JAN
Informing Practice and Policy Worldwide through Research and Scholarship

REVIEW PAPER

A systematic review of the psychometric properties of instruments for assessing the quality of the physical environment in healthcare

Marie Elf, Susanna Nordin, Helle Wijk & Kevin J. Mckee

Accepted for publication 20 January 2017

Correspondence to M. Elf:
e-mail: mel@du.se

Marie Elf PhD RN
Associate Professor
School of Education, Health and Social Studies, Dalarna University, Falun, Sweden
Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden
Department of Architecture, Chalmers University of Technology, Göteborg, Sweden

Susanna Nordin PhD RN
Lecturer
School of Education, Health and Social Studies, Dalarna University, Falun, Sweden
Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

Helle Wijk PhD RN
Associate Professor
Sahlgrenska Academy, Institute of Health and Care Sciences, Gothenburg University, Göteborg, Sweden
Sahlgrenska University Hospital, Göteborg, Sweden

Kevin J. Mckee BSc PhD
Professor
School of Education, Health and Social Studies, Dalarna University, Falun, Sweden

ELF M., NORDIN S., WIJK H. & MCKEE K.J. (2017) A systematic review of the psychometric properties of instruments for assessing the quality of the physical environment in healthcare. *Journal of Advanced Nursing* **73**(12), 2796–2816. **doi: 10.1111/jan.13281**

Abstract

Aim. To identify instruments measuring the quality of the physical healthcare environment, describe their psychometric properties.

Background. The physical healthcare environment is regarded as a quality factor for health care. To facilitate evidence-based design there is a need for valid and usable instruments that can evaluate the design of the healthcare environment.

Design. Systematic psychometric review.

Data sources. A systematic literature search in Medline, CINAHL, Psychinfo, Avery index and reference lists of eligible papers (1990–2016).

Review method. Consensus based standards for selection of health measurement instruments guidelines were used to evaluate psychometric data reported.

Results. Twenty-three instruments were included. Most of the instruments are intended for healthcare environments related to the care of older people. Many of the instruments were old, lacked strong, contemporary theoretical foundations, varied in the extent to which they had been used in empirical studies and in the degree to which their validity and reliability had been evaluated.

Conclusions. Although we found many instruments for measuring the quality of the physical healthcare environment, none met all of our criteria for robustness. Of the instruments, The Multiphasic environmental assessment procedure, The Professional environment assessment protocol and The therapeutic environment screening have been used and tested most frequently. The Perceived hospital quality indicators are user centred and combine aspects of the physical and social environment. The Sheffield care environment assessment matrix has potential as it is comprehensive developed using a theoretical framework that has the needs of older people at the centre. However, further psychometric and user-evaluation of the instrument is required.

Keywords: evidence-based design, healthcare facilities, measurement instruments, nursing, older adults, physical healthcare environment, systematic psychometric review

Why is this review needed?

- The physical environment is an important component of a safe and high quality healthcare service. The difficulty of measuring design outcomes has gained interest internationally.
- The review addresses a problem that healthcare services face today: how can we assess the quality of the physical environment in a scientifically rigorous way?
- We reviewed published instruments that measure the quality of the physical healthcare environment on several criteria and evaluated their reported psychometric properties.

What are the key findings?

- The majority of the 23 instruments was developed during the early 90s and may be less relevant to a contemporary healthcare service, which is focused on person-centred care and interdisciplinary care.
- Few instruments have been subjected to satisfactory psychometric procedures.
- The limitations of the instruments constrain their ability to assess the quality of the physical environment and contribute to evidence-based design.

How should the findings be used to influence policy/practice/research/education?

- The study summarized the range of published measurement instruments as a resource for quality assessment of healthcare environments that support high quality and safe care.
- Much more research is needed to develop instruments that are theoretically well-grounded and predicated on current or emerging models of care and appropriate for measuring modern healthcare environments.
- Some of the identified instruments may have potential as the basis for the development of future instruments that can integrate environmental data on different levels, such as construction, sustainability, and person-environment factors.

Introduction

The physical healthcare environment (PHCE) is an important factor in the quality of health care (Henriksen *et al.* 2007, Eisen *et al.* 2008, Mourshed & Zhao 2012). Good environmental design is regarded as a therapeutic resource for promoting health and well-being (Nightingale 1820/1910, Evans & McCoy 1998, Gesler *et al.* 2004) and as support for the care and treatment of patients (Ulrich *et al.* 2008, Bromley 2012, Huisman *et al.* 2012, Janssen *et al.* 2014). What makes a good quality PHCE is still relatively unexplored, perhaps because the concept of good design is difficult to define and assess (Dewulf & van Meel 2004, Volker *et al.* 2008, Heylighen & Bianchin 2013). However, there is growing interest in developing valid methods to assess the quality of PHCEs. The United Kingdom National Health Service highlights in national protocols (Gesler *et al.* 2004) what should be assessed when considering the quality of PHCEs. Also, the Swedish Institute for Standardization stresses the need for supportive PHCEs and instruments for evaluation (Swedish Standard Institute 2014). To meet this burgeoning interest in the reliable assessment of PHCEs and to generate a useful resource for researchers and for those involved in the planning, design and building of PHCE, we conducted a systematic review of measurement instruments available.

Background

A healthcare environment can be conceptualized in both physical and psychosocial realms (Day *et al.* 2000, Charise *et al.* 2011, Edvardsson *et al.* 2012). The physical component concerns aspects such as space, distance, temperature, colour, and lighting, while the psychosocial component relates to people's interaction with and experience of the environment and their interaction with others in the environment (Dijkstra *et al.* 2006, Edvardsson 2008, Bromley 2012, Huisman *et al.* 2012). The concept of good design is complex in that it is a nexus for both relatively abstract notions (e.g. aesthetics and atmosphere) and pragmatic requirements (e.g. commissioning specifications and resource limitations), simultaneously subject to the technological and commercial fashions of the day and opinions of what good design should be (Gesler *et al.* 2004, Bromley 2012).

Developments in healthcare technology and methodology put high demands on the design of the PHCE (Bromley 2012). Increasing expectations and requirements from patients and staff relating to hospitality, privacy, accessibility, and security present challenges for healthcare design (Vischer 2008, Volker *et al.* 2008). Ultimately, good quality design is best understood in a specified context that relates the finished PHCE to the available options of the architects and builders, framed by the needs and demands of the users (Vischer 2008). Generally, quality in building design tends to be defined more in terms of technical criteria than by the functionality and suitability of the environment once occupied by people (Anåker *et al.* 2016, Vischer 2008, 2009).

Even though guidelines and building regulations exist for the design of specific high quality healthcare environments, they are rarely informed by research evidence and users' views

(Vischer 2008). In addition, there is little evaluation of new buildings once they have been occupied, with a consequent lack of feedback on how design features work in practice (Leaman *et al.* 2010). Research indicates that architects' and designers' ideas of users' preferences for building design features differ substantially from the users' actual preferences (Gifford *et al.* 2000, Arneill & Devlin 2002, Gesler *et al.* 2004).

To ensure a high quality environment, the concept of evidence-based design (EBD) has been introduced (Stankos & Schwartz 2007, Hamilton & Watkins 2009). EBD is defined as a critical and reflective process where decisions about the design of the PHCE is based on the best available information from credible research and evaluation of completed buildings (Stankos & Schwartz 2007, Ulrich *et al.* 2010), in particular the impact of different architectural design solutions on people, costs, and management (Codinhoto *et al.* 2009).

EBD is closely related to continuous quality improvements, where the expected outcomes of the care environment are presented at the beginning of a project, defined by users' needs in relation to the best available research, knowledge, and experience in the field. This allows for an evaluation when the building is completed and is in use, also known as postoccupancy evaluation (POE) (Zimmerman & Martin 2001). The idea behind POE is that by assessing how the design is appraised by users and how it supports certain activities, new knowledge is generated that can be included when new environments are planned (Zimmerman & Martin 2001). As part of POE, various approaches to generate feedback have been used, such as interviews with users. The primary focus has been on the users' experiences and opinions of the environment rather than on predetermined quality criteria and there has been less emphasis on the use of standardized and validated measurement instruments to support the process.

To facilitate EBD for healthcare environments there is a need for valid and usable instruments that can evaluate environmental design on the basis of features and building elements that are known to relate to positive healthcare outcomes (Craik & Femer 1987). Information from such instruments can be used to support better decision-making in new building projects and ultimately improve the overall quality of healthcare buildings. Appropriate instruments can: provide standardized information that allows for the comparison of different environments; identify strengths and weaknesses in the environment; and offer insights into how environments can be better adapted to patients' and staff needs. An acceptable measurement instrument needs to meet established criteria for reliability and validity and be simple to administer by users before widespread deployment can be recommended (Craik & Femer 1987).

The review

Aims

The aims of this systematic review were to: (i) identify instruments that assess the quality of the physical healthcare environment; (ii) describe their psychometric properties, and (iii) evaluate their applicability and feasibility for use in practice and research.

Design

A systematic psychometric review was conducted and framed according to the Consensus based standards for selection of health measurement instruments (COSMIN) (Mokkink *et al.* 2010). In addition, the study followed the preferred schema for systematic reviews and meta-analysis (Liberati *et al.* 2009). The study search and selection process is presented in Figure 1.

