HOW TO ASSESS ALERTING EFFECTS OF DAYLIGHT AT THE WORKPLACE? LEARNINGS FROM SEMI-CONTROLLED STUDIES

Soto Magán, V.E.[1], Andersen, M.[1]
[1] Laboratory of Integrated Performance in Design (LIPID), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, SWITZERLAND
victoria.sotomagan@epfl.ch

Abstract

Daylight, in which we evolved, is naturally rich in the blue component of the spectrum and has key properties when it comes to impacts on body functioning. Current lifestyles are driving the emergence of a 24-hour society that spends most of the time indoors (around 90%), where lighting conditions are a result of design and operation priorities derived from both comfort and energy criteria, which often lead to reduced access to daylight (smaller or shaded openings) to manage solar gains and glare risks. This may result in an insufficient (day)light exposure in daily life from a physiological perspective, light being an essential cue to properly entrain our internal circadian clock and increase subjective alertness. But it is still unclear whether it can have a significant beneficial effect when compared to artificial light from a psychophysiological standpoint. Most of the studies on acute alerting effects have been conducted in well-controlled laboratory settings, where somewhat extreme and narrowly defined lighting conditions have been tested. This paper proposes assessment and monitoring techniques that would apply to semi-controlled studies instead, and focuses on the impact of daylighting in work settings by exploring ways to investigate alertness neurobehaviour and physiology in realistic indoor conditions.

Keywords: Daylight, non-visual effects, alertness, sleepiness, arousal, sustained attention

1 Introduction

1.1 Background

Scientific evidence put forward the existence of various neurobehavioral and physiological processes, also known as non-visual responses. They are reactive to environmental stimuli through a novel type of photoreceptor in the retina, the intrinsic photosensitive retinal ganglion cells (ipRGCs) (Berson et al., 2002; Hattar et al., 2002), even in the absence of rods and cones (Provencio et al., 2000).

Light, and especially daylight, is important for many of these biological functions, but understanding its effects encompasses multiple dimensions. Some of the most established responses can be categorized in two main groups: indirect or circadian effects, which are not immediately obvious and for which extensive monitoring is needed in order to observe changes in the system, such as the entrainment of the internal biological clock to 24-hour light-dark cycles; and acute effects, which can be observed directly in healthy people and do not necessarily influence the circadian system, such as melatonin suppression, cortisol levels or alertness (Cajochen, 2007). But light exposure is important not only from a theoretical perspective; it may also lead to practical applications. It is known that impairment of alertness impacts quality of life in the form of daytime sleepiness, circadian misalignment or sleep disorders, but it also affects cognitive performance, perceptual skills or reasoning and decision-making capabilities, which are all likely to affect productivity (Lok et al., 2018).

Arousal, alertness, vigilance or attention are concepts that have been widely used as synonyms to express different states of cortical activation, but depending on the context and the field, the definition may change. Psychologists and cognitive neuroscientists usually refer to alertness to denote a state of vigilance or the ability to sustain attention to a certain task during a period of time, during which a cortical activation is implied. Neurophysiologists, however, refer to it as an arousal level on the sleep-wake spectrum without involving any cognitive or behavioural responsiveness (Oken et al., 2006). In the context of circadian research, alertness is commonly used to denote the opposite of sleepiness, which overlaps to some extent with arousal when used as a synonym of wakefulness (Cajochen, 2007).

1.2 Influential factors and light characteristics

On the one hand, it is important for studies on alertness to evaluate the conditions associated with alterations in the normal state of the system (i.e. stimuli). It is equally relevant to consider the underlying brain processes and psychological assembles behind this construct, which may be the partial cause of the effect. It is known that, to some extent, both sleep-wake and alerting systems can overlap partially, but still, there are substantial differences between them. Psychological factors can explain some of these. Motivation, stress or aspects related to the stimuli such as habituation, intensity, duration or background could be some of the altering psychological agents for alertness state. Working memory, cognitive control, decision-making response criteria or arithmetic ability, which are executive brain functions, may also interfere with alertness. Sleep disorders or neurobehavioral disorders, such as sleep apnoea, narcolepsy or attention-deficit hyperactivity, may also impact declines in vigilance. For a complete overview of brain systems and underlying mechanisms see Oken et al., 2006.

