A comparative field usability study of two lighting measurement protocols

Roberto G. Rodriguez*

Instituto de Ambiente, Hábitat y Energía (INAHE) – CONICET,
CCT Mendoza CONICET – Ruiz Leal N/N Avenue,
Mendoza, Argentina
and
Facultad de Arquitectura,
Urbanismo y Diseño – Universidad de Mendoza,
Boulogne Sur Mer 683, Mendoza, Argentina
Email: rgrodriguez@mendoza-conicet.gob.ar
*Corresponding author

Juan Manuel Monteoliva and Andrea E. Pattini

Instituto de Ambiente, Hábitat y Energía (INAHE) – CONICET,
CCT Mendoza CONICET – Ruiz Leal N/N Avenue,
Mendoza, Argentina
Email: jmonteoliva@mendoza-conicet.gob.ar
Email: apattini@mendoza-conicet.gob.ar

Abstract: We developed a lighting measurement protocol (PC-SRT) for Android mobile devices as a replacement of the Argentinean currently mandatory pen-and-paper measurement protocol (SRT-P), and compared their usability in a field study (n = 26) by means of the system usability scale (SUS). Descriptive statistics showed that PC-SRT (SUS = 58.7; SD = 14.9; lower marginally acceptable usability) outscored SRT-P (SUS=47.5; SD=14.5; unacceptable usability). The PC-SRT also performed better in both ease of use and learnability factors. An item-by-item analysis compared the behaviour of each standardised item score; Mann-Whitney test results showed statistically significant differences in SUS-Q1 (U = 36; p = 0.011), SUS-Q2 (U = 35; p = 0.01), SUS-Q4 (U = 46.5; p = 0.05), and SUS-Q7 (U = 39; p = 0.019). These results show that although PC-SRT gathers more data in a wider variety of aspects of the visual environment, users tended to perceive it as easier to use, encouraging the adoption of our proposed lighting measurement protocol for mobile devices.

Keywords: usability; comparative field study; lighting measurement protocol; system usability scale; SUS; human factors.

Reference to this paper should be made as follows: Rodriguez, R.G., Monteoliva, J.M. and Pattini, A.E. (xxxx) 'A comparative field usability study of two lighting measurement protocols', *Int. J. Human Factors and Ergonomics*, Vol. X, No. Y, pp.xxx–xxx.

Biographical notes: Roberto G. Rodriguez is an Industrial Designer, who graduated in 2003 from the Cuyo National University (Mendoza, Argentina). In 2012 he completed his PhD in Visual Environments and Efficient Lighting

from the Tucuman National University, Tucuman, Argentina. He is currently an assistant researcher at the CONICET INAHE – CCT Mendoza, Argentina and he is also a Tenured Professor of Ergonomics Applied to Design in the University of Mendoza. He is an author of several national and international peer-reviewed papers.

Juan Manuel Monteoliva is an Industrial Designer, who graduated in 2010 from the Cuyo National University, Mendoza, Argentina. In 2014 he completed his PhD in Visual Environments and Efficient Lighting from the Tucuman National University, Tucuman, Argentina. He is currently an assistant researcher at the CONICET INAHE – CCT Mendoza, Argentina and also a Professor of Industrial Legislation and Lighting Products Design in the Cuyo National University. He is an author of several national and international peer-reviewed papers.

Andrea E. Pattini is an Industrial Designer, who graduated in 1985 from the Cuyo National University, Mendoza, Argentina. In 2007 she completed her PhD in Visual Environments and Efficient Lighting from the Tucuman National University, Tucuman, Argentina. She is currently a main researcher at the CONICET and the Director of the INAHE – CCT Mendoza, Argentina. She is the Academic Co-Director of the Master in Sustainable Development of the Human Habitat at the National Technological University, Mendoza, Argentina. She directed several graduate students and authored several national and international peer-reviewed papers, book chapters, and a book on daylighting.

This paper is a revised and expanded version of a paper entitled ‘Desarrollo de una aplicación de dispositivos móviles para la medición de iluminación en espacios de trabajo’ presented at the 2nd Congreso Iberoamericano de Ergonomía, Montevideo, Uruguay, June 2017.

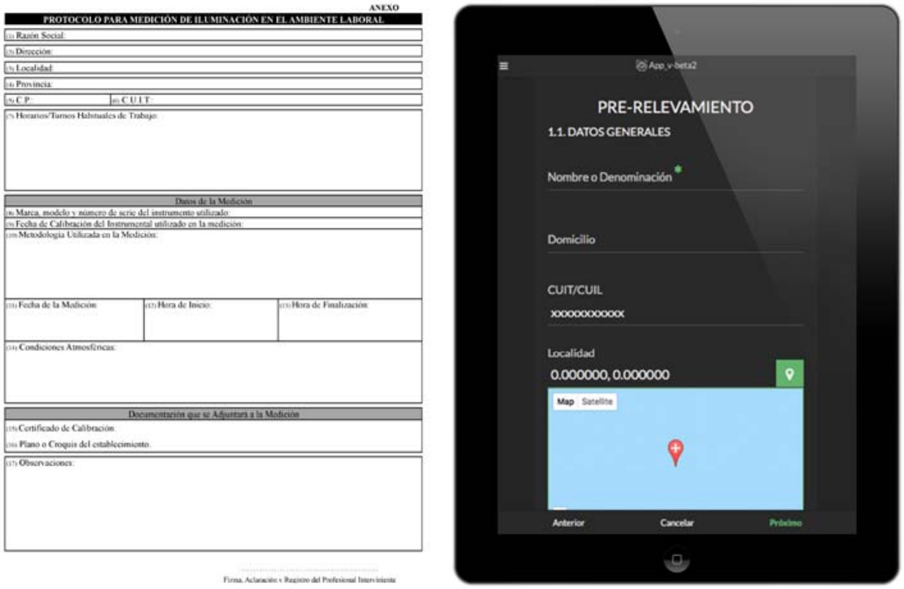
1 Introduction

A visual environment conceived to suit the capabilities and limitations of human vision may help to reach higher employee productivity in terms of visual performance, fewer errors and accidents, better safety, and lower absenteeism. But more importantly, good lighting has health and well-being advantages for the workers themselves (Van Bommel and Van den Beld, 2004). In other words, good lighting allows specific users to achieve specific goals in a highly usable visual environment, understanding usability as quality of use within a context (ISO 9241/11, 1998).

Scientific-based knowledge is made available to lighting practitioners through methods, indexes, indicators, guidelines and recommendations. Then, some of them may reach the status of standards or codes, the former usually of voluntary application, and the latter mandatory. However this process is neither linear nor straightforward, since it involves various actors and brings many – sometimes competing-forces into play. According to Boyce (1996) such forces are both practical (e.g., technological) and political (e.g., financial and emotional). Hence, scientific findings are weighted by a balance of those forces. A clear example of this is the wide variety of horizontal illuminance (E) level recommendations for the same given task that can be found around the globe (Mills and Borg, 1999). The Argentinean legal framework of lighting for working environments also followed the aforementioned path. Our lighting standards, the

IRAM-AADL series date from the late 1960s, and they are essentially faithful copies of their contemporary European counterparts. The recommended minimum E levels from IRAM-AADL j20-06 (1969) became mandatory through the resolution 351/79 (1979). Finally, in 2012, the resolution 84/12 mandated a pen-and-paper measurement protocol to verify in-field the compliance of resolution 351/79. We will refer to this protocol as the SRT protocol (SRT-P).