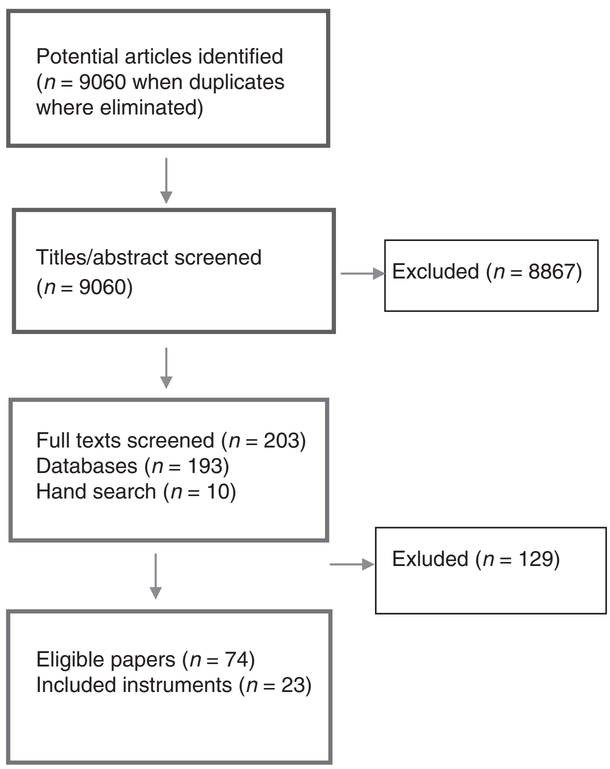

Figure 1 Flow chart of the search process.

Search methods

A systematic literature search from 1990 to 2016 was performed in: Medline; CINAHL; Psychinfo; and Avery index. In addition, we screened the reference list of eligible papers and a second search was performed in the selected databases by using the name of instruments and their developers as identified in the first search. The search period was chosen because it covers the timespan when instruments for measuring quality in healthcare environments have emerged (Fleming 2011).

A Boolean search strategy was adopted incorporating the following truncated search terms and potential synonyms supplemented by appropriate free-text terms entered in various combinations: Tool, Instrument, Scale, Assessment, Measurement, Evaluate, Screening, Physical healthcare environment, Healthcare space, Healthcare setting, Hospital, Healthcare architecture, Healthcare building, Healthcare design (File S1 for further detail).

To be included in the review, papers should be published in English and concerned with measurement instruments addressing the design of healthcare environments. We also choose to include the leading environmental certification instruments even if they were not primarily designed for use in health care. Literature was excluded if it concerned instruments for evaluating private dwellings (Iwarsson *et al.* 2005) or non-healthcare environments or described an instrument that assessed only a single aspect of the healthcare environment (for example, only air quality, or noise, or lighting).

Two reviewers (ME and SN) independently assessed the inclusion eligibility of retrieved papers. The screening process involved: (i) an initial selection for inclusion based on the title and abstract and all duplicates were deleted; (ii) abstracts were screened to determine relevance; (iii) relevant papers were retrieved in full-text; (iv) papers detected by screening the reference lists and by the second search were retrieved; and (v) full-text copies of the papers were assessed by ME and SN to determine whether they fulfilled the inclusion criteria.

Search outcomes

The title and abstract scan resulted in >9000 papers that were judged to meet the inclusion criteria. After full evaluation, 74 papers qualified for the review, which described a total of 23 measurement instruments (Figure 1, Table 1 & File S2).

Quality appraisal

The psychometric properties of instruments were assessed using the COSMIN checklist (Mokkink *et al.* 2010,

Table 1 Names, abbreviations and frequency of references of included instruments.

Name of instrument	Abbreviation	No. of references
Achieving Excellence Design Evaluation Toolkit	AEDET Evolution	4
A Staff and Patient Environment Calibration Toolkit	ASPECT	4
Birthing Unit Design Spatial Evaluation Tool	BUDSET	3
Building Research Establishment Environmental Assessment Method	BREEAM	4
Dementia Design Audit Tool	DDAT	3
Design Quality Indicator	DQI	5
Environmental Audit Tool	EAT	4
Environmental Audit Tool-High Care	EAT-HC	1
Environment-Behaviour (E-B) model for Alzheimer special care units	E-B Model	3
Environment Quality Assessment for Living	EQUAL	2
Evaluation of Older people's Living Environments	EVOLVE	2
Leadership in Energy and Environmental Design	LEED	3
Multiphasic Environmental Assessment Procedure	MEAP	14
Nursing Unit Rating Scale	NURS	3
Perceived Hospital Environment Quality Indicators	PHQI	3
Physical and Architectural Features Checklist (part of MEAP)	PAF	1
Professional Environmental Assessment Protocol	PEAP	14
Rating Scale (part of MEAP)	–	1
Sheffield Care Environment Assessment Matrix	SCEAM	7
Special Care Unit Environmental Quality Scale (a summary scale of TESS-NH)	SCUEQS	1
Swedish version of the Sheffield Care Environment Assessment Matrix	S-SCEAM	2
Therapeutic Environment Screening Survey for Nursing Homes	TESS-NH	8
Therapeutic Environment Screening Survey for Nursing Homes and Residential Care	TESS-NH/RC	1

Terwee *et al.* 2012, Evans *et al.* 2015), consisting of 10 aspects to determine good methodological quality standards such as internal consistency, reliability and content validity, presented in boxes with related items rated on a 4-point scale (where 0 = poor, 1 = fair, 2 = good, and 3 = excellent).

Data abstraction

The included papers were read in full and summarized using a data extraction sheet covering information about the instrument such as its name and source, the setting where it was deployed, purpose, method of administration, items and scoring of items and subscales. Information regarding applicability and feasibility in terms of time to complete and ease of use of the instrument was extracted as well (Table 2 & File S3). Psychometric properties regarding the validity and reliability of measurements were extracted if provided. All data were extracted by ME and SN and checked by KM and HW.

Synthesis

At first, the extracted data was analysed and interpreted by ME and SN independently to gain an overview of the respective instruments' content and quality. Subsequently, the data were analysed to produce a secondary level of conceptualization guided by the research questions. Similarities and contradictions were discussed by the research team, which guided the final results and conclusions.

Results

General characteristics of included instruments

Twenty-three instruments were found (Table 1). The included instruments are summarized in Table 2 and further in File S3. The instruments originate from North America (n = 8), the UK (n = 9), Australia (n = 3), and Europe (n = 3), demonstrating a global interest in measuring PHCEs. Most of the instruments (n = 17) had been developed for healthcare environments related to the care of older people such as SCEAM (Parker *et al.* 2004) and MEAP (Lawton *et al.* 1997). Among these, seven instruments were specifically developed for use in dementia care settings like EAT (Fleming 2011) and E-B model (Zeisel *et al.* 1994). Only two instruments addressed the PHCE in acute care BUDSET and PHQI (Sheehy *et al.* 2011, Andrade *et al.* 2012). However, several of the instruments had a broad area of application for example ASPECT (Abbas & Ghazali 2011), AEDT (Ghazali & Abbas 2012) and DQI (Gann *et al.* 2003) and the environmental benchmarking instruments focusing mainly on green houses such as BREEAM (Schweber & Haroglu 2014) and LEED (Shulman 2003). Several instruments have been developed further into new versions such as TESS (Sloane *et al.* 2002) and SCEAM (Parker *et al.* 2004). MEAP (Moos & Lemke 1996) contains part instruments i.e. PAF, Rating Scale. DQI has recently been developed to provide a version specific for health care (Design Quality Indicator Group 2014).

The instruments varied in the extent to which they had been used in empirical studies and in the degree to which their validity and reliability had been evaluated. The instruments that had been used in the most studies were MEAP (Lawton *et al.* 1997), TESS (Sloane *et al.* 2002) and PEAP (Lawton *et al.* 2000). Certain instruments that were developed some time ago were a reference point, or form the basis, for the development of other instruments e.g. MEAP (Lawton *et al.* 1997).

Dimensions and structure

The instruments varied considerably in terms of their size, with the number of individual items contained in the instruments ranging from >400 to <20. Both SCEAM (Parker *et al.* 2004) and MEAP (Lemke & Moos 1986) contained many items, structured into a series of domains. The instruments also varied in scope, some focusing on the assessment of a few specific dimensions of the physical environment, others assessing a more comprehensive range of dimensions. Aspects of the environment assessed included functionality (use, access, space), impact (materials, character, and impression) and build quality (engineering, construction, and performance). Several of the instruments such as SCEAM (Parker *et al.* 2004), TESS (Sloane *et al.* 2002) and AEDET (Abbas & Ghazali 2011) additionally assessed if e.g. the environment could support privacy, comfort and choice or control.

Aim of the instruments

The main uses of the instruments could be identified as being for: evaluating existing building design to improve the physical environment (Fornara *et al.* 2006) and/or planning new healthcare environments (Whyte & Ganna 2003) and/or providing a quantitative evaluation of the building, often for research purposes (Lawton *et al.* 1997).

Well-known instruments in the fields of architecture and construction are LEED (Steinke *et al.* 2010) and BREEAM (Steinke *et al.* 2010). These are specific benchmarking

Table 2 General information of instruments included in the review.