On the other hand, intensity and spectral distribution are among the main qualities of light that drive non-visual responses, together with duration or temporal characteristics, circadian timing of exposure and prior light history, but only the first two have been widely explored (Pachito et al., 2016). In line with this statement, it is known that the blue part of the visible spectrum (wavelengths peaking at 460-480 nm) is more effective activating certain brain regions (Phipps-Nelson et al., 2009) that might affect alertness and cognitive performance both indirectly, by modifying circadian rhythms, and directly, giving rise to acute effects (Cajochen, 2007). This last property has been less explored (Amundadottir, 2016), whereas its phase-shifting capacity is better established, especially with high intensity levels (Badia et al., 1991). It is therefore generally agreed that shorter wavelengths or higher correlated colour temperature (CCT) and brighter environments lead to stronger acute alerting effects during the evening or at night (Souman et al., 2018). Nevertheless, other publications have demonstrated that alertness and/or performance can be enhanced also during daytime when melatonin levels are almost unnoticeable. However, a closer look at the literature reveals that these daytime effects have not been explored thoroughly for the case of daylight (Amundadottir, 2016), which has a very dynamic nature compared to artificial light, and more important, is naturally rich in the blue component of the spectrum, assuming therefore that it may have the “right” properties when it comes to impacts on body functioning.

1.3 Paper objectives

In this paper we provide an overview of different methods used in daytime empirical studies on acute alerting effects of light, while assessing their feasibility for replication in semi-controlled or uncontrolled experimental conditions using daylight. We evaluate both subjective and objective indicators of alertness (i.e. alertness understood as the opposite of sleepiness for self-reports, and alertness as a synonym of physiological arousal/wakefulness or sustained attention/vigilance for objective markers). Only studies with significant results, as highlighted in two different literature reviews on acute alerting effects of light done by Souman et al., 2018 and Lok et al., 2018, have been considered. The overall goal of this paper is to propose a novel methodology to assess direct non-visual effects of daylight in this case, in semi-controlled realistic conditions outside the laboratory confinement.

2 Methods

2.1 Quantification of alertness

In the context of circadian research, alertness – understood as the opposite of sleepiness or as a synonym of physiological arousal/wakefulness or sustained attention/vigilance – can be assessed in two different ways: by using subjective self-evaluated instruments, which are inexpensive, simple and not time consuming, or with objective measures, which involve more rigorous, and typically (but not always) more intrusive and expensive techniques. The most popular ones in the current literature have been summarized in Table 1, and categorized based on feasibility of use, monitoring and administration for the different types of experimental design conditions (i.e. controlled, semi-controlled, uncontrolled).

Table 1 – Instruments for measuring sleepiness, arousal or wakefulness and sustained attention or vigilance, categorized by experimental design

	Method	Experimental design	Type of evaluation	Indicator name	Recording technique
Subjective measures	Non-clinical	• Controlled • Semi-controlled • Uncontrolled	Short term (point-in-time)	KSS	9 levels / likert scale
				SSS	7 levels / likert scale
			Short term – combined scales	GV+GA (vigour/affect)	8-item / VAS (0-100)
				FACES (alertness+fatigue)	50-item /checklist
			Long term (preceding weeks)	DSS+NSOS	8+2-item / 4-points scale
			Long term – combined scales	SWAI	59-item / 9-points scale
				THAT	10-item /6-points scale
			Global	ESS	8-item / 4-points scale
Objective measures	Non-clinical	• Controlled • Semi-controlled • Uncontrolled	Short term – point-in-time	PVT	Reaction time / auditory or visual task
				MCT	Reaction time / auditory or visual task
	Clinical	• Controlled	Short term/ long term – continuous	MSLT	EEG activity
				MWT	EEG activity
				Pupil size	Pupillography
				Blink rate	EOG activity
				Eye movement	EOG activity
		• Controlled • Semi-controlled	Short term/ long term – continuous	HR	ECG activity
				HRV	ECG activity

2.1.1 Subjective measures

A non-clinical but broadly used method to measure alertness levels on individuals is to resort to self-rated questionnaires. These are often easy and quick to conduct, which is a key feature when targeting groups with large sample sizes and/or either uncontrolled or semi-controlled experimental conditions. Three types of evaluations are possible in this case: short-term (also called point-in-time) measures, where participants are asked to report their current perceived level of alertness on a predefined anchored scale; long-term reports, where participants are asked to express perceived alertness during a given period of time (e.g. in the last two weeks); and overall or global measures, where participants have to assess their feelings of sleepiness in general, implying long-term experiences that could be influenced by personality or other characteristics.