Figure 1 (a) SRT-P pen-and-paper protocol (b) PC-SRT Alternative mobile protocol (see online version for colours)



ANEXO

PROTOCOLO PARA MEDICIÓN DE ILUMINACIÓN EN EL AMBIENTE LABORAL

(1) Razón Social:

(2) Dirección:

(3) Localidad:

(4) Provincia:

(5) C.P.: (6) C.U.I.T.:

(7) Horarios/Turnos Habituales de Trabajo:

Datos de la Medición

(8) Marca, modelo y número de serie del instrumento utilizado:

(9) Fecha de Calibración del Instrumental utilizado en la medición:

(10) Metodología Utilizada en la Medición:

(11) Fecha de la Medición: (12) Hora de Inicio: (13) Hora de Finalización:

(14) Condiciones Atmosféricas:

Documentación que se Adjuntará a la Medición

(15) Certificado de Calibración:

(16) Plano o Croquis del establecimiento:

(17) Observaciones:

Firma, Aclaración y Registro del Profesional Interviniente

App_v-beta2

PRE-RELEVAMIENTO

1.1. DATOS GENERALES

Nombre o Denominación *

Domicilio

CUIT/CUIL

xxxxxxxxxxx

Localidad

0.000000, 0.000000

(a) (b)

A critical analysis of the methodology proposed by resolution 84/12 defined the SRT-P as a useful tool to systematise the gathering of horizontal illuminance in a grid, allowing through 41 items a quantitative analysis of some aspects of artificial lighting alone (Pattini et al., 2012). However, good lighting is much more than ensuring a sufficient and uniform level of artificial lighting in the working plane. Satisfying only visual requirements results in dull lighting scenarios. There are many other aspects to consider in order achieving good quality lighting, which positively impacts on task performance, energy consumption, and human health. A comparative analysis of the Argentinean legal framework on lighting in relation to both international regulations (ISO 8995, 2002; CEN, 2002) and good lighting practices for indoor workplaces (IES, 2011), allowed us to conceptualise an alternative lighting measurement protocol, and then to materialise it in the form of a software application for Android mobile devices, which we called PC-SRT. This novel protocol aims to provide new metrics and better criteria for the comprehensive assessment of the visual environment in work spaces. Figure 1 shows the mandatory and the alternative protocols. Modern ergonomics is evolving very dynamically (Gašová et al., 2017), and our developed app helps to solve the problem of performing quick and non-intrusive evaluations of workplace lighting from the view of ergonomics. Within this development, we coined the term ‘m-ergonomics’ (mobile ergonomics) which must not be understood in the straightforward sense of ergonomics of mobile devices but as the practice of ergonomics supported by mobile technology.

2 Part A: alternative lighting protocol PC-SRT

2.1 Background

Cuttle (2010) proposed different stages of the lighting profession. Synthetically, the first stage was mainly interested in determining the amount of illumination required for visual performance. The second stage termed the notion of lighting quality, which incorporated economical and psychological aspects to lighting design, as well as sustainability concerns (Veitch and Newsham, 1998). Finally, the third stage, currently in progress, addresses issues related to the non-visual and physiological effects of light, aiming to achieve healthy lighting (Rea et al., 2002). The SRT-P clearly belongs to the first era of lighting because of the prominence of E levels at the work-plane. Being easy to calculate, to measure, and convenient to codify, E is still nowadays nearly the one and only criterion used in lighting design. Illuminance is a very robust criterion; however, the search of single light level targets alone to satisfy all users all the time has proved to be too simplistic. Boyce (1996) defined illuminance selection based on visual performance as a fairy story. We define illuminance levels as a burden in lighting codes and regulations. Considerable efforts have been done to alleviate this burden during the second and third eras of the lighting profession, resulting in a more comprehensive approach to lighting for indoor working environments in newer normative. The European normative CEN (2002) ‘lighting of indoor workplaces’ seeks for good lighting that enables people to see, to move around safely and perform visual tasks efficiently, accurately and safely without causing undue visual fatigue and discomfort, regardless the lighting sources, which could be daylight, electric light or a combination of both. Based on ISO 8995 (2002) and CEN (2002) specifies lighting requirements for indoor workplaces (including visual display terminal work) in terms of quantity and quality. It defines the luminous environment from:

1. illuminance (average E, E uniformity and E ratios)
2. luminance distribution (luminance ratios in the visual field)
3. discomfort glare (unified glare rating – UGR-tabular method, shielding angles)
4. light directionality (modelling)
5. colour aspects of the source (correlated colour temperature and colour rendering index – CRI)
6. flicker and stroboscopic effect
7. daylighting.

It presents tables where it establishes, by task and activity, minimum values of maintained E and of CRI, and UGR thresholds. Regarding the illuminance, EN12464-1 presents an E scale starting from 20 lx (minimum to discern features of a human face under normal conditions), up to 5,000 lx in jumps of a factor of 1.5 (minimum difference of E that can be perceived by the human visual system). This minimum value can be weighted if the visual conditions differ from the normal assumptions, climbing at least one point on the scale according to the expert's criteria.

Taking one step further towards good lighting practice as laid down in CEN (2002) and Dehoff (2007) proposed the ergonomic lighting indicator (ELI). This measure uses five major criteria related to the most important human aspects in lighting to describe the overall quality of a lighting installation:

1. visual performance
2. vista (view of a scene)
3. visual comfort
4. vitality
5. empowerment (to influence the lighting).

They are rated on a scale from 1 ('poor') to 5 ('excellent'), on the basis of specific checklists, totalising 27 sub-criteria in three levels of lighting quality: visual requirements, emotional requirements and biological requirements. In conjunction with lighting energy indicator (LENI), the balance between human and energy requirements of a lighting environment can be evaluated (Dehoff, 2012). Based on the calculation procedures of CEN (2007), LENI is a measure of the annual lighting energy (kWh/m^2 year) required to fulfil the illumination function and purpose in the building specifications. Both indicators, ELI and LENI, are practical tools based on specific standards that together help reach the goals of the third era of the lighting profession.

The Illuminating Engineering Society of North America (IES) is a leading authority on lighting technology. Since 1947 it has edited the lighting handbook, a recognised compendium of the state of the art in lighting. Its latest edition, the tenth (DiLaura et al., 2011), describes six types of factors to address in lighting design:

1. spatial
2. psychological
3. physiological
4. task
5. systemic
6. prescribed factors.

This categorisation shows the effort of IES since the 9th edition of its Handbook, to diminish the role of illuminance levels in the work plane as the main and only criterion in the design and evaluation of indoor lighting. In relation to the visual task factor, which heavily relies on photometric conditions, it includes the following criteria:

1. luminance (task L, background L, L limits, L contrasts and ratios, as well as L patterns and gradients)
2. chromatic contrast

3 veiling reflections

4 illuminance (horizontal, vertical, and E ratios)

5 colour reproduction.

The 10th edition proposes a new method for determining levels of illuminance, a section incorporated into the handbook since its 6th edition (1979). Given that it is the most consulted section of the handbook, this method has been modified, particularly in the last three editions. The proposed illuminance level selection method considers three factors:

1 the characteristics of the visual task (physical and photometric properties of the visual stimulus)

2 the relative importance of the task (in relation to its interaction with other tasks)

3 the characteristics of the observer (limited to the age of his visual system, defining three ranges: <25 years, 25–65 years, >65 years).

We highlight IESs approach to daylighting, which is considered a lighting source on equal footing with artificial lighting, rather than an optional, complementary or desirable component of the visual environment. The comprehensiveness of IES lighting design considerations positions the lighting handbook as a reference source material of the third stage of lighting.

When compared to international standards (ISO 8995, 2002; CEN, 2002) and the state of the art in lighting (IES handbook), the Argentinean legislation (DR 351/79; SRT resolution 84/12) and its protocol for legal verification (SRT-P) are outdated. As a readily available solution while our entire legal and normative framework is revised and updated, we developed the PC-SRT as an amend to the current SRT-P upon the premise that a lighting diagnostic tool should gather together all relevant information of the visual environment, the visual tasks performed, and the users beyond any specific legal constraint. This will lead to make evidence-based corrective actions towards good lighting quality in terms of what is expected within the third era of lighting.

