of the papers provide detailed information about the exact formulation of the questions meant to elicit participants' responses. Thus, constructs are frequently not accompanied by information about the stated questions. Similarly, reviewed papers at times did not make clear what type of scale they used to measure the constructs. This represents a problem because, as mentioned in 2.1., the type of scale highly influences the outcome, and constructs with identical or similar labels can elicit different responses if accompanied by different questions or different scales.

With the exception of some constructs (mainly in the thermal domain), which appear to have been adapted based on established standards in this area, one does not observe a high level of consistency and harmonization in the constructs and their scales. Moreover, identically labelled constructs have at times different associated questions or scales. The inherent differences and overlaps between such constructs as sensation, perception, comfort, and satisfaction are not consistently delineated. The variations in the labelling of constructs and definition of scales do not appear to be motivated by well-argued necessities. Instead, they can impede conducting meta-analyses and generating cumulative knowledge.

As to the construct validation issue, the findings point to paucity of direct and conclusive validation tests in the reviewed papers. A considerable fraction of the constructs (40%) was not accompanied by references. In those cases where references were provided, the most common type was other papers, except for the thermal domain, where well-established standards were frequently cited as the source of the applied construct. However, the IEQ standards themselves do not necessarily provide direct evidence for constructs, scales, or their mandated values. In those cases, where papers were cited for the applied constructs, they did not entail detailed and conclusive evidence for the constructs' validity.

A further relatively common shortcoming pertains to insufficiently explicated analytical treatment of different scale types, which future efforts must address more consistently. In this context, it is important to emphasize that the manner in which scale-based research findings are derived and compared in terms of percentages, mean values, and other statistical indicators are essential for a reliable interpretation of the results. For instance, linear translation of ordinal scales towards uniform numeric treatment could be misleading. Likewise, uniform numeric treatment of scale with and without neutral positions can also lead to conceptual confusion and quantitative misrepresentation of the results. Progress in this area and potential guidelines and standards in this area could contribute to the improvement of future studies' overall quality.

In light of these observations, it must be concluded that a considerable share of experimental IEQ research, especially multi-domain IEQ exposure studies, display non-negligible deficits with regard to the specification, documentation, and validation of constructs and respective scales. It is thus impingent on the collective efforts on the part of research community in this area to adapt a more systematic approach towards the application of constructs in their experimental investigations of subjective human response to various dimensions of indoor-environmental exposure.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [10.55776/I5993] (Project MuDoCo). For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. Jian Pan and Isabel Mino-Rodriguez are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the project number 498385769.

References

Aparicio-Ruiz, P., Barbadilla-Martín, E., Guadix, J., & Muñuzuri, J. (2021). A field study on adaptive thermal comfort in Spanish primary classrooms during summer season. *Building and Environment*, *203*, 108089. doi:10.1016/j.buildenv.2021.108089

Apgar, V., Holaday, D. A., James, L. S., Weisbrot, I. M., & Berrien, C. (1958). Evaluation of the newborn infant-second report. *Journal of the American Medical Association*, *168*(15), 1985–1988. doi:10.1001/jama.1958.03000150027007

ASHRAE Standard 55-2023. (2023). *Thermal Environmental Conditions for Human Occupancy*.

Baniya, R. R., Tetri, E., Virtanen, J., & Halonen, L. (2016). The effect of correlated colour temperature of lighting on thermal sensation and thermal comfort in a simulated indoor workplace. *Indoor and Built Environment*, *27*(3), 308–316. doi:10.1177/1420326X16673214

Berger, C., Mahdavi, A., Ampatzi, E., Bandurski, K., Hellwig, R. T., Schweiker, M., … Zgank, M. (2023a). The role of user controls with respect to indoor environmental quality: From evidence to standards. *Journal of Building Engineering*, *76*, 107196. doi:10.1016/j.jobe.2023.107196