Instruments and references	Aim and target environment	Administration and scoring	Items, subscales/domains
AEDET Evolution (Abbas & Ghazali 2011)	To assess design quality of a broad range of buildings	Self-assessment form Can be used together with ASPECT or alone 6-point Likert-scale ranging from agree completely to not agree	Fifty seven items, in three areas Impact: form, materials Build quality: engineering Functionality: use, access
ASPECT (Abbas & Ghazali 2011)	To evaluate the quality of design of staff and patient environments in healthcare buildings in general.	Self-assessment form Can be used to support AEDET or alone 6-point Likert-scale ranging from agree completely to not agree	Forty-seven items, eight domains Privacy, company, dignity, views, nature, outdoors, comfort, control, interior appearance
BREEAM, www.breeam.org (Schweber & Haroglu 2014)	To assess environmental and sustainability issues in a broad range of buildings	Rating is made through site visits, audits and document review by licensed assessors in collaboration with the design team. The sum of the scores results in a 5-level classification from pass to outstanding	Eight main categories Energy, materials, innovation, waste, pollution, health, water, transport
BUDSET (Foureur *et al.* 2011)	To assess the quality of the design of hospital birthing units	Direct observation and survey Each item is marked as present or absent with a total score calculated for each domain and an overall score for the facility	Eighty four items, four domains Fear cascade, facility appearance, aesthetics, and support
DDAT, www.dementia.stir.ac.uk (Cunningham 2009, Kelly *et al.* 2011)	To provide consistent guidance in the design of facilities for people with dementia	Direct observations 3-point scale ranging from standard not met to standard fully met Final scores are weighted according to the category. Essential category represents 30% of the total score; Recommended category represents 70% of the total score	181 items, two categories (essential and recommended), 11 building areas Hall/entrance/way-finding, lounge/day room, meaningful occupation and activity, bedrooms, toilet area, bathroom/shower room (en-suite), dining room, treatment areas, lighting
DQI (Gann & Whyte 2003)	To assess design quality of buildings in general	Self-assessment form Likert scale. Scores are aggregated to a total sum	90 items, 10 sections Character and innovation, form and materials, staff and patient environment, urban and social integration, build quality, performance, engineering, construction, functionality, use, access, space
EAT (Smith *et al.* 2012)	To assess the quality of residential care facilities for persons with dementia	Direct observations Dichotomous scale (Yes/No) The total score is the mean of the ten domain percentage scores	72 items, 10 domains Safety and security, small size, visual access features, stimulus reduction features, highlighting useful stimuli, provision for wandering and access to outside area, familiarity, privacy and community, community links, domestic activity
EAT-HC (Fleming & Bennett 2015)	To assess the quality of residential care facilities for persons with dementia, including those who are immobile or in palliative care	Direct observations Dichotomous scale (Yes/No) The total score is the mean of the ten domain percentage scores	Seventy seven items, 10 domains Safety and security, small size, visual access features, stimulus reduction features, highlighting useful stimuli, provision for wandering and access to outside area, familiarity, privacy and community, community links, domestic activity

Table 2 (*Continued*).

Instruments and references	Aim and target environment	Administration and scoring	Items, subscales/domains
E-B Model (Zeisel *et al.* 2003)	To describe and organize the influences that the physical environment has on residents and caregivers in Alzheimer special care units (SCUs)	Self-score form two dimensions of each domain scoring on a 3-point scale ranging from excellent to poor environmental features	Sixty-one items, eight dimensions Exit control, wandering path, individual away places, common space structure, outdoor freedom, residential character, autonomy support, sensory comprehension
EQUAL (Cutler *et al.* 2006)	To assess physical environments for older people with or without dementia	Observation checklist Dichotomous scale (yes/no) for a majority of items, some multiple-choice options, a few require measurement or count	387 items, 3 sections, 11 domainsAutonomy, dignity, privacy, meaningful activity, enjoyment, relationships, comfort, security, functional competence, spiritual well-being, individuality
EVOLVE (Lewis *et al.* 2010, Orrell *et al.* 2013)	To evaluate the design of institutional housing for older people, and how well a building contributes to the physical support and personal well-being	Direct observations	487 items for a single dwelling; 2020 items for a housing scheme, two categories (universal needs and support for older age) which are further divided into 13 subdomains Personal realization and choice, dignity and privacy, comfort and control, personal care, social support inside building, social contact outside, accessibility, physical support, sensory support, health and safety, security, working care
LEED, www.usgbc.org (Happio & Viitaniemi 2008, Steinke *et al.* 2010)	To identify, implement and measure green building and neighbourhood design, construction, operations and maintenance	Used for environmental certification for private or institutional buildings. Can be applied to a broad range of healthcare facilities	Buildings can be qualified into four certification levels: certified, silver, gold or platinum Energy efficiency, indoor environmental quality, materials selection, sustainable site development, water savings
MEAP (Moos & Lemke 1996)	To evaluate the physical features and social environments in residential facilities for older people	Direct observation, questionnaires and document analysis Comprises of five parts with different aspects of residential care facilities that can be used separately A profile of the building is created and compared to a standard score mean of 50 and SD 10	474 items, 33 dimension (five subscales) 1. RESIF (resident and staff information form; 104 items) 2. PAF (physical and architectural features checklist; 153 items) 3. POLIF (Policy and program information form; 130 items) 4. SCES (Sheltered care environment scale; 63 items) 5. Rating Scale (24 items)
NURS (Morgan *et al.* 2004)	To assess policy and programme features of dementia specific care units	Observations and analysis of documents (policy and programme features) and interviews with staff 5-point Likert scale ranging from always to never, or a 4-point Likert-type scale from not at all to a great deal. Each dimension is the sum of item scores divided by 5 or 4. No total score is obtained	81 items, 6 dimensions Separation, stability, stimulation, complexity, control/tolerance, continuity

Table 2 (*Continued*).

Instruments and references	Aim and target environment	Administration and scoring	Items, subscales/domains
PAF (a part of MEAP) (Linney *et al.* 1995)	To measure physical resources of residential facilities for older people	Direct observation supplemented by information from administrators or staff. Dichotomous scale (yes/no) for a majority of questions The raw scores are percentage scores reflecting the number of features present out of the total number. A profile of the building is created and compared to a standard score mean of 50 and SD 10.	153 items, 8 domains Community accessibility, physical amenities, social-recreational aids, prosthetic aids, orientation aids, safety features, staff facilities, space availability.
PEAP (Lawton *et al.* 2000, Slaughter & Morgan 2012)	To provide a standardized method of expert evaluation of special care units for people with dementia. The physical setting is the primary focus, but the assessment is conducted within an understanding of the larger context of the social, organizational, and policy environment.	Interview with administrative staff and 2-hour participant observation in the special care unit5- point Likert scale for each dimension ranging from unusual low support to exceptionally high support.A score on dimension levels can be obtained as well as an overall summary score	Nine dimensions Maximize safety and security, maximize awareness and orientation, support functional ability, facilitation of social contact, provision of privacy, opportunities for personal control, stimulation and coherence (regulation), stimulation and coherence (quality), continuity of the self.
PHQI (Fornara *et al.* 2006, Andrade *et al.* 2012)	To assess design and social attributes that are expected to have a role in assessing the quality of the healthcare environment	A self-assessment questionnaire is filled in by hospital users (patient, relatives and staff). One observational grid is filled in by experts (architects and engineers) about architect's technical attributes 5-point Likert response scales ranging from totally disagree to fully agree. The instrument includes equal numbers of positive- and negative-worded statements.	71–80 items (the instrument is still in development phase), three scales Spatial-physical aspects of the external spaces of the hospital, spatial-physical aspects of the care unit and waiting areas, social-functional aspects of the care unit
Rating Scale (part of MEAP) (Morgan *et al.* 2004)	To measure physical environment and resident and staff functioning in residential facilities for older people.	Many items are overlapping two parts in MEAP; RESIF and PAF, but is intended to tap more subjective aspects of the setting 4-point response scale.	24 items, 4 subscales Attractiveness (odour, noise, cleanliness), environmental diversity (stimulation, variation, view, private rooms for residents), resident function (resident appearance, activity level, interaction), staff functioning (reflects quality of interaction between staff and residents, organization of the facility, amount of conflict among staff members)

Table 2 (*Continued*).

Instruments and references	Aim and target environment	Administration and scoring	Items, subscales/domains
SCEAM (Parker *et al.* 2004)	To assess the physical environment of residential care facilities for older people	Assessment checklist completed by direct observation The assessor completes a checklist of items by indicating yes(1)/no (0) to their presence/ absence. Scores are summed to provide an overall score or scores by home area or domain	337 items, 11 domains Privacy, personalization, choice and control, community, safety and health, physical support, comfort of the environment, cognitive support, awareness, normalness authenticity, and provision for staff
S-SCEAM (Nordin *et al.* 2015)	To assess the physical environment of residential care facilities for older people	Assessment checklist completed by direct observation. Guided by checklists the assessor answer yes/no questions by observation	210 items, 8 domains Integrity, choice, openness and integration, safety, physical support, comfort, cognitive support, normalness
SCUEQS (a summary scale comprised of 18 TESS-NH variables) (Sloane *et al.* 2002)	To measure the ability of physical environments to address therapeutic goals for persons with dementia	Self-assessment form via direct observation	18 items (a summary scale comprised of TESS-NH variables) within seven domains. Maintenance, cleanliness, safety, lighting, physical appearance/homelikeness, orientation/cueing, noise.
TESS-NH (Slaughter *et al.* 2006, Fleming 2011)	To assess the physical environment of institutional facilities for persons with dementia	Self-assessment form via direct observation Scale 0–3 (0 = absent, 1 = present) for the 84 items. The higher number represents a more favourable attribute of the environment The global item: scoring on a Likert- scale ranging from 1 (low, distinctly unpleasant, negative, and non-functional) to 10 (high quite pleasant, positive, and functional) The global item gives a summary of the quality of the environment, but the 84 items do not combine to form a scale and a summary of the quality of the environment cannot be obtained	84 items, 13 domains plus a global item. Unit autonomy, outdoor access, privacy, exit control, maintenance, cleanliness, safety, lighting, noise, visual/tactile stimulation, space/seating, familiarity/homelikeness, orientation/cueing.
TESS-NH/RC (Sloane *et al.* 2002)	To assess the physical environment of long-term care settings	Self-assessment form via direct observation Scores may be 0, 1, or 2 resulting in a summary score ranging from 0 to 30. Higher score indicate better quality.	Contains mostly items from TESS-NH. The items reflect 15 domains. Facility maintenance, cleanliness, handrails, call buttons, light intensity, light glare, light evenness, hallway length, homelikeness, room autonomy, telephones, tactile stimulation, visual stimulation, privacy, outdoor areas.

instruments focused on green buildings and technical aspects such as energy consumption, water use, or materials. The instruments have been used in professional practice and there is a track record of their use but there is little reference to them in the research literature.