For point-in-time assessments of alertness, there are two well known validated scales, the Karolinska Sleepiness Scale (KSS) (Åkerstedt and Gillberg, 1990) and the Stanford Sleepiness Scale (SSS) (Hoddes and Dement, n.d.). These are Likert-type scales where participants have to define their perceived level of alertness on a 1-item predefined anchored scale, with 9 or 7 options to choose from respectively. This same idea could be also replicated on a visual analog scale (VAS) by simply asking the participants to report their

feelings on a 10-cm line that goes from 0 to 100. However, there are other short-term evaluations that combine assessment of alertness with other acute effects. Some of these are the Global Vigour and Global Affect scale (GV and GA) (Monk, 1989), a VAS that includes 8 items (being alertness and sleepiness two of them), or the FACES (Shapiro et al., 2006), which includes five subscales (i.e. fatigue, energy, consciousness, energized and sleepiness), among others.

For long-term assessments, the Daytime Sleepiness Scale and Nocturnal Sleep Onset Scale (DSS and NSOS) was designed to report previous 2-weeks feelings of alertness. It was originally part of the Sleep Wake Activity Inventory (SWAI) (Rosenthal et al., 1993) that combines assessment of multidimensional components of sleepiness over the last week, similar to the Toronto Hospital Alertness Test (THAT) (Shapiro et al., 2006). To date, the only known scale to measure general levels of daytime sleepiness is the Epworth Sleepiness Scale (ESS) (Johns, 1991) , which rates daily life situations on a 4-points scale.

2.1.2 Objective measures

There are also different objective ways to evaluate an individual' neurophysiological need for sleep, or in other words, the strength of the arousal system on the sleep-wake spectrum. We find tests such as the Multiple Sleep Latency Test (MSLT) (Carskadon and Dement, 1977), where the participant is required to lie down in a dark room to fall asleep, or the Maintenance of Wakefulness Test (MWT), in which subjects are instructed to sit in a dimly lit room for 30 min and attempt to stay awake. Both of them assess state of wakefulness by standard electroencephalography (EEG) activity, which reflects the central nervous system (CNS).

On the other hand, changes in activity in the autonomic nervous system (ANS) are hypothesized to also serve as a predictor of alertness. In particular, the sympathetic division, which promotes arousal and energy generation, is often associated with higher alertness, whereas increases in parasympathetic division, which promotes rest activities, could be translated into increases in sleepiness (Kaida et al., 2006). This transition from wakefulness to sleep (or vice versa) in the ANS may be associated with various parameters, which are also clinical measures but potentially less intrusive, including heart rate (HR) and heart rate variability (HRV) as measured with electrocardiogram (ECG) activity, eye movement and blink rate measured with electrooculography (EOG) activity, or pupil size with pupillography. A number of other physiological biomarkers have also been suggested to correlate to some extent with alertness, such as changes in body temperature (i.e. core body temperature (CBT), skin temperature (ST) or skin conductance (SCL), measured with electrodermal (EDA) activity), and melatonin or cortisol concentration levels (Amundadottir, 2016).

But, as mentioned earlier, in psychology and neuroscience, alertness is also used as an indicator of sustained attention – which implies not only physiological arousal but also the ability to perform over extended periods – and of executive performance – which reflects alertness in combination with other cognitive functions such as inhibitory control or working memory. For the first case, simple non-clinical tests such as the Psychomotor Vigilance Task (PVT) (Dinges and Powell, 1985) or the Mackworth Clock Test (MCT) (Mackworth, 1948) are able to account for symptoms of sleepiness by recording reaction times of participants to a repetitive visual stimulus. The duration of the PVT test can be personalized from 1 to 10 minutes, and intervals of randomized appearance of stimuli can range from 1 to 10 seconds. An auditory and a visual version of this test exist. For the MCT the task duration is restricted to 30 minutes. In general, the shorter the reaction time (in milliseconds), the more ability to sustain attention (i.e. higher levels of alertness). Some other performance tests exist that measure executive performance, such as the Go/NoGo task or the N-back task (among others) that will not be discussed in this paper since alertness real contribution is difficult to isolate from other cognitive processes involved.

2.2 Evaluation of protocols based on existing studies

As we live in a 24-hour society with very dynamic routines, the extent to which results from nighttime laboratory studies are comparable to daytime situations applicable in real-life scenarios remains unclear. An overview of different methods used in daytime (from 6:00 am to 6:00 pm) empirical studies on acute alerting effects of light is provided in this section to assess the state of knowledge on that topic, and assess their feasibility for replication in other

contexts. Only studies which reported statistically significant results (whether positive or negative), as highlighted in Souman et al., 2018 and Lok et al., 2018, were considered, and in particular, only those that included either subjective or objective indicators (or both) applicable to semi-controlled experimental conditions (as described in Table 1).