Berger, C., Mahdavi, A., Ampatzi, E., Crosby, S., Hellwig, R., Khovalyg, D., … Vellei, M. (2023b). Thermal conditions in indoor environments: Exploring the reasoning behind standard-based recommendations. *Energies*, *16*(*4*), 1587. http://doi.org/10.3390/en16041587

Berger, C., Mahdavi, A., Azar, E., Bandurski, K., Bourikas, L., Harputlugil, T., … Schweiker, M. (2022). Reflections on the evidentiary basis of indoor air quality standards. *Energies*, *15*(20), 7727. doi:10.3390/en15207727

Berglund, L. G., & Cain, W. S. (1989). Perceived air quality and the thermal environment. *The Human Equation: Health and Comfort Proceedings IAQ*, *89*, 93–99.

Brambilla, A., Hu, W., Samangouei, R., Cadorin, R., & Davis, W. (2020). How correlated colour temperature manipulates human thermal perception and comfort. *Building and Environment*, *177*, 106929. doi:10.1016/j.buildenv.2020.106929

Candas, V., & Dufour, A. (2005). Thermal comfort: Multisensory interactions? *Journal of Physiological Anthropology and Applied Human Science*, *24*(1), 33–36. V10.2114jpa.24.33

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, *3*(3), 106–116. doi:10.3844/jssp.2007.106.116

Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. New York: Oxford University Press.

Chen, J. (n.d.). *What is an index? Examples, how it's used, and how to invest*. Investopedia. https://www.investopedia.com/terms/i/index.asp

Chinazzo, G., Andersen, R. K., Azar, E., Barthelmes, V. M., Becchio, C., Belussi, L., … Wei, S. (2022). Quality criteria for multi-domain studies in the indoor environment: Critical review towards research guidelines and recommendations. *Building and Environment*, *226*, 109719. doi:10.1016/j.buildenv.2022.109719.

Chinazzo, G., Pastore, L., Wienold, J., & Andersen, M. (2018a). *A field study investigation on the influence of light level on subjective thermal perception in different seasons*. Proceedings of 10th Windsor Conference: Rethinking comfort, Windsor.

Chinazzo, G., Wienold, J., & Andersen, M. (2018b). Combined effects of daylight transmitted through coloured glazing and indoor temperature on thermal responses and overall comfort. *Building and Environment*, *144*, 583–597. doi:10.1016/j.buildenv.2018.08.045

Chinazzo, G., Wienold, J., & Andersen, M. (2020). Influence of indoor temperature and daylight illuminance on visual perception. *Lighting Research & Technology*, *52*(3), 350–370. doi:10.1177/1477153519859609

Chinazzo, G., Wienold, J., & Andersen, M. (2021). Effect of indoor temperature and glazing with saturated color on visual perception of daylight. *LEUKOS*, *17*(2), 183–204. doi:10.1080/15502724.2020.1726182

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302. doi:10.1037/h0040957

Crosby, S., & Rysanek, A. (2021). Correlations between thermal satisfaction and non-thermal conditions of indoor environmental quality: Bayesian inference of a field study of offices. *Journal of Building Engineering*, *35*, 102051. doi:10.1016/j.jobe.2020.102051.

DeCastellarnau, A. (2018). A classification of response scale characteristics that affect data quality: A literature review. *Quality & Quantity*, *52*(4), 1523–1559. doi:10.1007/s11135-017-0533-4

de Grijs, R. (2011). *An introduction to distance measurement in astronomy*. Chichester: Wiley.

Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics*, *18*(3), 1–13.