A fundamental distinction could also be made between those instruments that assessed the physical environment from a user-centred perspective such as SCEAM (Parker *et al.* 2004)and TESS (Fleming 2011) and PHQI (Fornara *et al.* 2006) and those instruments, such as LEED (Steinke *et al.* 2010) and BREEAM (Schweber & Haroglu 2014) that addressed technical aspects of buildings with little or no reference to a building's users.

Conceptual framework

Some of the instruments had a strong theoretical foundation for their development such as MEAP (Lawton *et al.* 1997), SCEAM and TESS while others had been developed on a more empirical basis like ASPECT and AEDET (Abbas & Ghazali 2011). Overall, the instruments were rarely embedded in explicit conceptual frameworks, making it difficult to establish conceptual comparability between instruments. The instruments' most common conceptual framework was Lawton's ecological model (Lawton & Nahemow 1973), which stipulates that for an older person to maintain independence and quality of life there is a need for congruence between the older person's capacity and the demands of the environment. According to this model, the environment interacts with the persons in it and there are relationships between the design of a building and therapeutic outcomes. The model originates from the idea that ageing is connected with increasing levels of impairment and therefore the environment must be adjusted to these new conditions to support independence and well-being in the frail older people.

For example MEAP (Lawton *et al.* 1997) explicitly uses the ecological model as a framework. For other instruments, while it was not explicitly expressed that they derived from Lawton's ecological model, the model can be discerned in the description of the instrument. For example, TESS-NH (Aiken *et al.* 2002) is conceptualized in terms of interactions between a physical space and the persons in it. Several instruments were predicated on the evidence-based needs of older people e.g. SCEAM (Parker *et al.* 2004) and some of these had a specific focus on persons with dementia such as TESS-NH (Aiken *et al.* 2002). Both TESS-NH and SCEAM are expressions of a theoretical framework where quality of life and well-being are regarded as influenced by the environment.

The instruments that were developed in the construction industry have imprecise conceptual frameworks. The development of the instrument(s) was often justified by reference to established links between health and well-being and the environment without further theoretical background.

Psychometric properties

Data extracted for the psychometric evaluation of the selected instruments is summarized in Table 3. A general and important limitation of all the included instruments was the low level of validation work that had been carried out. The respective instrument developers and/or study authors in many cases indicated that the instruments satisfied various reliability and validity criteria, but for the most part this was not supported by the presentation of data. Several instruments had been pilot tested in the course of their development, which did address some aspects of their validity.

Face and/or content validity were described for most the instruments, even if no tests or figure were presented. Many of the instruments had been developed systematically and rigorously both according to literature reviews for generating items and through the use of expert panels for assessing the relevance of the items included in the instrument. For example, MEAP (Lawton *et al.* 1997) was developed through a careful literature review and a pool of items were generated and judged by experts indicating that content and face validity were met. The same procedure is described for SCEAM (Parker *et al.* 2004), EVOLVE (Orrell *et al.* 2013), TESS-NH (Fleming & Purandare 2010), EAT (Fleming 2011) and PEAP (Cutler *et al.* 2006).

TESS-NH (Fleming & Purandare 2010), EAT (Fleming 2011), PEAP (Slaughter & Morgan 2012), and MEAP (Moos & Lemke 1996) have been examined in relation to criterion validity, with good results. Studies that have used MEAP and the E-B model produce data that suggest a good match between the instruments and their respective conceptual frameworks and the researchers responsible for the studies use this as a basis to argue for the instruments' high construct validity (Linney *et al.* 1995, Zeisel *et al.* 2003).

Reliability data were available for many of the instruments although there was a lack of rigorous reliability testing reported. The reliability tests that were mostly used were inter-rater reliability and Cronbach's alpha (internal consistency). For example EAT (Fleming 2011), PEAP (Slaughter *et al.* 2006), MEAP (Linney *et al.* 1995) and TESS-NH (Sloane *et al.* 2002) were all presented with data that indicated moderate to strong inter-rater reliability for the instruments. Stability, or test–retest, reliability was reported for three of the instruments: TESS-NH (Sloane *et al.* 2002), S-SCEAM (Nordin *et al.* 2015), and PHEQI (Fornara *et al.* 2006).

No psychometric data was reported for the instruments developed for the construction sector, i.e. LEED (Steinke *et al.* 2010) and BREEAM (Steinke *et al.* 2010) and their reliability and validity can therefore be questioned. Instruments which were developed for use in research, such as AEDET (Abbas & Ghazali 2011) and ASPECT (Abbas & Ghazali 2011) and DQI (Gann & Whyte 2003), all have

Table 3 Results psychometric properties rated by COSMIN checklist.

Instrument*	References	COSMIN assessment†	Reliability/Validity
AEDET	Abbas & Ghazali (2011), Ghazali & Abbas (2012)	NA	NR
ASPECT	Abbas & Ghazali (2011), Ghazali & Abbas (2012)	NA	NR
BREEAM	Crawley & Aho (2006), Schweber (2013), Schweber & Haroglu (2014)	NA	NR
BUDSET	Sheehy *et al.* (2011), Foureur *et al.* (2010), Foureur *et al.* (2011)	Box B: Fair Box D: Good Box F: Fair	*Content validity* Expert groups assessed the relevance by using content validity index (CVI) for items (I-CVI) and for scale (S-CVI) and interviews. CVI was reached in all domains (0·89–0·97) *Interclass correlation coefficient (ICC)* The ICC was acceptable (at a level of >0·60) for 9 (50%) of the 18 characteristics measured by the instrument *Construct validity* Hypotheses testing. Not formulated but possible to deduce what was expected. No comparator instrument (s) used
DDAT	Cunningham (2009), Kelly *et al.* (2011)	Box A: Fair Box B: Fair Box D: Fair Box E: Fair Box F: Fair	*Content validity* Item generation was based on expert consultation and extensive literature review followed by pilot studies. No figures presented *Construct validity* Could discriminate between various dementia settings as presumed *Concurrent validity* Strong concurrent validity when compared to the global score of TESS-NH (0·89), and the sum score of SCUEQS (0·87) *Cronbach's alpha* Five of the sub-scales did not reach 0·60 *Interclass correlation coefficient (ICC)* Ranged from −0·12 to 1 (20·1% of items had ICC <0·4; 28·8% had ICC higher than 0·70) *Inter-rater reliability* The average of agreement between two raters was 79% (range 43–100%) The inter-rater reliability of the total score was 95%
DQI	Gann *et al.* (2003), Markus (2009), Thomson *et al.* (2003), Whyte & Gann (2003)	Box D: Fair	*Content validity* The tool is reported to have been tested for face and content validity in several projects with good results. No figures are reported
EAT	Fleming & Purandare (2010) Fleming (2011) Fleming *et al.* (2012) Smith *et al.* (2012)	Box A: Good Box B: Good Box D: Fair Box F: Good	*Content validity* Item generation was based on literature review and earlier instruments. No figures are presented *Construct validity* EAT sufficiently differentiates between traditional and purpose-built facilities in principles of design that are necessary in environments of people with dementia *Concurrent validity* Showed strong concurrent validity when compared to the global score of TESS-NH (0·82), and the sum score of SCUEQS (0·85) *Cronbach's alpha* Two of the domains did not reach 0·60 during the development phase *Interclass correlation coefficient (ICC)* Ranged from −0·05 to 1 (13·8% of items had ICC <0·4; 54·2% had ICC higher than 0·70) *Inter-rater reliability* The average of absolute agreement between two raters was 86·8% (range 46·6–100%). The inter-rater reliability of the total score was 97%

Table 3 (*Continued*).

Instrument*	References	COSMIN assessment†	Reliability/Validity
EAT-HC	Fleming & Bennett (2015)	Box A: Good Box D: Fair Box F: Good	*Content validity* Item generation was based on literature review and earlier instruments *Concurrent validity* The Pearson correlations between the Total EAT-HC score and the TESS-NH Global 0·72, and SCUEQS 0·34 *Cronbach's alpha* Internal consistency assessed with Cronbach's alpha, were satisfactory, ranging from 0·57 to 0·88
E-B Model	Zeisel (2003)	Box D: Fair Box F: Fair	*Content validity* Item generation was based on literature review and earlier instruments. No figures presented *Construct validity* A study testing the instrument shows that the measure could discriminate among various facilities and correlates to older person's behaviour and health status e.g. persons score lower on the psychotic problem scale when living in a facility supporting privacy-personalization
EQUAL	Cutler *et al.* (2006), Cutler & Kane (2009)	Box A: Poor Box B: Fair Box E: Poor	*Construct validity* A cognitive rating process was performed. Experts assigned each item to predefined domains *Inter-rater reliability* Extensive tests of inter-rater reliability during the development phase using kappa statistics. Items with low *k* were deleted from the tool
EVOLVE	Lewis *et al.* (2010), Orrell *et al.* (2013)	Box B: Poor Box D: Poor	*Content validity* Support for face and content validity. No figures presented *Reliability* Strong inter-rater reliability when testing the instrument in three care facilities, no figures presented
LEED	www.usgbc.org; Happio and Viitaniemi (2008), Steinke *et al.* (2010)	NA	NR
MEAP	Benjamin & Spector (1990), Benjamin & Spector (1992); Braun (1991), Davidson *et al.* (1996), Field *et al.* (2005), Izal (1992), Fleming & Purandare (2010), Fonda *et al.* (1996), Linney *et al.* (1995), Moos & Lemke (1996); Sikorska-Simmons (1996), Timko & Moos (1990), Timko & Moos (1991), Wells & Taylor (1991)	Box A: Fair Box D: Fair Box F: Fair	*Content validity* Item generation was based on literature review and earlier instruments. No figures presented *Construct validity* The tool has been able to discriminate between various environments in a range of studies *Cronbach's alpha* The 5 scales had a Cronbach's alpha that ranged from 0·50 to 0·85
NURS	Grant (1996), Morgan *et al.* (2004), van Hoof *et al.* (2010)	Box A: Fair	*Cronbach's alpha* Four of six dimensions have showed good alpha coefficients (from 0·83 to 0·95)
PAF	Linney *et al.* (1995), Davidson *et al.* (1996)	Box A: Fair Box F: Fair	*Construct validity* The scale has discriminated between various stakeholders' (staff and clients) views of important features of an environment *Cronbach's alpha* The different subscales alpha coefficient ranged from 0·83 to 0·94 or 0·62 to 0·84

Table 3 (*Continued*).