This criteria led to restrict the pool of considered papers to 12 studies based on subjective indicators, which all demonstrated, one way or another, that alertness can be affected during daytime by manipulating lighting conditions, either with higher intensities of polychromatic white light (Huiberts et al., 2017, 2016; Maierova et al., 2016; Phipps-Nelson et al., 2003; Rüger et al., 2006; Smolders et al., 2012; Smolders and de Kort, 2014; te Kulve et al., 2017; Weisgerber et al., 2017), with shifts in spectrum towards higher CCT (Münch et al., 2016), or with shorter wavelengths from monochromatic light sources (Rahman et al., 2014; Revell et al., 2006).

The idea behind this collection was to underline which parameters from Table 1 pertaining to semi-controlled experimental designs were more frequently used in these studies, but also, to understand how they were used and to what extent they were successful in each case. IA summary of the main findings is provided in Table 2.

Table 2 – Overview of subjective and objective indicators for alerting effects

Authors, year	Light source	Light manipulation	Sample size	Subjective indicator	Objective indicator(s)	
					Performance	Physiology
Phipps-Nelson et al., 2003	Polychromatic white	Intensity	16	**KSS**	**PVT** (visual + auditory)	(not applicable)
Rüger et al., 2006			12	**KSS**	-	ECG - HR
Smolders et al., 2012			32	**KSS**	**PVT** (auditory)	**ECG - HR + HRV**
Smolders et al., 2014			28	**KSS**	**PVT*** (auditory)	ECG - HR + HRV
Huiberts et al., 2016			34	KSS	PVT (auditory)	**ECG – HR***
Maierova et al., 2016			32	**Not specified**	PVT (not specified)	-
Huiberts et al., 2017			34+39	KSS	PVT (auditory)	ECG - HR
te Kulve et al., 2017			19	**KSS**	**PVT*** (auditory)	-
Weisgerber et al., 2017			19	**KSS***	**PVT*** (not specified)	-
Revell et al., 2006	Monochromatic blue vs red LED	Wavelength	12	**VAS**	-	-
Rahman et al., 2014	Monochromatic (blue vs green)		16	KSS	**PVT** (auditory)	(not applicable)
Münch et al., 2016	Polychromatic (blue vs orange)	Spectral shifts	18	**VAS***	**PVT (not specified)**	(not applicable)

NOTE Indicators are represented in bold when significant effects (either positive or negative) are reported; indicators in bold with an asterisk * reported significant results not moderated by lighting condition itself, but by other variable.

2.2.1 Tested assessment methods

All these studies have in common that certain light manipulations were reported. According to the light source used, studies were classified in three categories: polychromatic white with changes on intensity, polychromatic "coloured" light with spectral shifts and monochromatic light with changes in wavelength. Most of the studies fall under the first category. Only a few explored other variables than light such as mental condition (fatigue vs control) (Smolders et al., 2012), chronotype (extreme morning vs evening) (Maierova et al., 2016), season (autumn/winter vs spring) (Huiberts et al., 2017), room temperature (cool vs neutral vs warm) (te Kulve et al., 2017), sleep condition (deprived vs rested) (Weisgerber et al., 2017) or longitudinal effects of prior light history (day 1 vs day 2 vs day 3) (Münch et al., 2016).

In nine out twelve studies, the KSS was used as a measure of subjective alertness. Of those, only five reported significant results as an effect of light condition (Phipps-Nelson et al., 2003; Rüger et al., 2006; Smolders et al., 2012; Smolders & de Kort, 2014; te Kulve et al., 2017), and two of them, as an effect of mental condition (Smolders & de Kort, 2014) or room temperature (te Kulve et al., 2017) as well. One study reported significant results without specifying the scale used for the assessment (Maierova et al., 2016), and another one reported significant effects on the KSS due to sleep condition, but not due to light manipulation (Weisgerber et al., 2017). Only one study using a VAS reported significant results (Revell et al., 2006) and none of them used the SSS to assess subjective alertness.

Except for two studies (Rüger et al., 2006; Revell et al., 2006), they all assessed alertness objectively as a measure of sustained attention using simple reaction time tasks. None of them used the MCT as an indicator, but instead, either an auditory or a visual version of the PVT was used. Four of those study reported significant effects due to light condition (Phipps-Nelson et al., 2003; Smolders et al., 2012; Münch et al., 2016; Rahman et al., 2014), and the other three achieved different performances due to the effect of mental condition (Smolders & de Kort, 2014), room temperature (te Kulve et al., 2017) or sleep condition (Weisgerber et al., 2017). Only five studies used physiological indicators valid for semi-controlled studies (i.e. HR or HRV measured with ECG activity), out of which two reported significant results, one due to the effect of light (Smolders et al., 2012) but the other one only due to the effect of time of the day (Huiberts et al., 2016).