Du, Y., Li, N., Zhou, L., A, Y., Jiang, Y., & He, Y. (2022). Impact of natural window views on perceptions of indoor environmental quality: An overground experimental study. *Sustainable Cities and Society*, *86*, 104133. doi:10.1016/j.scs.2022.104133

Fang, L., Clausen, G., & Fanger, P. O. (1998). Impact of temperature and humidity on the perception of indoor air quality. *Indoor Air*, *8*(2), 80–90. doi:10.1111/j.1600-0668.1998.t01-2-00003.x

Fang, L., Wyon, D. P., Clausen, G., & Fanger, P. O. (2004). Impact of indoor air temperature and humidity in an office on perceived air quality, sbs symptoms and performance. *Indoor Air*, *14*(s7), 74–81. doi:10.1111/j.1600-0668.2004.00276.x

Favero, M., Luparelli, A., & Carlucci, S. (2023). Analysis of subjective thermal comfort data: A statistical point of view. *Energy and Buildings*, *281*, 112755. doi:10.1016/j.enbuild.2022.112755

Fayers, P. M., & Hand, D. J. (2002). Causal variables, indicator variables and measurement scales: An example from quality of life. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *165*(2), 233–253. doi:10.1111/1467-985X.02020

Fayers, P., & Machin, D.. (2002). Quality of life: Assessment, analysis and interpretation. *In Quality of Life: Assessment, Analysis and Interpretation*, 28–42. doi: 10.1002/0470846283.ch2

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378. doi:10.1177/1948550617693063

Fuchs, X., Becker, S., Schakib-Ekbatan, K., & Schweiker, M. (2018). Subgroups holding different conceptions of scales rate room temperatures differently. *Building and Environment*, *128*, 236–247. doi:10.1016/j.buildenv.2017.11.034

Galasiu, A. D., & Veitch, J. A. (2006). Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: A literature review. *Energy and Buildings*, *38*(7), 728–742. doi:10.1016/j.enbuild.2006.03.001

Garretón, J. Y., Rodriguez, R., & Pattini, A. (2016). Effects of perceived indoor temperature on daylight glare perception. *Building Research & Information*, *44*(8), 907–919. doi:10.1080/09613218.2016.1103116

Gauthier, S., Liu, B., Huebner, G., & Shipworth, D. (2015). *Investigating the effect of CO2 concentration on reported thermal comfort* (J. Scartezzini, Ed.; pp. 315–320). InfoScience. https://eprints.soton.ac.uk/381850/

Geng, Y., Ji, W., Lin, B., & Zhu, Y. (2017). The impact of thermal environment on occupant IEQ perception and productivity. *Building and Environment*, *121*, 158–167. doi:10.1016/j.buildenv.2017.05.022

Gu, J., Li, N., Hu, X., & A, Y. (2023). Influence of artificial windows in simulated underground spaces on thermal and light perceptions, physiological and work performance. *Energy and Buildings*, *297*, 113440. doi:10.1016/j.enbuild.2023.113440

Gujar, S., & Deshmukh, A. (2023). Methodology for visual comfort analysis in intermediate open spaces of residential areas, case study: Nagpur, India. *Architectural Engineering and Design Management*, *19*(6), 602–619. doi:10.1080/17452007.2022.2125490

Haldi, F., & Robinson, D. (2010). On the unification of thermal perception and adaptive actions. *Building and Environment*, *45*(11), 2440–2457. doi:10.1016/j.buildenv.2010.05.010

Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, *18*(3), 66–67. doi:10.1136/eb-2015-102129

Hettiarachchi, A., & Emmanuel, R. (2017, July 3). *Colour as a psychological agent to manipulate perceived indoor thermal environment for low energy design; cases implemented in Sri Lanka*. Passive Low Energy Architecture (PLEA) 2017 - Design to Thrive.

Huang, L., Zhu, Y., Ouyang, Q., & Cao, B. (2012). A study on the effects of thermal, luminous, and acoustic environments on indoor environmental comfort in offices. *Building and Environment*, *49*, 304–309. doi:10.1016/j.buildenv.2011.07.022

Jia, C., Rong, Y., Geng, X., Wang, M., Zhang, Z., Han, S., & Bie, X. (2023). People's psychological and physiological responses to the combined smell-thermal environments. *Building and Environment*, *241*, 110510. doi:10.1016/j.buildenv.2023.110510

Keithley, J. F. (1999). *The story of electrical and magnetic measurments: From 500 bc to 1940s*. New York: IEEE Press.