Instrument*	References	COSMIN assessment†	Reliability/Validity
PEAP	Barnes (2004), Campo & Chaudhury (2012), Cutler (2007), Cutler *et al.* (2006), Fleming & Purandare (2010), Fleming (2011), Lawton *et al.* (2000), Lawton (2001), Morgan *et al.* 2004, Schwarz *et al.* (2004), Slaughter & Morgan (2012), Sloane *et al.* (2002), Teresi *et al.* (2000), Weisman (1994)	Box A: Fair Box B: Fair Box D: Fair Box E: Fair Box F: Fair	*Content validity* Item generation was based on literature review and earlier instruments. No figures presented *Construct validity* Correlations among the dimensions ranged from 0·45 to 0·85. Variation of the environments in special care units for dementia care was reflected. The summary scores discriminated between special care units and integrated facilities in comparison of rural nursing homes *Factor analysis* Principal components analysis generated a single factor structure for the nine dimensions accounting for 67% of the total variance *Concurrent criterion validity* Global scores showed strong correlation with TESS-NH global rating (r = 0·71)
PHQI	Andrade *et al.* (2012), Andrade *et al.* (2013), Fornara *et al.* (2006)	Box A: Good Box B: Good Box D: Good	*Construct validity* The tool could discriminate between settings with different quality *Criterion validity* Showed high correlation with three global questions on design quality *Cronbach's alpha* The four scales had an alpha ranging from 0·64 to 0·91 *Factor analysis* Repeated principal components analysis revealed 12 factors of quality environment perception. The factors had a total explained variance of 54·3–58·3 (only one scale had a lower explained variance: 44·4) *Test–retest reliability (%)* The various scales showed satisfactory to very good reliability 0·64–0·85 (Andrade *et al.* 2012)
Rating Scale	Morgan *et al.* (2004), Davidson *et al.* (1996)	Box A: Poor	*Cronbach's alpha* The subscale demonstrates a value of 0·67–0·82
SCEAM	Barnes (2004), Parker *et al.* (2004), Popham & Orrell (2012), Torrington *et al.* (2004), Torrington (2007)	Box D: Fair Box F: Fair	*Content validity* Item generation was based on literature review and earlier instruments. No figures presented *Construct validity* SCEAM was shown to possess construct validity to some extent. Hypotheses testing. Not formulated but possible to deduce what was expected. No comparator instrument (s) used. No figures are The tool has been able to discriminate between various environments in a range of studies
S-SCEAM	Nordin *et al.* (2015)	Box A: Fair Box D: Fair Box G: Good Box F: Good	*Content validity* Expert groups assessed the relevance by using content validity index (CVI) for items (I-CVI) and for scale (S-CVI) and interviews I-CVI above 0·89; S-CVI above 0·90 *Test–retest reliability* Test–retest reliability was examined by two independent raters showing high stability: 96% and 95% (κ = 0·903 and 0·869) Inter-rater reliability was measured on two rating occasions demonstrating high levels of agreement: 95% and 94% (κ = 0·851 and 0·832)

Table 3 (*Continued*).

Instrument*	References	COSMIN assessment[†]	Reliability/Validity
SCUEQS	Sloane *et al.* (2002)	Box A: Good Box B: Good Box D: Good Box F: Good	*Content validity* Item generation was based on literature review and earlier instruments. No figures presented *Concurrent criterion validity* Showed strong correlation with EAT (r = 0·85), and moderately strong correlation when compared with PEAP global scores (r = 0·52, P < 0·01) A significant negative correlation was found between SCUEQS scores and prevalence of residents agitation (r = −0·34, P < 0·01) *Inter-rater reliability* The inter-rater reliability was r = 0·84 *Cronbach's alpha* Cronbach's alpha was 0·78 in non-SCU dementia units and 0·63 for the non-SCU units *Interclass correlation coefficient (ICC)* Ranged from −0·07 to 0·88
TESS-NH	Bicket *et al.* (2012) Campo & Chaudhury (2012) Fleming & Purandare (2010) Fleming (2011) Slaughter *et al.* (2006) Sloane *et al.* (2002) Teresi *et al.* (2000)	Box A: Good Box B: Good Box D: Good Box F: Good Box E: Good	Validity tests were foremost performed with the shorter form of TESS-NH SCUEQS (see above) *Construct validity* TESS-NH could discriminate between different dementia care units *Concurrent validity* Global rating showed strong correlation with PEAP global scores (r = 0·71) Light meter levels at four locations correlate significantly with PEAP (r = 0·29–0·38) Showed strong concurrent validity when compared to the global score of TESS-NH (0·82), and the sum score of SCUEQS (0·85) Showed strong concurrent validity when compared to the global score of SCUEQS (0·92), and the sum score of SCUEQS (0·82) *Cohen's kappa* for 74% of the items was above 0·60 *Inter-rater reliability* The average percentage of absolute agreement between two raters was 84·4% (range 43–100%) *Test–retest reliability* Items indicated environmental factors that are fixed such as floor surface demonstrated high levels of test–retest reliability (above 0·80). Those items that reflect behaviour such as adequacy/evenness of lighting demonstrated moderate to substantial agreement *Cronbach's alpha* Four of the subscales have a Cronbach's alpha below the usually acceptable level of 0·6; two were not calculable; and seven were above the acceptable level *Interclass correlation coefficient (ICC)* Ranged from −0·05 to 1. 39·8% of the items exceeded 0·7 The global score had an ICC of 0·81
TESS-NH/RC	Sloane *et al.* (2002)	Box A: Good Box F: Good	*Construct validity* Factor analysis resulted in two factors; Dignity and Sensitivity that the 15 items logical could be divided into. The tool could discriminate between persons with more severe Alzheimer diagnose and quality of life and fall risks. Reported good internal reliability

*Abbreviation of instruments.
NA = not applicable, NR = not reported.
[†]Internal consistency (Box A), reliability (Box B), measurement error (Box C), content validity (Box D), structural validity (Box E), hypothesis testing (Box F), cross-cultural validity (Box G), criterion validity (Box H), responsiveness (Box I).

associated websites where the instruments are described and case studies using the instruments reported, but there was little available information regarding their validity and reliability.

Applicability and feasibility

Most the instruments demonstrated a rather weak empirical base. The instruments have not often been used outside of their period of development, or by actors other than their original developers or authors. This means that there is a weak basis for critically assessing both the applicability and feasibility of the instruments. The review identified only three instruments that had more widespread use: MEAP (Moos & Lemke 1996); PEAP (Lawton *et al.* 2000); and TESS-NH (Sloane *et al.* 2002). Of these instruments both MEAP and PEAP are rather old, having been developed during late 90s.

Information regarding e.g., the time needed for completion, usage costs, perceived difficulties in administration, training needs or availability of a user's guide was reported for some but far from all of the instruments. In many cases, the authors themselves described the instruments as easy to use and that no training was required before use. Both MEAP (Moos & Lemke 1996) and PEAP (Cutler *et al.* 2006) are described as complex in that a minimum of a 2-day course is required to learn about the instrument, followed by time-consuming data collection. The instruments are not recommended for use by non-researchers. EAT (Fleming 2011) and TESS-NH (Sloane *et al.* 2002) on the other hand are described as easier to use with guidance from published articles. SCEAM (Nordin *et al.* 2015) is comprehensive, involving many items but not complex to complete: it has been reported that it takes around 2 hours to complete the instrument depending on the size of facility being assessed and no specific training is needed.

Discussion

This is the first review of the reliability and validity of measurement instruments for assessing the quality of the physical environment in health care. The results demonstrate that there exists a rather large body of published instruments for measuring the quality of PHCEs. However, the review also illustrates several problems with the available instruments, with perhaps the most significant being that few appear to have been subjected to satisfactory validation procedures. The majority of the instruments were also developed during the early 90s and thus could be less relevant to a contemporary healthcare service that is focused on concepts such as person-centred care and interdisciplinary care. In addition, contemporary health care increasingly includes more knowledge from several disciplines such as nursing, which is not visible or highlighted in the early instruments.

Valid instruments are important for many reasons. First, rigorous assessment with valid instruments can contribute to the general development of high quality healthcare environments by discovering poor and inadequate design (Baird 2001, Gesler *et al.* 2004). Second, the assessment of design quality in healthcare environments can be integrated with routine strategic improvement work (Preiser 1995). A lack of valid instruments seriously constrains the ability to assess the quality of the PHCE and contribute to EBD.