It is important to note that not all indicators of alertness used in these studies have been reported in Table 2. This is because they did not fulfil the criteria of 'applicable parameters for semi-controlled studies', which was the priority criteria of the evaluation. The same happened with other biomarkers (e.g. melatonin suppression, cortisol levels, etc.) or other acute effects evaluated (e.g. mood, vitality, executive performance, etc.) that did not pertain to alertness. However, an overview of these other parameters is shown in Table 3.

Table 3 – Other acute effects, biomarkers or physiological measures used in daytime studies

Authors, year	Other acute effects	Other biomarkers	Other indicators of alertness
Phipps-Nelson et al., 2003	-	Melatonin levels	EEG and EOG
Rüger et al., 2006	Fatigue and energy	Cortisol concentration and core body temperature	EEG
Smolders et al., 2012	Mood, vitality, light appraisal, executive performance and tension	SCL	-
Smolders et al., 2014	Mood, vitality, light appraisal, executive performance, tension and self-control	SCL	-
Huiberts et al., 2016	Mood, vitality, light appraisal, executive performance and tension	SCL	-

Maierova et al., 2016	Mood, executive performance, wellbeing and relaxation	Melatonin levels and cortisol concentration	-
Huiberts et al., 2017	Mood, vitality, light appraisal and executive performance	-	
te Kulve et al., 2017	-	Melatonin levels, cortisol concentration, core body temperature and adrenaline	-
Weisgerber et al., 2017	-	Melatonin levels and oral temperature	-
Revell et al., 2006	Mood	-	-
Rahman et al., 2014	-	-	EEG
Münch et al., 2016	-	Melatonin levels	EEG

2.3 Lessons learned

Overall, research methodologies employed in earlier studies differed between them in a broad range of aspects, and in some cases, the problem was that certain details were missing, hampering replication. However, valuable lessons have been learned as demonstrated below.

It is not clear from the literature whether correlates of alertness are stronger with physiological indicators or with performance tasks. Therefore, it is a good practice to explore both paths together in a multi-measurement approach, as most of these studies did. Also, in order to better understand certain acute effects, and whether they are a consequence of the specific circadian phase of a person or not, it is important to (at least) keep track of wake-up and bed times by completing sleep and activity diaries, as done by Smolders et al., 2012 and Smolders & de Kort, 2014.

In some cases, various light manipulations were done during the experiment. For those studies testing the effect of monochromatic light or coloured polychromatic, intensity was not isolated. Therefore, it is unclear whether the effect is actually coming from the difference in colour or instead, from the difference in intensity. Also, duration of light exposure in all studies was on average 2.4 hours (SD±1.7) (except for Maierova et al., 2016, which was 16 hours). Consequently, this swings affected also the frequency of measurement blocks (in some cases, alertness was only measured twice, once at the beginning and again at the end).

Finally, most of the reviewed studies used a within-subject experimental design. In such cases, it is difficult to keep participants blind to light manipulations, arising bias from expectations or result associations, particularly in any kind of subjective measure of alertness. An a-priori power analysis is always recommended to determine an adequate sample size. However, none of the studies reported such an evaluation.

3 A consolidated protocol to assess alertness

Based on learnings from previous studies (Huiberts et al., 2017, 2016; Maierova et al., 2016; Phipps-Nelson et al., 2003; Rüger et al., 2006; Smolders et al., 2012; Smolders and de Kort, 2014; te Kulve et al., 2017; Weisgerber et al., 2017) summarized in section 2.3, a between-subjects experimental protocol was designed to test the effect of daylight spectral shifts and intensity changes throughout the day on various acute alerting responses for subjects in a semi-controlled but real scenario. Its main objectives were to evaluate the psycho-physiological effects of light exposure in realistic conditions (i.e. on individuals who spend prolonged hours and/or days exposed to it) with the following priorities:

- Focus exclusively on daylight i.e. no use of artificial light source.
- Test how a red-impoverished daylight spectrum compares to a neutral one if the visual illuminance level is maintained.

- Test how a brighter daylight intensity compares to a dimmer one if spectrum is maintained.
- Compare daylight manipulations in semi-controlled scenarios i.e. minimize confounding factors (e.g. effect of brightness change when observing effect of spectrum) while replicating realistic conditions.