Kim, J., & de Dear, R. (2012). Nonlinear relationships between individual IEQ factors and overall workspace satisfaction. *Building and Environment*, *49*, 33–40. doi:10.1016/j.buildenv.2011.09.022

Knez, I., & Hygge, S. (2002). Irrelevant speech and indoor lighting: Effects on cognitive performance and self-reported affect. *Applied Cognitive Psychology*, *16*(6), 709–718. doi:10.1002/acp.829

Ko, W. H., Schiavon, S., Zhang, H., Graham, L. T., Brager, G., Mauss, I., & Lin, Y.-W. (2020). The impact of a view from a window on thermal comfort, emotion, and cognitive performance. *Building and Environment*, *175*, 106779. doi:10.1016/j.buildenv.2020.106779

Lan, L., Wargocki, P., Wyon, D. P., & Lian, Z. (2011). Effects of thermal discomfort in an office on perceived air quality, SBS symptoms, physiological responses, and human performance: The effects of thermal discomfort on health and human performance. *Indoor Air*, *21*(5), 376–390. doi:10.1111/j.1600-0668.2011.00714.x

Laurentin, C., Berrutto, V., & Fontoynont, M. (2000). Effect of thermal conditions and light source type on visual comfort appraisal. *Lighting Research and Technology*, *32*(4), 223–233. doi:10.1177/096032710003200406

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x

Lawshe, C. H. (1985). Inferences from personnel tests and their validity. *Journal of Applied Psychology*, *70*(1), 237–238. doi:10.1037/0021-9010.70.1.237

Lievens, F. (1998). Factors which improve the construct validity of assessment centers: A review. *International Journal of Selection and Assessment*, *6*(3), 141–152. doi:10.1111/1468-2389.00085

Luo, W., Kramer, R., Kompier, M., Smolders, K., de Kort, Y., & van Marken Lichtenbelt, W. (2023a). Effects of correlated color temperature of light on thermal comfort, thermophysiology and cognitive performance. *Building and Environment*, *231*, 109944. doi:10.1016/j.buildenv.2022.109944

Luo, W., Kramer, R., Kompier, M., Smolders, K., de Kort, Y., & van Marken Lichtenbelt, W. (2023b). Personal control of correlated color temperature of light: Effects on thermal comfort, visual comfort, and cognitive performance. *Building and Environment*, *238*, 110380. doi:10.1016/j.buildenv.2023.110380

Mahdavi, A., & Berger, C.. (2023). *A critical appraisal of recent research in multi-domain indoor-environmental exposure*. 18th Healthy Buildings 2023 Europe Conference, Aachen, 1172.

Mahdavi, A., Berger, C., Bochukova, V., Bourikas, L., Hellwig, R. T., Jin, Q., … Schweiker, M. (2020). Necessary conditions for multi-domain indoor environmental quality standards. *Sustainability*, *12*(20), 8439, 10.3390/su12208439.

Mahdavi, A., Cappelletti, F., & Berger, C. (2023). Reflections on the scientific basis of building-related acoustic standards. *Journal of Building Engineering*, *74*, 106847. doi:10.1016/j.jobe.2023.106847

Mahdavi, A., & Eissa, H. (2002). Subjective evaluation of architectural lighting via computationally rendered images. *Journal of the Illuminating Engineering Society*, *31*(2), 11–20. doi:10.1080/00994480.2002.10748388

Mahdavi, A., & Gurtekin, B. (2001). *Computational support for the generation and exploration of the design-performance space*. 7th International Building Simulation (IBPSA) Conference, Vol. I, 669–676.

Mahdavi, A., & Gurtekin, B.. (2002). Shapes, numbers, and perception: Aspects and dimensions of the design performance space. Sixth Design and Decision Support Systems in Architecture and Urban Planning - Part one: Architecture Proceedings, Avegoor.