Psychometric issues

Many of the instruments have not often been used beyond the specific context where they were developed, nor by actors other than their respective developers. External validation of an instrument requires a demonstration that the instrument has reliability outside its original development context. In general, psychometric information on the instruments is lacking, so that information such as item sensitivity, internal consistency of scales and so forth, are not available. Nor for the most part is any data provided on inter-rater reliability and test–retest reliability. Few of the studies explicitly stated that consideration was given to measurement test theory in the development or testing of the instruments. However, many of the instruments had been tested in ways related to classical measurement theory such as Cronbach's alpha (Table 3). One reason for the lack of application of other methods relating to measurement theory such as factor analysis may be their requirement for large studies, which is often difficult to realize in studies of PHCEs.

Conceptual framework, aim, and applicability

We found the conceptual framework and definitional precision of the instruments to be limited. While many of the instruments were justified on the basis of the long-held understanding of the important relationship between healthcare environments, safe care and patient well-being, there was little explicit attempt to move beyond this model. This limited use of theory in the development and testing of the instruments included in this study may reflect the more general state of the science in EBD and POE. There is still a lack of rigorous research on design and its impact on health and few evaluations of completed new buildings (Steinke

2015). The dominant theory that explicitly or implicitly informed many of the instruments was Lawton ecological model of ageing (Lawton & Nahemow 1973).

Many years have passed since the ecological model was first proposed and since the development of many of the instruments found in this review. For example, TESS was developed in the USA in the early 90s, since when much useful literature on environmental design has been published (Ulrich *et al.* 2010). The instrument reflects an institutional approach to residential care that was prevalent at the time. Given the advances in healthcare technology and procedures and the knowledge generated in the past few decades on how the environment impacts on patients' health and well-being, there is a question as to whether relatively old instruments have satisfactory applicability to contemporary healthcare environments. The development of new care models in recent years also has implications for the way healthcare environments should be designed to facilitate good quality care. Recently, person-centred care has been implemented in many healthcare settings and in this care approach the environment is seen as a central component (Edvardsson *et al.* 2010, Chenoweth *et al.* 2011). New instruments are therefore required that are based on evidence of how PHCEs have an impact on health and well-being and for emerging models of care. Such instruments also need to be embedded in current policy and perspectives on ageing. For example, the ecological model emerged before the literature on successful ageing and healthy ageing burgeoned (McKee & Schüz 2015). Given the dominant position in social and healthcare policy held by the healthy ageing paradigm, instruments that mesh the environmental perspective with healthy ageing could be of considerable utility (Wahl *et al.* 2012).

The majority of the instruments obtained were developed for use in healthcare environments for older people, several specifically for dementia care environments. It is possible that instruments designed for use in older people care settings might have applicability in other healthcare environments, but the application of instruments intended for one form of healthcare environment in a different environment would require careful monitoring and, potentially, adaptation of the instrument.

Since LEED (Shulman 2003) and BREEAM (Schweber & Haroglu 2014) were developed, there has been a shift in focus from green buildings towards sustainability including a building's entire life span. Very little research has been carried out using LEED and BREEAM, especially in healthcare settings (Schweber & Haroglu 2014). When searching in databases using LEED and BREEAM, we found many articles describing the structure of the instruments and comparisons between them but very few studies on the use of them in real projects. In addition, authors have proposed LEED and BREEAM as design tools for supporting dialog among stakeholders and as vehicles for specification of sustainable values and goals, although few have studied their use in such contexts (Schweber 2013).

Strengths and limitations

We faced a particular challenge in that research concerning healthcare environments is still limited and a cohesive body of literature of measurement instruments is lacking. This area of research exists on the border between more science-based disciplines with traditional modes of publication and with a focus on validation and reliability and more practitioner-based and humanities-oriented disciplines where experience, expertise, and intuition are valued above scientific proof. Many of the instruments developed in the fields of architecture, planning, and construction have not been developed using research methods and used in research and therefore not easily found in regular research databases. Literature on instruments for assessing quality in healthcare environments has been published in a range of forms, from peer-reviewed academic journals to non-academic, non-peer-reviewed papers. Research on healthcare environments is poorly indexed, thus making it difficult to perform a sensitive and specific search. This is further complicated by diverse keywords and publication strategies. As a result and given the multidisciplinary focus of the review, a broad framework was required to gather data for the reviewing process drawn from various disciplines that use differing methodological approaches. Given our broad search strategy ensuring data retrieval across a wide range of databases and our manual review of the bibliography of retrieved papers, we are confident that most of the relevant papers and articles were captured. However, the authors are aware of the existence of 'centres of excellence' for EBD in healthcare environments, such as the Center for Health Design (https://www.healthdesign.org/), involved in the development of instruments and procedures to ensure quality in healthcare environments and improve healthcare outcomes, whose instruments unfortunately have yet to be documented in published research studies and which therefore fall out with the remit of this review.

Lastly, our data extraction was ambitious with respect to psychometric characteristics using the established COSMIN checklist but unfortunately this important information was mostly not reported to recommended standards.

Conclusions

We have summarized the range of published measurement instruments for PHCEs as a resource for quality assurance of environments that support high quality and safe care and good working conditions. The target groups for this review are healthcare managers, those responsible for planning or/and building healthcare environments and researchers in care and architecture. Although many instruments for measuring the quality of the PHCE have been published, none met all of our criteria for robustness. Most lacked strong, up-to-date theoretical foundations, while many instruments had been used to only limited extents in research contexts or beyond the settings where they were originally developed. In addition, psychometric data were found to be severely lacking for many of the instruments.

It would be wrong to select any one of the reviewed instruments as the 'most fit for purpose' since the instruments vary considerably in their aim, comprehensiveness, target environment, and level of use. However, some instruments performed better than others on our assessment criteria and in our psychometric evaluation and so can be cautiously recommended for use. PEAP, MEAP, and TESS-NH come with some validation or reliability data and are comprehensive instruments for measuring the quality of the PHCE, although primarily with application in care facilities for older people. PHQI is the newest instrument in this review and the developers have also conducted a relatively thorough validation procedure. The instrument represents one of the few instruments created to measure users' perception of environmental quality in hospitals and combine physical and social aspects of the environment. SCEAM is also quite new and has potential given its comprehensive nature, its development in a theoretical framework that has the needs of the older person at the centre and its initial psychometric performance. However, further information on all these instruments' reliability, validity, and applicability are clearly warranted.

More research is needed to develop instruments that are theoretically well-grounded and predicated on current or emerging models of care and appropriate for measuring modern healthcare environments. In particular, a broader understanding of the healthcare environment should inform further development work, so that in the future instruments emerge that can integrate data on engineering and sustainability factors with data on the interaction between environmental features and users and which are founded on a strong theoretical framework that has the needs of users in the centre. None of the instrument included in this review offers such a comprehensive engagement with the PHCE, but it is possible that some of the instruments could be used as a starting point in the development process.

Funding

This work was supported by internal research funds made available via the Health and Welfare research theme, Dalarna University.

Conflicts of interest

Kevin McKee was a co-investigator on the research project that developed the SCEAM instrument.

Author contributions

All authors have agreed on the final version and meet at least one of the following criteria [recommended by the ICMJE (http://www.icmje.org/recommendations/)]:

- substantial contributions to conception and design, acquisition of data or analysis and interpretation of data;
- drafting the article or revising it critically for important intellectual content.

Supporting Information

Additional supporting information may be found online in the supporting information tab for this article.

References

Abbas M.Y. & Ghazali R. (2011) Physical environment: the major determinant towards the creation of a healing environment? *Procedia – Social and Behavioral Sciences* **30**, 1951–1958.

Aiken L.H., Clarke S.P. & Sloane D.M. (2002) Hospital staffing, organization and quality of care: cross-national findings. *Nursing Outlook* **50**(5), 187–194.

Anåker A., Heylighen A., Nordin S. & Elf M. (2016) Design quality in the context of healthcare environments: a scoping review. *HERD: Health Environments Research & Design Journal* 1937586716679404.

Andrade C., Lima M.L., Fornara F. & Bonaiuto M. (2012) Users' views of hospital environmental quality: validation of the Perceived Hospital Environment Quality Indicators (PHEQIs). *Journal of Environmental Psychology* **32**(2), 97–111.

Andrade C.C., Lima M.L., Pereira C.R., Fornara F. & Bonaiuto M. (2013) Inpatients' and outpatients' satisfaction: the mediating role of perceived quality of physical and social environment. *Health & Place* **21**, 122–132.

Arneill A.B. & Devlin A.S. (2002) Perceived quality of care: the influence of waiting room environment. *Journal of Environmental Psychology* **22**(4), 345–360.

Baird G. (2001) Post occupancy evaluation and Probe: a New Zeeland perspective. *Building Research & Information* **29**(6), 469–472.

Barnes S. (2004) The design of caring environments and the quality of life of older people. *Ageing and Society* **22**(6), 775–789.

Benjamin L.C. & Spector J. (1990) Environments for the dementing. *International Journal of Geriatric Psychiatry* **5**(1), 1099–1166.

Benjamin L.C. & Spector J. (1992) Geriatric care on a ward without nurses. *International Journal of Geriatric Psychiatry* **7** (10), 743–750.

Bicket M.C., Samus Q.M., McNabney M., Onyike C.U., Mayer L.S., Brandt J., Rabins P., Lyketsos C. & Rosenblatt A. (2012) The physical environment influences neuropsychiatric symptoms and other outcomes in assisted living residents. *International Journal of Geriatric Psychiatry* **25**(10), 1044–1054.