3.1 Set-up

Two adjacent rooms with the same layout and characteristics are needed to conduct the semi-controlled studies. Orientation and window to wall ratio in these spaces should be kept constant between them, guaranteeing the same flow of light entering the rooms. As an example, the proposed protocol was applied to two similar south-oriented classrooms on the campus of EPFL.

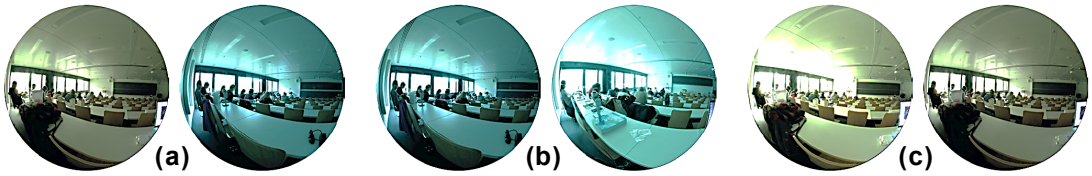

Figure 1 – Overview of all three daylight manipulations, corresponding to (a) neutral vs blue, (b) dim blue vs bright blue and (c) bright neutral vs dim neutral

Manipulations of daylight conditions illustrated in Figure 1 tested, separately, the effect of spectrum and intensity as follows: first, a neutral condition was compared to an environment illuminated with a red-impoverished spectrum (which looks “bluer”), while keeping both at the same low visual illuminance; in the other two experiments, changes in illuminance were applied with a constant spectrum (i.e. we tested the effect of daylight intensity and evaluated whether this effect was the same in spectrally neutral versus tinted (blue-shifted) conditions). Filters were used so as to keep a similar visual appearance (intensity or spectrum accordingly) in any pair of conditions (a, b or c). To maintain realistic indoor conditions, the rooms were equipped with commercially available electrochromic glazing technology, more and more widely used in architectural design nowadays, and which provides the opportunity to regulate daylight intensity and correlated colour temperature (CCT) through a change of tint in the glazing.

3.2 Design

Each study should be designed in batches of three consecutive days, including ideally seven continuous hours each day (starting at 9:00 until 16:00 h) and 8 measurement blocks (one per hour). This results in four measurements in the morning (9:00-12:00) and four in the afternoon (13:00-16:00), and should include an hour break for lunch (12:00-13:00) during which participants are not to be allowed to leave the rooms. Starting earlier could only be contemplated during summer months when daylight is available earlier.

The main reason to select this approach is due to the dynamic nature of daylight that undergoes continuous changes over 24-hour periods. As was later discovered from the findings introduced in Soto Magán et al., 2018, an exposure of at least two consecutive days also proved to be necessary to observe measurable differences, probably due to longitudinal effects of prior light history as described also in Münch et al., 2016.

Most of the reviewed studies used a within-subject experimental design. In such cases, it is difficult to keep participants blind to light manipulations, arising bias from expectations or result associations, particularly on any kind of subjective measure of alertness. To avoid these kind of problems, a between-subjects design should thus be followed, where half of the participants in each trial are to be randomly assigned to either a control group (classroom A) or an intervention group (classroom B). For the whole duration of the experiment, people should stay in the same classroom and in the same position, and should be exposed to one condition only. Different groups of participants should be used for the different experiments in order to reduce potential bias from personal differences.

3.3 Participants

For a between-subject design, with two independent groups, it would be necessary to have 88 participants per condition to detect a medium effect size (Cohen's d=0.5) as determined by an a-priori power analysis, but only 35 subjects per group for a large effect size (Cohen's d=0.8). Still, due to the complexity of this kind of studies and to both logistic and economic limitations (equipment availability, payment of fees to participants, etc.), it is not always easy to comply with such numbers Therefore these results should be only considered as a reference. In fact, in previous studies (Table 2) the total number of participants ranged from 10 to 39 in the case of within-subject experimental designs (which not always correspond to the sample size per condition), and was reduced to a sample size of 8 subjects per condition (16 participants in total) in the only two between-subject experiments (Phipps-Nelson et al., 2003 and Rahman et al., 2014).

Participants should be recruited with the following criteria: people from the same age group to minimize heterogeneity (for example, 18 to 32 years old to avoid differences in eye ageing), no abuse of caffeine or other stimulants, no use of drugs and no recent travels between time zones. Besides, volunteers should be asked to complete an extensive baseline questionnaire (around 30 min. duration) regarding their sleep quality, chronotype, general health and wellbeing, global perceived sleepiness, personality and overall vitality.