Mahdavi, A., & Gurtekin, B. (2004). Generating the design-performance space via simulation and machine learning. *Journal of Architectural and Planning Research*, *21*(4), 350–362.

Mahdavi, A., Mino-Rodriguez, I., Berger, C., Martínez-Muñoz, I., & Wagner, A. (2024). An exploration of experts' views on the relative importance of indoor-environmental quality parameters. *Building Research & Information*, *53*(1–2), 3–18. doi:10.1080/09613218.2024.2350563

Martins Gnecco, V., Pigliautile, I., & Pisello, A. L. (2024). Exploring office comfort and productivity in living labs: A yearlong structural equation modeling study. *Building and Environment*, *250*, 111147,10.1016/j.buildenv.2023.111147.

Melikov, A. K., & Kaczmarczyk, J. (2012). Air movement and perceived air quality. *Building and Environment*, *47*, 400–409. doi:10.1016/j.buildenv.2011.06.017

Nakamura, H., & Oki, M. (2000). Influence of air temperature on preference for color temperature of general lighting in the room. *Journal of the Human-Environment System*, *4*(1), 41–47. doi:10.1618/jhes.4.41

Ncube, M., & Riffat, S. (2012). Developing an indoor environment quality tool for assessment of mechanically ventilated office buildings in the UK – A preliminary study. *Building and Environment*, *53*, 26–33. doi:10.1016/j.buildenv.2012.01.003

Pan, J., Mahdavi, A., Mino-Rodriguez, I., Martínez-Muñoz, I., Berger, C., & Schweiker, M. (2024). The untapped potential of causal inference in cross-modal research. *Building and Environment*, *248*, 111074. doi:10.1016/j.buildenv.2023.111074

Peng, Y., Antanuri, N., Lau, S.-K., Jebelli, B., Jusuf, S. K., Miller, C., … Chong, A. (2023). Experimental assessment of thermal and acoustics interactions on occupant comfort in mixed-mode buildings. *Building and Environment*, *238*, 110342. doi:10.1016/j.buildenv.2023.110342

Phye, G. D., Saklofske, D. H., Andrews, J. J. W., & Janzen, H. L. (2001). *Handbook of psychoeducational assessment: A practical handbook a volume in the EDUCATIONAL PSYCHOLOGY series*. https://shop.elsevier.com/books/handbook-of-psychoeducational-assessment/phye/978-0-12-058570-0

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research—application to practice*. Upper Saddle River: Prentice-Hall Health.

Salamone, F., Bellazzi, A., Belussi, L., Damato, G., Danza, L., Dell'Aquila, F., … Vitaletti, W. (2020). Evaluation of the visual stimuli on personal thermal comfort perception in real and virtual environments using machine learning approaches. *Sensors*, *20*(6), 1627. doi:10.3390/s20061627

Schweiker, M., Ampatzi, E., Andargie, M., Andersen, R., Azar, E., Barthelmes, V., … Zhang, S. (2020). Review of multi-domain approaches to indoor environmental perception and behaviour. *Building and Environment*, *176*, 106804. doi:10.1016/j.buildenv.2020.106804

Schweiker, M., André, M., Al-Atrash, F., Al-Khatri, H., Alprianti, R. R., Alsaad, H., … Zomorodian, Z. (2020b-03-15). Evaluating assumptions of scales for subjective assessment of thermal environments – Do laypersons perceive them the way, we researchers believe? *Energy and Buildings*, *211*, 109761. doi:10.1016/j.enbuild.2020.109761

Seyedrezaei, M., Awada, M., Becerik-Gerber, B., Lucas, G., & Roll, S. (2023). Interaction effects of indoor environmental quality factors on cognitive performance and perceived comfort of young adults in open plan offices in north American Mediterranean climate. *Building and Environment*, *244*, 110743. doi:10.1016/j.buildenv.2023.110743

Smith, G. T. (2005). On construct validity: Issues of method and measurement. *Psychological Assessment*, *17*(4), 396–408. doi:10.1037/1040-3590.17.4.396