2.2 PC-SRT

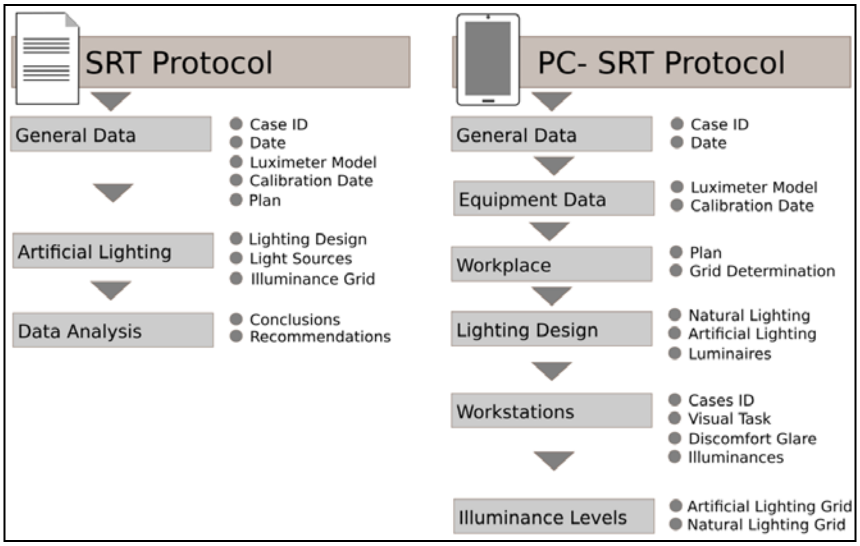
We developed different versions of the PC-SRT until reaching its current beta version for Android devices. The universal use of mobile communication technologies within developed and developing countries allowed the development of mHealth (Chib et al., 2015), an emerging field that uses mobile information and communication technologies as an economical and feasible solution in healthcare (Househ et al., 2012). As researchers are proposing more mHealth applications for many different health conditions (e.g., Georgsson and Staggers, 2015; Derks et al., 2017; Morey et al., 2017), Usability becomes a key factor in the adoption of those applications (Zapata et al., 2015; Derks et al., 2017). Also, in recent years a growing number of specific mobile applications aroused for the assessment of environmental factors at the workplace (Spitzhirn et al., 2016; Castillo-Martinez et al., 2018; Green et al., 2018). We propose to frame this emerging field by adopting the encompassing term ‘m-ergonomics’. To the best of our knowledge, this neologism has never been used before in the literature.

The rationale behind PC-SRT is based on the 4S model (safety, satisfaction, suitability, stimulation) by Gligor (2004). This lighting and productivity model states that high quality lighting does not and cannot depend on photometric quantities alone. Instead, a good luminous environment is suited to the specific work situation, providing optimal conditions for visibility, visual acuity, visual performance, task performance and productivity. It also positively stimulates the occupants, improving their alertness, interaction, mood and social climate, and meets the occupants' needs in terms of visual comfort, aesthetic judgement, acceptability and lighting control systems. Additionally, it takes into account health and safety parameters such as eyestrain, circadian rhythms, seasonal affective disorders and ageing. We structured our protocol around this reconciling luminous environment – human factors framework. The PC-SRT has two sections:

1. pre-site survey, which in turn is divided in general data and equipment data sub-sections
2. site survey, consisting in three sub-sections: case study general data, lighting systems, workstations, and lighting conditions.

Overall there are 64 items that need to be assessed by the users (Figure 2).

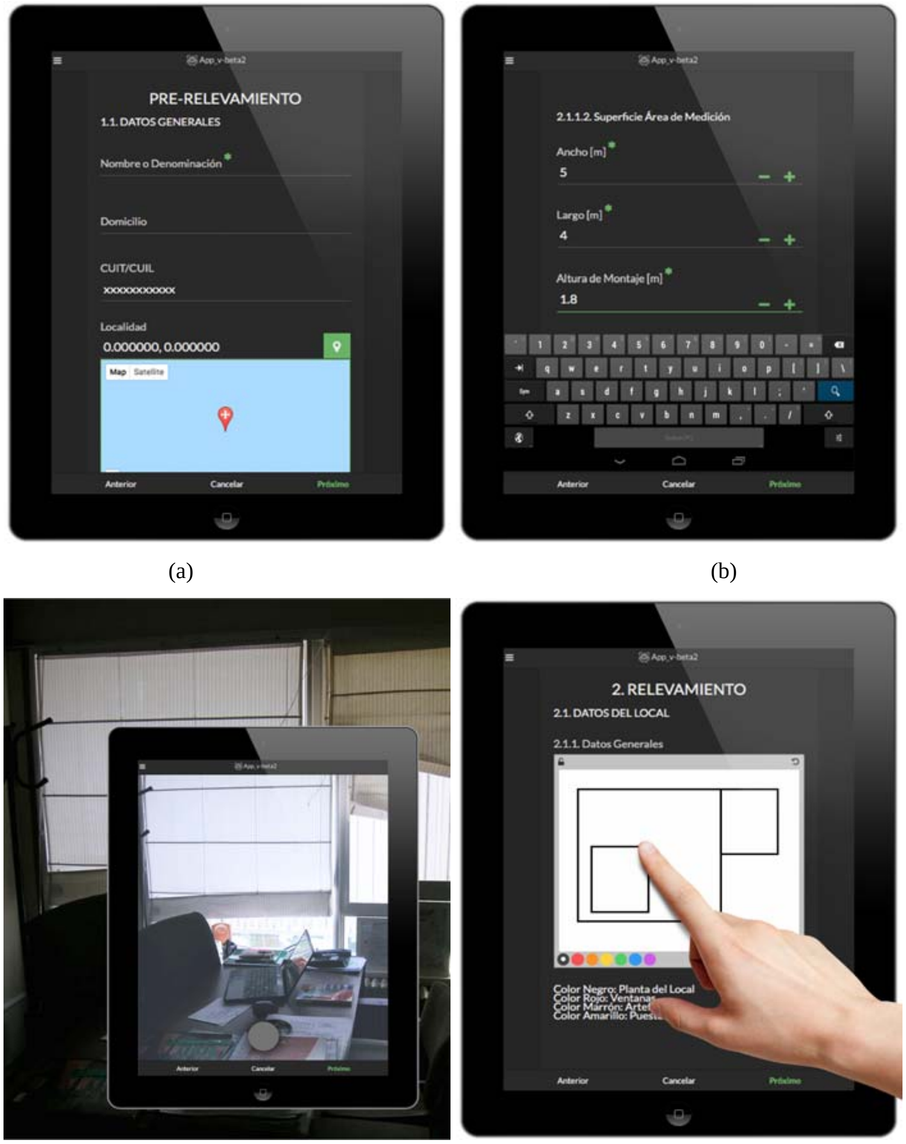
Figure 2 Structure of the SRT-P and PC-SRT protocols (see online version for colours)



2.2.1 *Features and user interface*

The PC-SRT interface is based on the post-WIMP natural interaction paradigm (Steinberg, 2012) characterised by hand gestural interaction. Navigation buttons at the bottom of the screen allow the access to the different sections of the protocol. To entry alphanumeric data, a keyboard pops up when necessary. The integration of functions of mobile devices allowed us to include in the PC-SRT several operations that previously required other equipment, such as the photographic record by means of the built-in camera of the mobile device. Another new functionality is the inclusion of a drawing module to sketch the plan of the room (Figure 3).

Figure 3 PC-SRT in use, (a) app installed in a tablet (b) alphanumeric data entry (c) photographic record with the built-in camera (d) drawing module (see online version for colours)



App_v-beta2

PRE-RELEVAMIENTO

1.1. DATOS GENERALES

Nombre o Denominación

Domicilio

CUIT/CUIL
xxxxxxxxxxx

Localidad
0.000000, 0.000000

(a)

App_v-beta2

2.1.1.2. Superficie Área de Medición

Ancho [m]
5

Largo [m]
4

Altura de Montaje [m]
1.8

(b)

(c)

App_v-beta2

2. RELEVAMIENTO

2.1. DATOS DEL LOCAL

2.1.1. Datos Generales

Color Negro: Planta del Local
Color Rojo: Ventanas
Color Marrón: Arte
Color Amarillo: Puest

(d)

2.2.2 PC-SRT lighting assessment metrics

The PC-SRT includes daylighting, illuminance, and glare metrics simplified to be user-friendly to practitioners.