Braun B.I. (1991) The effect of nursing home quality on patient outcome. *Journal of The American Geriatrics Society* **39**(4), 329–338.

Bromley E. (2012) Building patient-centeredness: hospital design as an interpretive act. *Social Science & Medicine* **75**(6), 1057–1066.

Campo M. & Chaudhury H. (2012) Informal social interaction among residents with dementia in special care units: Exploring the role of the physical and social environments. *Dementia* **11** (3), 401–423.

Charise A., Witteman H., Whyte S., Sutton E.J., Bender J.L., Massimi M., Stephens L., Evans J., Logie C., Mirza R.M. & Elf M. (2011) Questioning context: a set of interdisciplinary questions for investigating contextual factors affecting health decision making. *Health Expectations* **14**(2), 115–132.

Chenoweth L., King M., Luscombe G., Forbes I., Jeon Y.-H., Parbury J.S., Brodaty H., Fleming R. & Haas M. (2011) Study protocol of a randomised controlled group trial of client and care outcomes in the residential dementia care setting. *Worldviews on Evidence-Based Nursing* **8**(3), 153–165.

Codinhoto R., Aouad G., Kagioglou M., Tzortzopoulos P. & Cooper R. (2009) Evidence-based design of health care facilities. *Journal of Health Service Research Policy* **14**(4), 194–196.

Craik K. & Femer N. (1987) Environmental assessment. In *Handbook of Environmental Psychology* (Stockolos D. & Altman I., eds), John Wiley, New York.

Crawley D. & Aho I. (2006) Building environmental assessment methods: application and development trends. *Building Information and research* **27**(4–5), 300–308.

Cunningham C. (2009) Auditing design for dementia. *Journal of Dementia Care* **17**(3), 31–32.

Cutler L.J. (2007) Physical environments of assisted living: research needs and challenges. *The Gerontologist* **47**(suppl 1), 68–82.

Cutler L.J. & Kane R.A. (2009) Post-occupancy evaluation of a transformed nursing home: the first four Green House® settings. *Journal of Housing for the Elderly* **23**(4), 304–334.

Cutler L.J., Kane R.A., Degenholtz H.B., Miller M.J. & Grant L. (2006) Assessing and comparing physical environments for nursing home residents: using new tools for greater research specificity. *Gerontologist* **46**(1), 42–51.

Davidson L., Tebes J.K., Rakfeldt J. & Sledge W.H. (1996) Differences in social environment between inpatient and day hospital-crisis respite settings. *Psychiatric Services* **47**(7), 714–720.

Day K., Carreon D. & Stump C. (2000) The therapeutic design of environments for people with dementia: a review of the empirical research. *The Gerontologist* **40**(4), 397–416.

Design Quality Indicator Group (2014) Retrieved from http://www.dqi.org.uk/case-studies/healthcare/ on 25 February 2017.

Dewulf G. & van Meel J. (2004) Sense and nonsense of measuring design quality. *Building Research & Information* **32**(3), 247–250.

Dijkstra K., Pieterse M. & Pruyn A. (2006) Physical environmental stimuli that turn healthcare facilities into healing environments through psychologically mediated effects: systematic review. *Journal of Advanced Nursing* **56**(2), 166–181.

Edvardsson D. (2008) Therapeutic environments for older adults: constitutes and meanings. *Journal of Gerontological Nursing* **34** (6), 32–40.

Edvardsson D., Fetherstonhaugh D. & Nay R. (2010) Promoting a continuation of self and normality: person-centred care as described by people with dementia, their family members and aged care staff. *Journal of Clinical Nursing* **19**(17/18), 2611–2618.

Edvardsson D., Sandman P.-O. & Rasmussen B. (2012) Forecasting the ward climate: a study from a dementia care unit. *Journal of Clinical Nursing* **21**(7/8), 1136–1114.

Eisen S.L., Ulrich R.S., Shepley M.M., Varni J.W. & Sherman S. (2008) The stress-reducing effects of art in pediatric health care: art preferences of healthy children and hospitalized children. *Journal of Child Health Care: For Professionals Working with Children in the Hospital and Community* **12**(3), 173–190.

Evans G.W. & McCoy J.M. (1998) When buildings don't work: the role of architecture in human health. *Journal of Environmental Psychology* **18**(1), 85–94.

Evans K., Spiby H. & Morrell C. (2015) A psychometric systematic review of self-report instruments to identify anxiety in pregnancy. *Journal of Advanced Nursing* **71**(9), 1986–2001.

Field E., Walker M., Hancock G. & Orrell M. (2005) The needs of older people in sheltered housing: a comparison of inner city and new town areas. *Journal of Housing for the Elderly* **19**(2), 107–117.

Fleming R. (2011) An environmental audit tool suitable for use in homelike facilities for people with dementia. *Australasian Journal on Ageing* **30**(3), 108–112.

Fleming R. & Bennett K. (2015) Assessing the quality of environmental design of nursing homes for people with dementia: Development of a new tool. *Australasian Journal on Ageing* **34**(3), 191–194.

Fleming R. & Purandare N. (2010) Long-term care for people with dementia: environmental design guidelines. *International Psychogeriatrics/IPA* **22**(7), 1084–1096.

Fleming R., Fay R. & Robinson A. (2012) Evidence-based facilities design in health care: a study of aged care facilities in Australia. *Health Services Management Research: An Official Journal Of The Association Of University Programs In Health Administration* **25**(3), 121–128.

Fornara F., Bonaiuto M. & Bonnes M. (2006) Perceived hospital environment quality indicators: a study of orthopaedic units. *Journal of Environmental Psychology* **26**(4), 321–334.

Fonda S.J., Maddox G.L., Clipp E. & Reardon J. (1996) Design for a Longitudinal Study of the impact of an enhanced environment on the functioning of frail adults. *Journal of Applied Gerontology* **15**(4), 397–413.

Foureur M., Davis D., Fenwick J., Leap N., Iedema R., Forbes I. & Homer C.S. (2010) The relationship between birth unit design and safe, satisfying birth: developing a hypothetical model. *Midwifery* **26**(5), 520–525.

Foureur M.J., Leap N., Davis D.L., Forbes I.F. & Homer C.E. (2011) Testing the Birth Unit Design Spatial Evaluation Tool (BUDSET) in Australia: a pilot study. *Health Environment Research and Design* **4**(2), 36–60.

Gann D. & Whyte J. (2003) Design quality, its measurement and management in the built environment. *Building Research & Information* **31**(5), 314–317.

Gann D.M., Salter A.J. & Whyte J.K. (2003) Design quality indiators:tools for thinking. *Building Research & Information* **31** (5), 318–333.

Gesler W., Bell M., Curtis S., Hubbard P. & Francis S. (2004) Therapy by design: evaluating the UK hospital building program. *Health & Place* **10**(2), 117–128.

Ghazali R. & Abbas M.Y. (2012) Assessment of healing environment in paediatric wards. *Procedia – Social and Behavioral Sciences* **38**, 149–159.

Gifford R., Hine W.D., Muller-Clemm W., Reynolds J., D'Arcy J. & Shaw K.T. (2000) Decoding modern architecture. A lens model approach for understanding the aesthetic differences of architects and laypersons. *Environment and Behavior* **32**(2), 163–187.

Grant L. (1996) Assessing environments in alzheimer special care units. *Nursing Unit Rating Scale- Research on Aging* **18**(3), 275–291.

Hamilton K. & Watkins D.H. (2009) *Evidence-Based Design for Multiple Building Types*. John Wiley & Sons, New Jersey, USA.

Happio A. & Viitaniemi P. (2008) A critical review of buildning assssment tools. *Environmental Impact Assessment Review* **28**, 469–469.

Henriksen K., Isaacson S., Sadler L.B. & Zimring M.C. (2007) The role of the physical environment in crossing the quality chasm. *The Joint Commission Journal on Quality and Patient Safety* **33** (11 Supplement), 68–80.

Heylighen A. & Bianchin M. (2013) How does inclusive design relate to good design? Designing as a deliberative enterprise. *Design Studies* **34**, 93–110.

van Hoof J., Kort H.S.M., van Waarde H. & Blom M.M. (2010) Environmental interventions and the design of homes for older adults with dementia: an overview. *American Journal of Alzheimer's Disease & Other Dementias* **25**(3), 202–232.

Huisman E.R.C.M., Morales E., van Hoof J. & Kort H.S.M. (2012) Healing environment: a review of the impact of physical environmental factors on users. *Building and Environment* **58**, 70–80.

Iwarsson S., Nygren C. & Slaug B. (2005) Cross-national and multi-professional inter-rater reliability of the Housing Enabler. *Scandinavian Journal of Occupational Therapy* **12**(1), 29–39.

Izal M. (1992) Residential facilities for older adults: cross-cultural environmental assessment. *European Journal of Psychological Assessment* 8(2), 118–134.

Janssen H., Ada L., Bernhardt J., McElduff P., Pollack M., Nilsson M. & Spratt N.J. (2014) An enriched environment increases activity in stroke patients undergoing rehabilitation in a mixed rehabilitation unit: a pilot non-randomized controlled trial. *Disability & Rehabilitation* **36**(3), 255–262.

Kelly F., Innes A. & Dincarslan O. (2011) Improving care home design for people with dementia. *Journal of Care Services Management* **5**(3), 147–155.

Lawton M.P. & Nahemow L. (1973) *Ecology and the aging process*. In *The Psychology Of Adult Development and Aging* (Eisdorfer C. & Lawton M.P., eds), American Psychological Association, Washington DC, pp. 619–674.