None of the participants should present extreme chronotypes as assessed by the Morning-Evening Self-Assessed Questionnaire (MEQ-SA) (Horne and Östberg, 1976), or poor quality scores in the Pittsburgh Sleep Quality Index (PSQI) (Buysse et al., 1989). None of them should report eye discomfort or complaints about their general health. They should all be free from sleep or psychological disorders, as assessed by the Patient Health Questionnaire (PHQ) (Kroenke et al., 2001), nor should they be neurotics according to the Big Five Inventory test (BFI) (John and Srivastava, 1999). Potential confounding factors such as environmental conditions and views to the outside, or psychological components of the procedure itself such as habituation, motivation, mental fatigue or stress, should be kept constant between rooms.

3.4 Procedure

A single-blind procedure should be followed and participants are not to be informed about the specific objectives of the study. The method should consist of three identical sessions, one per day. Participants should sit at the same spot throughout the experiment, within the same room. Each session should consist of eight measurement-blocks, and every hour the same procedure is to be followed: during the first 5 minutes, subjects are to be required to perform a simple reaction time task such as the PVT (available in an app-based version called PVT-touch and (Kay et al., 2013) and a 2-minutes short questionnaire on the smartphone provided; during the remaining 55-minutes in each measurement-block, participants are allowed to engage quiet and non screen-based activities of their preference (i.e. study and/or research-related activities on paper like reading, writing, etc.). These seven hours should include one-hour break for lunch, from 12:00 pm to 1:00 pm, in which participants are allowed to stand-up, talk to each other and move around the classroom. The full protocol is described in Figure 2.

Every day, participants should complete a sleep and activity diary describing time spent outdoors, wake-up and bedtime, etc. Leaving the room should not be allowed at any moment during the sessions, including lunchtime, except to eventually go to the bathroom. As a general rule during each session, consumption of coffee, tea, carbonated drinks or any other stimulants is strictly forbidden other than water. Also, participants should not be allowed to talk to each other, move from their places or listen to music. More importantly, the use of any screen device (phone, tablet, laptop, etc.) is to be strictly prohibited, except for the smartphones provided by the researcher.

On the first day of experiments, subjects are to be randomly assigned to one classroom or the other (control vs. intervention), trying to respect gender balance. Researchers should make sure everyone is wearing all the devices, and that these are properly attached and functioning. Rooms should be equipped and ready. Ten minutes before the start of the session, participants should enter the rooms and are to be individually assigned to specific seats (ideally, only those immediately close to the windows should be used to maximize

amount of daylight exposure), where they should remain for the entire experiment. Subsequently, experience and measurements should start at 9:00 am.

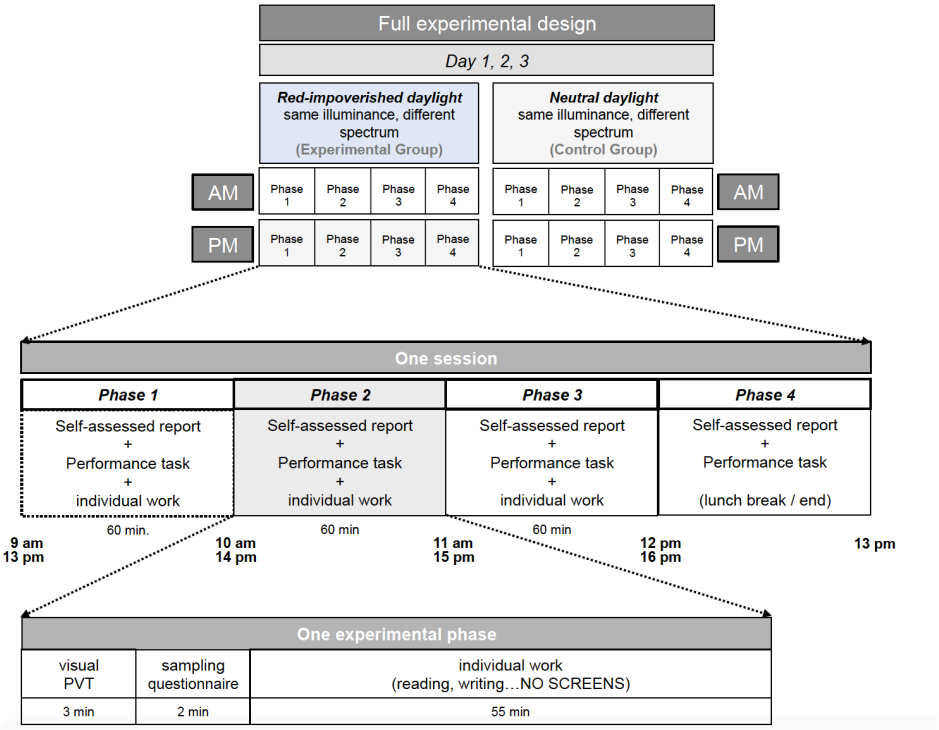

Figure 2 – Overview of the experimental design and procedure

3.5 Measurements

Each subject is to be individually equipped with the following material: one smartphone, one small wearable ECG monitoring device, a bag with 25 high-adherence electrodes, a wearable light sensor (recording at the eye level) and four wireless skin temperature sensors. Qualitative and quantitative methods should be used to assess psychophysiological alerting effects of daylight.