2.2.2.1 Daylighting

According to the CIE Technical Committee on daylight metrics, comfort performance goals must be established before energy use can be optimised (Heschong Mahone Group, 2011), hence the PC-SRT contains human-centric metrics over energy performance ones: daylight factor (DF) (Moon and Spencer, 1942; Mansfield, 2018), view factor (VF) (Gowri, 2004), and lighting areas (Velds and Christoffersen, 2001). The PC-SRT user documentation provides recommendations related to the proper moment and frequency of the monitoring according to the criticality of the required daylight analysis. In practice, the user would find workplaces with more or less contribution of daylight to illuminate, even work environments with no daylight contribution at all. We provide criteria to define the importance of daylighting, depending on the characteristics of use, the location and dimensions of windows, as well as the morphology of the room and its location. With these variables it will be possible to a-priori define when and how to perform the daylighting analysis. These recommendations are based on the International Energy Agency daylight monitoring protocol (Velds and Christoffersen, 2001).

2.2.2.1.1 Daylight factor

This metric expresses the relation between the illuminance in an interior point (Ei) and horizontal illuminance on an unobstructed outer surface (Ee). It is a measure the light from the sky efficiency to provide indoors horizontal illuminance. The DF has practical advantages: it is intuitive and easy to communicate and to compare with specific standards, also it requires relatively affordable equipment and is widely used, being the most widely used natural lighting metric worldwide and it is included in the Argentinean standard IRAM-AADL j20-02 (1969). The DF, however, has some limitations to consider when interpreting its results (Monteoliva and Pattini, 2013; Reinhart et al., 2006). The DF is a static metric that overlooks the temporal and spatial dynamism of daylight in terms of amount and spectrum. It was developed for cloudy sky conditions, so this metric does not consider the particularities of specific luminous climates.

To overcome these limitations, the dynamic daylight paradigm (Nabil and Mardaljevic, 2005; Reinhart et al., 2006) uses Climate-based daylight modelling (CBDM) to predict luminous quantities using sun and sky conditions derived from standard meteorological datasets. This computer simulation-based paradigm is currently well consolidated at the Academia, to accurately predict daylight behaviour in interior spaces (Monteoliva et al., 2017). This search of accuracy led to the development new metrics and indicators, mainly dynamic values of horizontal illuminance (e.g., daylight autonomy – DA, useful daylight illuminance – UDI, continuous daylight autonomy – DAc), as well as the research of 'new workplanes' such as vertical, semi-cylindrical or cylindrical (dynamic discomfort glare assessment, circadian entrainment, objects modelling) probabilistically defined by the whole working activity, not just by the visual tasks performed, which is limited to the foveal visual field.

The shift from the static to the dynamic paradigm in daylighting assessment is also an opportunity to develop daylight metrics to replace the DF in normative and regulations worldwide. Unfortunately it is not an easy endeavour. Boubekri (2004), states that daylighting legislation is beset by many problems, some germane to the general field of lighting and others more specific to the nature of daylight as a source of illumination. Also, the fact that dynamic daylighting analysis heavily relies on computational simulations constitutes a barrier for the development of a practical replacement of the DF. Considering this, we decided to include the DF in spite of its known limitations.

2.2.2.1.2 View factor

Perhaps the most interesting and most controversial characteristic of windows is the importance of view. The most widely acknowledged positive contribution of a window view involves contributions to eye health. Frequent changes in eye focus distance give eye muscles a chance to relax momentarily, but it requires attractive stimulus to favour it (Anshel, 2007; Gowrisankaran and Sheedy, 2015). Views of nature potentially improve people's health and well-being, reduce stress and improve attention (Beute and Kort, 2014). However, the existence of an observable landscape from a window does not necessarily mean that it would effectively be in the visual field of the workers. Therefore, to assess the view outside the PC-SRT includes the following criteria from LEED V3 IEQ 8.2 credit 'daylight and views' (Konstantzos et al., 2015):

1. the presence of flora, fauna or sky, or movement or objects at a distance of at least 7.5 m from the observer
2. a VF of at least 3 (California Energy Commission, 2003).

The VF defines the quantity and quality of the view outside a window on a scale of 0 to 5, where 0 is the absence of view outside. Although this methodology differentiates between the primary view and the break view, the PC-SRT only requires the latter. The primary view encompass the quantity and quality in a 90° cone from the usual working posture while the break view includes the whole 183° human visual field (i.e., the view that the workers would get when they wished to take a break from regular work for a few moments).

2.2.2.1.3 Lighting areas

This characterisation defines specific lighting zones in relation to the prevailing lighting source: the daylight area (DLA), the mixed area (MLA), and the artificial light area (ALA). This subdivision of a room is based on the effective window height and the effective window area, calculated from the surface of the window above 0.9 m from the ground, the width of the wall and the visible transmittance of the glazed area. The DLA in a room is the area with a high daylight level. This area starts at the façade and has a depth of approximately two times the effective window height. In general, this area will be sufficiently illuminated by daylight to perform a normal task. The MLA will need some supplementary artificial lighting to accomplish a satisfactory light level throughout the day. It starts at the inner border of the DLA and has a depth of approximately 1.5 times the effective window height. The remaining part of the room is the ALD, which will be illuminated by artificial lighting alone, being the daylight contribution too low.

2.2.2.2 Illuminance

Lighting must enable users to perform the ‘work’ they came to do. Task performance and visual performance are not synonymous; in fact, several non-visual factors contribute significantly to task performance. Training, motor skills, motivation, and many other human factors interact with visibility to affect the level of task performance. Illuminance selection, which will be discussed below, is largely based on visual performance, not on task performance. The current trend is to develop illuminance selection procedures rather than seeking for absolute target values (DiLaura et al., 2011). The SRT-P approaches the illuminance evaluation problem at two levels: general (i.e., horizontal illuminance grid) and specific (i.e., several illuminance values at the workstation). Our procedure requires general illuminance measurements in the usual working lighting conditions, and in another ‘artificial lighting off’ scenario to assess the contribution of daylighting on mean lighting levels and uniformity. In the next level of specificity, illuminance measures are taken at workstations. PC-SRT allows two workplane illuminance measures per workstation. There are four steps to define the target illuminance:

1 define the visual task

2 select an illuminance category

3 determine an illuminance range

4 establish the final illuminance target value considering the age of the observer, the room reflectance’s, and the task speed/accuracy requirements.

We also included vertical illuminance at the eye a proxy of discomfort glare.

2.2.2.3 Discomfort glare

Instead of complicated formula, we propose the assessment of discomfort glare by means of semantic differential scaling. This method has been widely used in discomfort glare studies since its early days of discomfort glare research (Luckiesh and Holladay, 1925; Hopkinson, 1950), and it is the gold standard used in the development of current glare indexes (Fotios, 2015; Carlucci et al., 2015). PC-SRT includes glare sensation vote (GSV), replicating the glare categories used by Wienold and Christoffersen (2006) translated to Spanish. We instructed our participants to associate the magnitude of glare with the approximate period of time they could stand their sensation of discomfort using a four-point scale with pre-defined glare criteria: unnoticeable glare (in Spanish ‘deslumbramiento imperceptible’), noticeable glare (in Spanish ‘deslumbramiento notable’), disturbing glare (in Spanish ‘deslumbramiento molesto’), and intolerable glare (in Spanish ‘deslumbramiento intolerable’). Ratings above noticeable glare are beyond the borderline between comfort and discomfort (BCD). Luckiesh and Guth (1949) established the BCD as the single sensation which could be interpreted by the participants as a relatively definite sensation and that would be meaningful from a practical viewpoint (Rodriguez et al., 2017).

3 Part B: usability evaluation

3.1 Objectives and scope

The objective of this work is to perform a comparative usability analysis between the pen-and-paper SRT-P and the PC-SRT for Android mobile devices.