Lawton M.P., Weisman G.D., Sloane P. & Calkins M. (1997) Assessing environments for older people with chronic illness. *Journal of Mental Health & Aging* **3**(1), 83–100.

Lawton M.P., Weisman G.D., Sloane P., Norris-Baker C., Calkins M. & Zimmerman S.I. (2000) Professional environmental assessment procedure for special care units for elders with dementing illness and its relationship to the therapeutic environment screening schedule. *Alzheimer Disease and Associated Disorders* **14**(1), 28–38.

Leaman A., Stevenson F. & Bordass B. (2010) Building evaluation: practice and principles. *Building Research & Information* **38**(5), 564–577.

Lemke S. & Moos R.H. (1986) Quality of residential settings for elderly adults. *Journal of Gerontology* **41**(2), 268–276.

Lewis A., Torrington J., Barnes S., Darton R., Holder J., McKee K., Netten A. & Orrell A. (2010) EVOLVE: a tool for evaluating the design of older people's housing. *Housing, Care and Support* **13**(3), 31–40.

Liberati A., Altman D.G., Tetzlaff J., Mulrow C., Gøtzsche P.C., Ioannidis J.P.A., Clarke M., Devereaux P.J., Kleijnen J. & Moher D. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology* **62**(10), e1–e34.

Linney J.A., Arns P.G., Chinman M.J. & Frank J. (1995) Priorities in community residential care: a comparison of operators and mental health service consumers. *Psychiatric Rehabilitation Journal* **19**(1), 27–34.

Markus T. (2009) Lessons from the design quality indicator. *Building Research & Information* **31**(5), 399–405.

McKee K.J. & Schüz B. (2015) Psychosocial factors in healthy ageing. *Psychology & Health* **30**(6), 607–626.

Mokkink L.B., Terwee C.B., Patrick D.L., Alonso J., Stratford P.W., Knol D.L., Bouter L.M. & de Vet H.C.W. (2010) The COSMIN study reached international consensus on taxonomy, terminology and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology* **63**(7), 737–745.

Moos R.H. & Lemke S. (1996) *Evaluating Rsesidential Facilities: The Multiphasic Environmental Assessment Procedure*. Sage Publications, Thousand Oaks, CA.

Morgan D.G., Stewart N.J., D'Arcy C. & Werezak L.J. (2004) Evaluating rural nursing home environments: dementia special care units versus integrated facilities. *Aging & Mental Health* **8** (3), 256–265.

Mourshed M. & Zhao Y. (2012) Healthcare providers' perception of design factors related to physical environments in hospitals. *Journal of Environmental Psychology* **32**(4), 362–370.

Nightingale F. (1820/1910) *Florence Nightingale's Notes on nursing: what it is, and what it is not*. Springer, New York.

Nordin S., Elf M., McKee K. & Wijk H. (2015) Assessing the physical environment of older people's residential care facilities: development of the Swedish version of the Sheffield Care Environment Assessment Matrix (S-SCEAM). *BMC Geriatrics* **15** (1), 3–3.

Orrell A., McKee K., Torrington J., Barnes S., Darton R., Netten A. & Lewis A. (2013) The relationship between building design and residents' quality of life in extra care housing schemes. *Health & Place* **21**, 52–64.

Parker C., Barnes S., McKee K.J., Morgan K., Torrington J.M. & Tregenza P.R. (2004) Quality of life and building design in residential and nursing homes for older people. *Ageing and Society* **24**(6), 941–962.

Popham C. & Orrell M. (2012) What matters for people with dementia in care homes? *Aging & Mental Health* **16**(2), 181–188.

Preiser W. (1995) Post-occupancy evaluation: how to make buildings work better. *Facilities* **13**(11), 19–28.

Schweber L. (2013) The effect of BREEAM on clients and construction professionals. *Building Research & Information* **41** (2), 129–145.

Schweber L. & Haroglu H. (2014) Comparing fit between BREEAM assessment and design processes. *Building Research & Information* **42**(3), 300–317.

Sheehy A., Foureur M., Catling-Paull C. & Homer C. (2011) Examining the content validity of the birthing unit design spatial evaluation tool within a woman-centered framework. *Journal of Midwifery & Women's Health* **56**(5), 494–502.

Shulman K. (2003) Art of the possible: a new ambulatory-care clinic for a remarkable institution seeks to become the nation's first LEED-certified healthcare facility. *Metropolis* **23**(2), 124–127.

Sikorska-Simmons E. (2006) The effects of organizational context on autonomy-enhancing policies in assisted living. *Journal of Aging Studies* **20**(3), 217–226.

Slaughter S.E. & Morgan D.G. (2012) Functional outcomes of nursing home residents in relation to features of the environment: validity of the professional environmental assessment protocol. *Journal of the American Medical Directors Association* **13**(5), 487.e1–7.

Slaughter S., Calkins M., Eliasziw M. & Reimer M. (2006) Measuring physical and social environments in nursing homes for people with middle- to late-stage dementia. *Journal of the American Geriatrics Society* **54**(9), 1436–1441.

Sloane P.D., Mitchell C.M., Weisman G., Zimmerman S., Foley K.M.L., Lynn M., Calkins M., Lawton M.P., Teresi J., Grant L., Lindeman D. & Montgomery R. (2002) The Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH): an observational instrument for assessing the physical environment of institutional settings for persons with dementia. *Journals of Gerontology Series B: Psychological Sciences & Social Sciences* **57**(2), 69–78.

Smith R., Fleming R., Chenoweth L., Jeon Y.-H., Stein-Parbury J. & Brodaty H. (2012) Validation of the Environmental Audit Tool in both purpose-built and non-purpose-built dementia care settings. *Australasian Journal on Ageing* **31**(3), 159–163.

Stankos M. & Schwartz B. (2007) Evidence-based design in healthcare: a theoretical dilemma. *Interdisciplinary Design and Research e-Journal* **1**(1).

Steinke C. (2015) Assessing the physical service setting: a look at emergency departments. *Health Environment Research and Design Journal* **8**(2), 31–42.

Steinke C., Webster L. & Fontaine M. (2010) Evaluating building performance in healthcare facilities: an organizational perspective. *Health Environment Research and Design* **3**(2), 63–83.

Swedish Standard Institute. (2014) Quality of care, service nursing and rehabilitation for elderly people with extensive needs in ordinary and residential care facilities - Certification scheme of the services of SS 872500:2015.

Teresi J.A., Holmes D. & Ory M.G. (2000) The therapeutic design of environments for people with dementia: further reflections and recent findings from the national institute on aging collaborative studies of dementia special care units. *The Gerontologist* **40**(4), 417–421.

Terwee C.B., Mokkink L.B., Knol D.L., Ostelo R.W., Bouter L.M. & de Vet H.C. (2012) Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research* **21**(4), 651–657.

Thomson D., Austin S., Devine-Wright H. & Mills G. (2003) Managing value and quality in design. *Building Research & Information* **5**(3), 334–345.

Timko C. & Moos R.H. (1990) Determinants of interpersonal support and self-direction in group residential facilities. *Journal of Gerontology* **45**(5), 184–192.

Timko C. & Moos R.H. (1991) Assessing the quality of residential programs: Methods and applications. *Adult Residential Care Journal* **5**(1), 113–129.

Torrington J. (2007) Evaluating quality of life in residential care buildings. *Building Research & Information* **35**(5), 514–528.

Torrington J., Barnes S., McKee K., Parker C., Morgan K. & Tregenza P. (2004) The influence of building design on the quality of life of older people. *Architectural Science Review* **47**, 193–197.

Ulrich R.S., Zimring C., Zhu X., DuBose J., Seo H.B. & Choi Y.S. (2008) A review of the research literature on evidence-based health-care design. *Health Environment Research and Design* **1** (3), 61–125.

Ulrich R.S., Berry L.L., Quan X. & Parish J.T. (2010) A conceptual framework for the domain of evidence-based design. *Health Environment Research and Design* **4**(1), 95–114.

Vischer J. (2008) Towards a user-centred theory of built environment. *Building Research and Information* **36**(3), 231–240.

Vischer J. (2009) Applying knowledge on building performance: from evidence to intelligence. *Intelligent Building International* **1**, 239–248.

Volker L., Lauche K., Heintz J.L. & de Jonge H. (2008) Deciding about design quality: design perception during a European tendering procedure. *Design Studies* **29**(4), 387–409.

Wahl H.W., Iwarsson S. & Oswald F. (2012) Aging well and the environment: toward an integrative model and research agenda for the future. *Gerontologist* **52**(3), 306–316.

Weisman G. (1994) American journal of alzheimer's disease & other dementias. *Nursing Homes: Long Term Care Management* **43**(7), 46–47.

Wells L. M. & Taylor L. E. (1991) Empowering older people in residential care. *Adult Residential Care Journal*.

Whyte J. & Gann D. (2003) Design Quality Indicators: work in progress. *Building Research & Information* **31**(5), 387–398.

Zeisel J. (2003) Marketing therapeutic environment for alzheimers care. *Journal of Architectural and Planning Research* **20**(1), 75–86.

Zeisel J., Hyde J. & Levkoff S. (1994) Best practices: an Environment Behavior (EB) model for Alzheimer special care units. *American Journal of Alzheimer's Disease and Other Dementias* **9**(2), 4–21.

Zeisel J., Silverstein N.M., Hyde J., Levkoff S., Lawton M.P. & Holmes W. (2003) Environmental correlates to behavioral health outcomes in Alzheimer's special care units. *Gerontologist* **43**(5), 697–711.

Zimmerman A. & Martin M. (2001) Post-occupancy evaluation: benefits and barriers. *Building Research & Information* **29**(12), 168–174.