Stinson, J. N., Kavanagh, T., Yamada, J., Gill, N., & Stevens, B. (2006). Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain*, *125*(1–2), 143–157. doi:10.1016/j.pain.2006.05.006

Sun, X., Wu, H., & Wu, Y. (2020). Investigation of the relationships among temperature, illuminance and sound level, typical physiological parameters and human perceptions. *Building and Environment*, *183*, 107193. doi:10.1016/j.buildenv.2020.107193

Tang, H., Ding, Y., & Singer, B. (2020). Interactions and comprehensive effect of indoor environmental quality factors on occupant satisfaction. *Building and Environment*, *167*, 106462. doi:10.1016/j.buildenv.2019.106462

Torresin, S., Pernigotto, G., Cappelletti, F., & Gasparella, A. (2018). Combined effects of environmental factors on human perception and objective performance: A review of experimental laboratory works. *Indoor Air*, *28*(4), 525–538. doi:10.1111/ina.12457

Torriani, G., Torresin, S., Lara-Ibeas, I., Albatici, R., & Babich, F. (2024). Perceived air quality (PAQ) assessment methods in office buildings: A systematic review towards an indoor smellscape approach. *Building and Environment*, *258*, 111645. doi:10.1016/j.buildenv.2024.111645

Tweed, C., Dixon, D., Hinton, E., & Bickerstaff, K. (2014). Thermal comfort practices in the home and their impact on energy consumption. *Architectural Engineering and Design Management*, *10*(1–2), 1–24. doi:10.1080/17452007.2013.837243

Vardaxis, N.-G., & Bard, D. (2018). Review of acoustic comfort evaluation in dwellings: Part II—impact sound data associated with subjective responses in laboratory tests. *Building Acoustics*, *25*, 1351010X–11877202. doi:10.1177/1351010X18772026

Varjo, J., Hongisto, V., Haapakangas, A., Maula, H., Koskela, H., & Hyönä, J. (2015). Simultaneous effects of irrelevant speech, temperature and ventilation rate on performance and satisfaction in open-plan offices. *Journal of Environmental Psychology*, *44*, 16–33. doi:10.1016/j.jenvp.2015.08.001

Witterseh, T., Wyon, D. P., & Clausen, G. (2004). The effects of moderate heat stress and open-plan office noise distraction on SBS symptoms and on the performance of office work. *Indoor Air*, *14*(s8), 30–40. doi:10.1111/j.1600-0668.2004.00305.x

Wong, L. T., Mui, K. W., & Hui, P. S. (2008). A multivariate-logistic model for acceptance of indoor environmental quality (IEQ) in offices. *Building and Environment*, *43*(1), 1–6. doi:10.1016/j.buildenv.2007.01.001

Wu, H., Sun, X., & Wu, Y. (2020). Investigation of the relationships between thermal, acoustic, illuminous environments and human perceptions. *Journal of Building Engineering*, *32*, 101839. doi:10.1016/j.jobe.2020.101839

Wu, H., Wu, Y., Sun, X., & Liu, J. (2020). Combined effects of acoustic, thermal, and illumination on human perception and performance: A review. *Building and Environment*, *169*, 106593. doi:10.1016/j.buildenv.2019.106593

Yang, W., & Jeon, J. Y. (2021). Usability of visual analogue scales in assessing human perception of sound with university students using a web-based tablet interface. *Sustainability*, *13*(16), 9207. doi:10.3390/su13169207

Yang, W., Moon, H. J., & Jeon, J. Y. (2019). Comparison of response scales as measures of indoor environmental perception in combined thermal and acoustic conditions. *Sustainability*, *11*(14), 3975. doi:10.3390/su11143975

Zhang, H., Huizenga, C., Arens, E., & Wang, D. (2004). Thermal sensation and comfort in transient non-uniform thermal environments. *European Journal of Applied Physiology*, *92*(6), 728–733. doi:10.1007/s00421-004-1137-y