3.2 Materials and methods

We measured the general level of perceived usability by means of the system usability scale (SUS) (Brooke, 1996). This 'quick and dirty' instrument consists in ten Likert items, it is free and easy to administrate. This versatile instrument has allowed to measure usability in a wide variety of products and services (Kortum and Bangor, 2013). The SUS scale has a known factor structure (Lewis and Sauro, 2009) and good psychometric properties, reaching a Cronbach alpha of 0.91 (Bangor et al., 2008). It is scored from 0 to 100, with scores above 68 points desirable (Bangor et al., 2009). This instrument was developed and validated in English language, so in this research we implemented the Spanish version of SUS validated by Aguilar and Villegas (2016). The measurement of usability in mobile contexts poses new challenges due to factors external to the product itself (Kortum and Sorber, 2015). Usability is considered a key factor in achieving customer loyalty (Fetaji et al., 2011). Aspects related to the quality of the internet connection, screen size and resolution, and processing speed of the hardware (central processing units and available RAM memory) may affect the perceived quality of use (Ahmad et al., 2018). In practice, the combination of each of these factors generates hundreds of possible scenarios that are beyond the evaluator's control. Still, a controlled laboratory study in a single device lacks of ecological validity in relation to field measurement, where usability can be assessed in a variety of mobile devices in actual contexts of use. Following the decision tree proposed by Ji et al. (2006), we decided to collect usability data during post-occupational evaluations (POEs) using the participants' own mobile devices. We included a technical context of use survey (TECU) for our PC-SRT participants in order to analyse the possible effects of the aforementioned external factors in SUS scores. The TECU survey makes a qualitative description of three external factors by means of five-point semantic differential scales:

1 quality of the internet connection (1 very bad–5 very good)

2 screen size and resolution (1 very small–5 very large)

3 the processing speed of the hardware (1 very bad–5 very good).

Additional Likert-scale questions allowed us to gather data about the perceived degree of adequacy of each of those external factors for the specific context of use of our study (e.g., "the quality of the internet connection is adequate to use this app": 1 strongly disagree–5 strongly agree).

We carried out our study in two consecutive sessions during May 2017 as partial credits of applied ergonomics in design with our students at University of Mendoza (Mendoza, Argentina). In the first session, we presented our study and introduced the main features of the SRT-P and PC-SRT protocols to the students along with supporting material in PDF format. We reviewed the main features of both protocols and we instructed our participants in the installation process of the PC-SRT. On the second session, we briefed some methodological aspects of lighting measurement in a POE context and then we randomly divided our participants into two groups: one of them would use the SRT-P and the other one would use the PC-SRT. Finally, our participants teamed up in two or three persons and started the POE of lighting in different places of the university (classrooms, workshops, and offices) with the help of a LMT 200 luxmeter.

3.3 Results and discussion

26 volunteers (mean age = 23.4 years; SD = 5.1) participated in this study, 17 of them were females while 9 were males. From the literature review, effects of this uneven gender distribution on the SUS scale scores are not expected (Kortum and Sorber, 2015). We randomly divided our sample into two groups with the same number of participants. The first group carried out the POE following the pen-and-paper SRT-P (n = 13 with a distribution of 54.5% women and 45.5% men) while the second group performed the POE with the PC-SRT (n = 13 with a distribution of 76.9% of women and 23.1% of men). Regarding the devices used in the second group, 69.2% were Samsung, 23.1% were Motorola, and 7.7% were Huawei.

Figure 4 TECU survey results (see online version for colours)

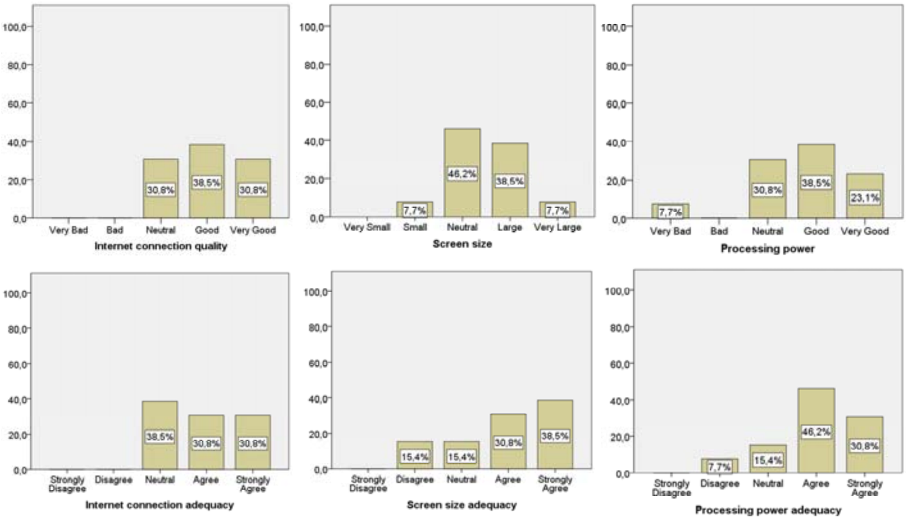


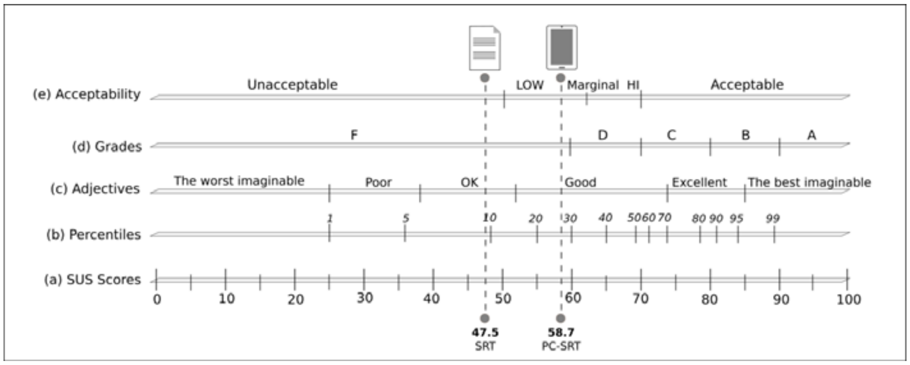
Figure 4 shows the results of the TECU survey for the PC-SRT group. Negative evaluations in these aspects would negatively influence the user experience; therefore, we sought to identify possible usability problems unrelated to the design of the application itself. In relation to the quality of the internet connection, it was mostly rated as ‘good’ (38.5%), with no negative assessments (mean = 4.0; SD = 0.82, median = 4 ‘good’). In relation to its adequacy (mean = 3.92; SD = 0.86; median = 4; ‘good’), it also received only neutral (38.5%) and positive ratings (61.5%). Regarding the processing capacity (mean = 3.69; SD = 1.11; median = 4 ‘good’), 7.7% of our participants rated it as ‘very bad’, 30.8% of them gave a neutral score and the remaining 61.5% rated it positively. Most of the participants (77%) considered that the processing capacity of their device for the use of PC-SRT was adequate (mean = 4 ‘agree’; SD = 0.91; median = 4 ‘agree’) while only 7.7% disagreed.

Finally, 7.7% of our volunteers considered their screen size as ‘small’ or ‘too large’, 38.5% as ‘large’, and 46.2% chose a neutral rating (mean = 3.46, SD = 0.78, median = 3 ‘neutral’). Most of the participants (69.3%) considered that the screen size of their devices was adequate (mean = 3.92; SD = 1.12; median = 4 ´agree’) while only 15.4% considered it as inadequate. In summary, it can be said that from the perspective of the participants, our participants used the PC-SRT within conditions of adequate connectivity, on screen sizes of appropriate size for the PC-SRT interface, with the processing demands of the application adequately satisfied by the hardware of the devices used.

3.3.1 Global usability scores

Figure 5 graphically summarises SUS results. The mean usability score of the SRT-P was 47.5 (SD = 14.5). This score corresponds to the 9th percentile according to the standardisation of Sauro (2011). Bangor et al. (2009) proposed a qualitative scale associated to the SUS score. According to this scale, the SRT-P score was in an unacceptable range (SUS < 50) that is associated to the adjective ‘OK’. This adjective should be interpreted as a mean score that at the same time indicates a mean level of usability that in some way is not acceptable and must be improved (Bangor et al., 2009). The mean usability score of the PC-SRT protocol was 58.7 (SD = 14.9), which is the 29th percentile (Sauro, 2011). This score is in the lower marginally acceptable range of usability (50 < SUS < 70), and corresponds to the adjective ‘good’. Mann-Whitney test results showed that the difference in SUS scores between the SRT protocol and the PC-SRT were not statistically significant (U = 51.5; p = 0.091).

Figure 5 (a) SRT-P and PC-SRT SUS scores (b) SUS Percentiles (Sauro, 2011) (c)–(e) SUS score interpretation (Bangor et al., 2009)



According to the psychometric analysis by Lewis and Sauro (2009), the SUS scale measures two factors: usability (US) and learnability (LE). LE is measured by items 4 and 10, while the rest of the items determine the US factor. Table 1 shows the absolute score reached for each factor and their relative performance as percentages of its maximum possible scores.

Table 1 SUS factors results

	LE	*%*	*US*	*%*	*SUS*
SRT-P	5.8	28.8	41.7	52.2	47.5
PC-SRT	10.2	51.0	48.5	60.6	58.7
Difference	4.4	22.2	6.8	8.4	11.2

For the SRT-P, the LE factor contributes an average of 5.8 points (28.8%) out of a possible total of 20 points, while for the PC-SRT this factor averages 10.2 points (51.0%). Mann-Whitney test results show that the difference in LE scores between SRT and PC-SRT are not statistically significant (U = 47.5; p = 0.055). Despite this lack of statistical significance, we highlight that despite that PC-SRT protocol was designed to gather more data in a wider variety of aspects of the visual environment (64 items in 15 factors, divided in six sections), users tended to perceive it as easier to learn in relation to the SRT-P (42 items in ten factors, divided in three sections). For the US factor, for a maximum of 80 possible points, the SRT protocol averaged 41.7 points (52.2%), while the PC-SRT showed an average of 48.5 points (60.6%). Mann-Whitney test results show that those differences in US scores were not statistically significant (U = 57; p = 0.157). The descriptive statistical analysis shows a better performance of the PC-SRT, however the score differences in global usability (11.2%), as well as in the LE and US factors (22.2% and 8.4% respectively) were not enough to reach statistical significance.

Kortum and Sorber (2015) evaluated the usability of the 10 most downloaded applications in Android mobile devices. The average usability score (n = 778) obtained by those applications was 82.7 (SD = 15.7). Also, Bangor et al. (2009) reported a mean SUS score of 65.9 (n = 593) for user interfaces of mobile phones. In relation to these benchmarks, we expected a low learnability and lower usability scores for the PC-SRT, because our application is a technical protocol designed for the legal verification of lighting in working environments that requires specific knowledge, and previous training in POE. Although we designed a user manual that was previously handed to our participants, this documentation appears to be insufficient given the methodological rigor of the PC-SRT and its SUS scores obtained.

3.3.2 Item-by-item analysis

Despite Brooke (1996, p.194) himself cautioned that 'scores for individual items are not meaningful on their own', Bangor et al. (2008, p.579) stated that 'clients and practitioners alike tend to ignore this admonition' and look into individual statements. In this research we performed a SUS item-by-item analysis in order to gain further information about the differences between SRT-P and PC-SRT, so we compared SUS performance of both protocols through their adjusted item scores (i.e., after completing the item scoring procedure so that positive responses in even items are associated with a larger number, like in the odd items). Table 2 compares the resulting descriptive statistics of SRT-P and PC-SRT protocols for every SUS item. When comparing their medians, PC-SRT outscores SRT-P in 6 out of 10 items and they tie in the remaining four. However, Mann-Whitney test for two independent samples showed statistically significant differences between SRT-P and PC-SRT in SUS-Q2: "I found the system unnecessarily complex" (U = 37.5; p = 0.013), SUS-Q4: "I think that I would need the support of a technical person to be able to use this system" (U = 34.5; p = 0.008), and SUS-Q7: "I would imagine that most people would learn to use this system very quickly" (U = 36; p = 0.011). Lewis and Sauro (1998) were surprised that SUS-Q7 did not align with the LE factor, and it is still surprising nowadays in the context of our current comparative study.

Table 2 Descriptive statistics of SRT-P and PC-SRT protocols for every SUS item

	SRT-P					*PC-SRT*				
	Median	*Mean*	*SD*	*STD mean*	*STD SD*	*Median*	*Mean*	*SD*	*STD mean*	*STD SD*
SUS-Q1	2	2.31	0.63	0.48	0.57	2	1.92	1.19	–0.42	0.89
SUS-Q2	1	1.31	1.32	–0.52	0.80	3	2.62	1.04	0.27	0.85
SUS-Q3	3	2.38	0.96	0.55	1.11	3	2.46	1.13	0.12	0.79
SUS-Q4	0	0.76	1.17	–1.06	0.85	2	2.15	1.28	–0.19	1.28
SUS-Q5	2	2.69	0.85	0.86	0.92	3	2.62	0.65	0.27	0.86
SUS-Q6	2	2.23	1.17	0.4	1.16	3	2.54	0.66	0.19	0.78
SUS-Q7	1	1.31	1.44	–0.52	0.98	3	2.77	0.93	0.42	0.79
SUS-Q8	1	1.62	0.96	–0.21	0.69	2	2.23	1.36	–0.12	1.06
SUS-Q9	2	2.15	1.34	0.32	0.87	2	2.23	1.09	–0.12	0.82
SUS-Q10	2	1.54	1.13	–0.29	1.08	2	1.92	1.49	–0.42	1.07

Following Bangor et al. (2008) analysis, we standardised SUS item scores (Table 2). Mann-Whitney test results of the standardised item scores showed statistically significant differences between SRT-P and PC-SRT in SUS-Q1: "I think that I would like to use this product frequently" (U = 36; p = 0.011) and confirmed the statistical differences in SUS-Q2 (U = 35; p = 0.01), SUS-Q4 (U = 46.5; p = 0.05), and SUS-Q7 (U = 39; p = 0.019). Score standardisation shows how much each item score differs from the average item response. Wilcoxon rank test results show for SRT-P that relative to all responses a participant gave on their survey, SUS-Q1 (Z = –2.351; p = 0.019) and SUS-Q5: "I found the various functions in the product were well integrated" (Z = –2.769; p = 0.006) tended to have a more positive rating relative to their other ratings, while SUS-Q2 (Z = –1.960; p = 0.05) and SUS-Q4 (Z = –2.559; p = 0.011) tended to have a more negative ranking. PC-SRT showed no statistically significant tendency in any of its SUS-items. And this contrast is the main reason why item-by-item analysis might be valuable after all. We were warned that SUS item scores are not meaningful on their own, however when the behaviour of each standardised item score is compared among products; it highlights specific positive or negative aspects of the user experience that should be consciously interpreted by the design team. There are four contextual factors that should be considered within mobile usability: the user, the environment, technology and task/activity (Baharuddin et al., 2013). The TECU survey deals with a series of technical characteristics that may pose constraints to usability. In order to gain a deeper understanding of the role of mobile technology in SUS item scores, we correlated them with TECU survey results by means of Spearman correlation coefficient. We found a moderate and positive statistically significant correlation (rs = 0.627, p = 0.022) between SUS-Q5: "I found the various functions in the product were well integrated" and TEQU-Q2: "Rate the internet connection (speed, stability, reliability)." We also found a strong and positive statistically significant correlation (rs = 0.860, p = <0.0001) between

SUS-Q5 and TECU-Q6: "The screen size is adequate for visualisation and interaction with the elements of the interface." In the context of this study, a better internet connection and a more adequate screen size positively relates with the functionality of the software as an integrated whole. We also found a positive and moderate statistically significant correlation (rs = 0.577, p = 0.039) between SUS-Q7: "I would imagine that most people would learn to use this system very quickly" and TECU-Q4: "Rate the processing capacity of your device (multitasking, stability, speed)." A faster hardware helps the user to feel more confident when interacting with the software encouraging an active exploration of the system by novice users.

4 Conclusions

After a critical analysis of the Argentinean currently mandatory pen-and-paper lighting measurement protocol (SRT-P), we developed an alternative protocol (PC-SRT) for Android mobile devices. Designed at the academia for practitioners, PC-SRT allows a comprehensive assessment of the lighting environment in work spaces. This ubiquitous diagnostic tool includes the latest advances in human factors in lighting in a trade-off between reliability and simplicity. There is a trend in lighting codes and good practice advice against following slavishly recommendations which simply specify illuminance on a mythical horizontal plane. Laws, regulations, codes, guides and practices should never be a substitute for thought. Within this context, our goals are:

1. to ease the work of practitioners while measuring the visual environment in field
2. to take the most of available mobile technology to improve and apply new data gathering methodologies
3. promote, through objectives 1 and 2, work spaces with good lighting quality that promote efficiency, effectiveness, and satisfaction of its users.

Although it is specifically designed for our technical and economical context (i.e., a developing country), the PC-SRT is not the only and one development around the globe seeking those goals. While working on this manuscript, we realised that our development is part of a trend that seeks to support the practice of ergonomics by mobile technology. Furthermore we consider this trend as an emerging field in Human Factors and Ergonomics, which we termed M-ergonomics (or mergonomics, after Mobile Ergonomics).The objective of this research was to perform a comparative field usability study between SRT-P (n = 13) and PC-SRT (n = 13) by means of the SUS scale and the TECU survey. TECU survey results show that our participants used the PC-SRT within conditions of adequate of connectivity, screen size, and hardware processing capabilities. The descriptive statistical analysis showed a better performance of the PC-SRT, however the score differences found in global usability (11.2%), as well as in the LE and US factors (22.2% and 8.4% respectively) were not enough to reach statistical significance. However, Mann-Whitney test results of the standardised item scores showed statistically significant differences between SRT-P and PC-SRT in SUS-Q1, SUS-Q2, SUS-Q4, and SUS-Q7, mainly due to a different statistically significant single item behaviour between SRT-P (item positive and negative trends in relation to the mean item score) and PC-SRT (no item trends found in relation to the mean item score). Perhaps every SUS item is not meaningful on its own, however a comparative item-by-item analysis (i.e., standardised item behaviour within and between products) proved to reveal valuable insights that should be consciously interpreted by the design team. The results obtained in this instance are encouraging, because even though PC-SRT gathers more data in a wider variety of aspects of the visual environment than SRT-P, users tended to perceive it as easier to use. The fact that we recruited a young sample (mean age = 23.4 years; SD = 5.1) can shorten the scope of our conclusions. Further studies with a sample of older workers are needed, because the learning curve in the use of mobile devices might be different and the perception of usability is age related. Comparison of results could lead to improvements in the application, which could enable a wider range of users to be successfully reached. The results of this comparative study, along with an upcoming study with actual practitioners instead of students are a direct input into the development of first stable version of PC-SRT (v1.0), which is currently in progress.

Acknowledgements

This research was funded by National Council of Scientific and Technical Research (Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET) and National Agency of Scientific and Technological Research (Agencia Nacional de Promoción Científica y Tecnológica – ANPCyT) under contracts PICT 2013-2089 and PIP 2009 – A Pattini – CONICET and by the 2013–2015 Scientific Research Program of University of Mendoza.

References

Aguilar, M.H. and Villegas, A.G. (2016) 'Análisis comparativo de la Escala de Usabilidad del Sistema (EUS) en dos versiones', *Revista Iberoamericana de las Ciencias Computacionales e Informática*, Vol. 5, No. 10, pp.44–58.

Ahmad, N., Rextin, A. and Kulsoom, U.E. (2018) 'Perspectives on usability guidelines for smartphone applications: an empirical investigation and systematic literature review', *Information and Software Technology*, September, Vol. 94, pp.130–149.

Anshel, J.R. (2007) 'Visual ergonomics in the workplace', *AAOHN Journal*, Vol. 55, No. 10, pp.414–420.

Baharuddin, R., Singh, D. and Razali, R. (2013) 'Usability dimensions for mobile applications – a review', *Research Journal of Applied Sciences, Engineering and Technology*, Vol. 5, No. 1, pp.2225–2231.

Bangor, A., Kortum, P. and Miller, J.T. (2009) 'Determining what individual SUS scores mean: adding an adjective rating scale', *Journal of Usability Studies*, Vol. 4, No. 3, pp.114–123.

Bangor, A., Kortum, P.T. and Miller, J.T. (2008) 'An empirical evaluation of the system usability scale', *International Journal of Human-Computer Interaction*, Vol. 24, No. 6, pp.574–594.

Beute, F. and Kort, Y.A. (2014) 'Salutogenic effects of the environment: review of health protective effects of nature and daylight', *Applied Psychology: Health and Well-Being*, Vol. 6, No. 1, pp.67–95.

Boubekri, M. (2004) 'An overview of the current state of daylight legislation', *Journal of the Human-Environment System*, Vol. 7, No. 2, pp.57–63.

Boyce, P. (1996) 'Illuminance selection based on visual performance – and other fairy stories', *Journal of the Illuminating Engineering Society*, Vol. 25, No. 2, pp.41–49.

Brooke, J. (1996) 'SUS: a Quick and dirty usability scale', in Jordan, P.W., Thomas, B., Weerdmeester, B.A. and McClelland, I.L. (Eds.): *Usability Evaluation in Industry*, Taylor & Francis, London.

California Energy Commission (2003) *Windows and Offices: a Study of Office Worker Performance and the Indoor Environment*, Technical Report, p.159.

Carlucci, S., Causone, F., De Rosa, F. and Pagliano, L. (2015) 'A review of indices for assessing visual comfort with a view to their use in optimization processes to support building integrated design', *Renewable and Sustainable Energy Reviews*, Vol. 47, No. 1, pp.1016–1033.

Castillo-Martinez, A., Medina-Merodio, J.A., Gutierrez-Martinez, J.M., Aguado-Delgado, J., de-Pablos-Heredero, C. and Otón, S. (2018) 'Evaluation and improvement of lighting efficiency in working spaces', *Sustainability*, Vol. 10, No. 4, p.1110.

Chib, A., van Velthoven, M.H. and Car, J. (2015) 'mHealth adoption in low-resource environments: a review of the use of mobile healthcare in developing countries', *Journal of Health Communication*, Vol. 20, No. 1, pp.4–34.

Cuttle, C. (2010) 'Towards the third stage of the lighting profession', *Lighting Research and Technology*, Vol. 42, No. 1, pp.73–93.

Dehoff, P. (2007) 'ELI – ergonomic lighting indicator, lighting criteria at a glance', in *Proceedings of the 26th CIE Session*, Beijing, China.

Dehoff, P. (2012) 'Lighting quality and energy efficiency is not a contradiction', *Light and Engineering*, Vol. 20, No. 3, pp.34–39.

Derks, Y.P., Visser, T.D., Bohlmeijer, E.T. and Noordzij, M.L. (2017) 'mHealth in mental health: how to efficiently and scientifically create an ambulatory biofeedback e-coaching app for patients with borderline personality disorder', *International Journal of Human Factors and Ergonomics*, Vol. 5, No. 1, pp.61–92.

DiLaura, D.L., Houser, K.W., Mistrick, R.G. and Steffy, G.R. (2011) '*The Lighting Handbook: Reference and Application*, p.1328, Illuminating Engineering Society of North America, New York, NY.

Comité Européende Normalisation (CEN) (2002) *EN 12464-1:2002: Light and Lighting-Lighting of Work Places, Part 1: Indoor Work Places*, Brussels.

Comité Européende Normalisation (CEN) (2007) *EN 15193-1:2008: Energy Performance of Buildings – Energy Requirements for Lighting*, Brussels

Fetaji, B., Fetaji M. and Kaneko K. (2011) 'Comparative study of efficiency among the developed MLUAT methodology in comparison with qualitative user testing method and heuristics evaluation, information technology interfaces (ITI)', in *Proceedings of the ITI 2011 33rd International Conference*, pp.269–274.

Fotios, S. (2015) 'Research note: uncertainty in subjective evaluation of discomfort glare', *Lighting Research and Technology*, Vol. 47, No. 3, pp.379–383.

Gašová, M., Gašo, M. and Štefánik, A. (2017) 'Advanced industrial tools of ergonomics based on Industry 4.0 concept', *Procedia Engineering*, Vol. 192, pp.219–224.

Georgsson, M. and Staggers, N. (2015) 'Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics', *Journal of the American Medical Informatics Association*, Vol. 23, No. 1, pp.5–11.

Gligor, V. (2004) *Luminous Environment and Productivity at Workplaces*, Thesis (Licentiate), Helsinki University of Technology, Espoo.

Gowrisankaran, S. and Sheedy, J.E. (2015) 'Computer vision syndrome: a review', *Work*, Vol. 52, No. 2, pp.303–314.

Gowri, K. (2004) 'Green building rating systems: an overview', *ASHRAE Journal*, Vol. 46, No. 11, pp.56–60.

Green, P., Cheydleur, G., Li, K., Deng, R. and Sanchez, R. (2018) 'Free smartphone apps to support human factors/ergonomics work', *Ergonomics in Design*, Vol. 26, No. 1, pp.4–16.

Heschong Mahone Group (2011) *Daylight Metrics*, California Energy Commission, Publication Number: CEC-500-2012-053.

Hopkinson, R.G. (1950) 'The multiple criterion technique of subjective appraisal', *Quarterly Journal of Experimental Psychology*, Vol. 2, No. 3, pp.124–131.

Househ, M., Borycki, E.M., Kushniruk, A.W. and Alofaysan, S. (2012) 'mHealth: a passing fad or here to stay?', in *Telemedicine and E-Health Services, Policies, and Applications: Advancements and Developments*, IGI Global, pp.151–178.

Instituto Argentino de Normalizacion y Certificacion (IRAM-AADL j20-02) (1969) *Iluminación natural en edificios: condiciones generales y requisitos especiales*, IRAM, Buenos Aires

International Organization for Standardization (ISO 8995-1) (2002) *Lighting of Indoor Work Places – Part 1: Indoors*, ISO, Geneva.

International Organization for Standardization (ISO 9241-11) (1998) *Ergonomic Requirements for Office Work with Visual Display Terminals (VDT) – Part 11: Guidance on Usability*, ISO, Geneva.

Ji, Y.G., Park, J.O., Lee, C. and Yun, M.H. (2006) 'A usability checklist for the usability evaluation of mobile phone user interface', *International Journal of Human-Computer Interaction*, Vol. 20, No. 3, pp.207–231.

Konstantzos, I., Chan, Y.C., Seibold, J. C., Tzempelikos, A., Proctor, R.W. and Protzman, J.B. (2015) 'View Clarity Index: a new metric to evaluate clarity of view through window shades', *Building and Environment*, Vol. 90, No. 1, pp.206–214.

Kortum, P. and Bangor, A. (2013) 'Usability ratings for everyday products measured with the system usability scale (SUS)', *International Journal of Human Computer Interaction*, Vol. 29, No. 2, pp.67–76.

Kortum, P. and Sorber, M. (2015) 'Measuring the usability of mobile applications for phones and tablets', *International Journal of Human-Computer Interaction*, Vol. 31, No. 8, pp.518–529.

Lewis, J.R. and Sauro, J. (2009) 'The factor structure of the system usability scale', in *Human-Centered Design*, pp.94–103, Springer, Berlin, Germany.

Luckiesh, M. and Guth, S.K. (1949) 'Brightness in visual field at borderline between comfort and discomfort (BCD)', *Illuminating Engineering*, Vol. 44, No. 11, pp.650–670.

Luckiesh, M. and Holladay, L.L. (1925) 'Glare and visibility', *Northwest Journal of Optometry*, Vol. 2, No. 6, p.175.

Mansfield, K.P. (2018) 'Architectural lighting design: a research review over 50 years', *Lighting Research and Technology*, Vol. 50, No. 1, pp.80–97.

Mills, E. and Borg, N. (1999) 'Trends in recommended illuminance levels: an international comparison', *Journal of the Illuminating Engineering Society*, Vol. 28, No. 1, pp.155–163.

Monteoliva, J.M. and Pattini, A. (2013) 'Iluminación natural en aulas. Análisis predictivo dinámico del rendimiento lumínico-energético en clima soleados', *Ambiente Construido*, Vol. 13, No. 4, pp.235–248.

Monteoliva, J.M., Villalba, A. and Pattini, A.E. (2017) 'Variability in dynamic daylight simulation in clear sky conditions according to selected weather file: satellite data and land-based station data', *Lighting Research and Technology*, Vol. 49, No. 4, pp.508–520.

Moon, P. and Spencer, D.E. (1942) 'Illumination form a non-uniform sky', *Illuminating Engineering*, Vol. 37, No. 1, pp.707–726.

Morey, S.A., Barg-Walkow, L.H. and Rogers, W.A. (2017) 'Managing heart failure on the go: usability issues with mHealth apps for older adults', in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications, Los Angeles, CA, September, Vol. 61, No. 1, pp.1–5.

Nabil, A. and Mardaljevic, J. (2005) 'Useful daylight illuminance: a new paradigm to access daylight in buildings', *Lighting Research and Technology*, Vol. 37, No. 1, pp.41–59.

Pattini, A., Rodriguez, R., Monteoliva, J.M. and Garretón, J.Y. (2012) 'Iluminación en espacios de trabajo. Propuestas al protocolo de medición del factor iluminación de la Superintendencia de

Riesgos de Trabajo', *Avances en Energías Renovables y Medio Ambiente*, Vol. 16, No. 1, pp.81–88.

Rea, M.S., Figueiro, M.G. and Bullough, J.D. (2002) 'Circadian photobiology: an emerging framework for lighting practice and research', *Lighting Research and Technology*, Vol. 34, No. 3, pp.177–187.

Reinhart, C., Mardaljevic, J. and Rogers, Z. (2006) 'Dynamic daylight performance metrics for sustainable building design', *Leukos*, Vol. 3, No. 1, pp.7–31.

Rodriguez, R.G., Yamín Garretón, J.A. and Pattini, A.E. (2017) 'An epidemiological approach to daylight discomfort glare', *Building and Environment*, Vol. 113, No. 1, pp.39–48.

Sauro, J. (2011) *Measuring Usability with the System Usability Scale (SUS)* [online] https://measuringu.com/sus/ (accessed 12 September 2018).

Spitzhirn, M., Merkel, T. and Bullinger, A.C. (2016) 'Measuring work environment factors by everyone using smartphones', in *Advances in Ergonomic Design of Systems, Products and Processes*, pp.281–299, Springer Vieweg, Berlin, Heidelberg.

Steinberg, G. (2012) 'Natural user interfaces', in *ACM SIGCHI Conference on Human Factors in Computing Systems*.

Van Bommel, W.J. and Van den Beld, G.J. (2004) 'Lighting for work: a review of visual and biological effects', *Lighting Research and Technology*, Vol. 36, No. 4, pp.255–266.

Veitch, J.A. and Newsham, G.R. (1998) 'Lighting quality and energy-efficiency effects on task performance, mood, health, satisfaction, and comfort', *Journal of the Illuminating Engineering Society*, Vol. 27, No. 1, pp.107–129.

Velds, M. and Christoffersen, J. (2001) 'Monitoring procedures for the assessment of daylighting performance of buildings', *IEA SHC TASK 21/ECBCS ANNEX 29*.

Wienold, J. and Christoffersen, J. (2006) 'Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras', *Energy and Buildings*, Vol. 38, No. 7, pp.743–757.

Zapata, B.C., Fernández-Alemán, J.L., Idri, A. and Toval, A. (2015) 'Empirical studies on usability of mHealth apps: a systematic literature review', *Journal of Medical Systems*, Vol. 39, No. 2, p.1.