3.5.1 Subjective indicators

Subjective sleepiness should be evaluated, on an hourly basis, using the KSS and/or the SSS. For the first scale, ratings range from (1) ‘extremely alert’ to (9) ‘extremely sleepy’, whereas in the second one levels go from (1) ‘feeling active, vital, alert or wide awake’ to (7) ‘no longer fighting sleep, sleep onset soon, having dream-like thoughts’. Other combined scales could be used to assess subjective vigour and affect, such as the GV+GA, which is a VAS that ranges from (0) ‘not at all’ to (100) ‘very much’. Four items assess subjective vigour (‘alert’, ‘sleepy’, ‘weary’ and ‘effort’) and another four items assess subjective affect (‘happy’, ‘sad’, ‘calm’ and ‘tense’). Subjective vitalizing effects could also be assessed using a six-item test called Vitality Scale (VS) (Ryan and Frederick, 1997). Participants indicate how they are feeling, either (a) ‘alive and vital’, (b) ‘so alive I just want to burst’, (c) ‘I have energy and spirit’, (d) ‘look forward to each new day’, (e) ‘alert and awake’ or ‘energized’, on a 7-point scale ranging from (1) not true to (7) very true.

3.5.2 Objective indicators

A 1 to 10-minute visual PVT is to be used to measure reaction times and to assess the ability of participants to sustain attention. During the test, visual stimuli in the form of a geometrical black shape are presented on a white screen, at random intervals of 1-9 s. The goal is for participants to react as fast as possible by touching anywhere on the screen right after seeing the image. Ideally, HR and HRV are the indicators to be used to estimate physiological arousal of participants in response to the daylit environment. ECG signal should be recorded continuously during the experiment using wearable non-intrusive heart rate monitoring devices (for example, Firstbeat Bodyguard2 technology) and disposable electrodes. Subjects

are to be asked to wear their ECG device every day, not only during the sessions in the classrooms, but from wake-up to bedtime for the entire duration of the study. They should be also required to wear these devices two days previous to the experiment in order to ensure enough baseline data.

4 Conclusions

This paper investigates methods adopted in previous studies to assess acute non-visual responses to light. Main findings are that, for example, within-subject comparisons might include bias from expectations or result associations, and therefore between-subject designs are preferred. In line with this, a-priori power analyses are always recommended to determine an adequate sample size for the study. Also, that it is safer to use both qualitative and quantitative methods at the same time, namely subjective and objective evaluations, and to keep track of individuals' sleep patterns during the experiment. Finally, that it is important to clearly define the independent variables of the study, and that when testing different light manipulations, this should be done isolating the different qualities to be able to draw definitive conclusions.

From these findings, a new protocol is proposed, that aims to address the following issues: to compare daylight manipulations while minimizing confounding factors, such as changes in brightness when observing the effect of spectrum (i.e. only one variable at a time); to understand the potential of daylight alone as an alerting natural source; and to evaluate the effectiveness of red-impoverished daylight spectra compared to a neutral one, but also, whether a brighter daylight intensity compared to a dimer one (both in neutral and bluer conditions) is more effective when it comes to impacts on body functioning.

Whether results from nighttime laboratory studies on acute alerting effects of light are applicable to daytime situations still remains unclear. A lack of alertness could be a symptom of sleep deprivation, depression, circadian misalignment or sleep disorders, but also, impairment of alertness is often connected to daytime sleepiness and decreases in cognitive performance in the form of decision-making capabilities, which is likely to ultimately affect productivity. This protocol will hopefully help other researchers to consistently monitor effects of different daylighting conditions for prolonged hours on neurobehaviour and physiology, under semi-controlled conditions, but with limited intrusion for participants. Ideally, this would lead to increase comparability and interpretation of results, which are still needed in the field to draw more robust conclusions.

***Acknowledgements**:* this work was supported by the EPFL and in part by the SoLaCe project funded by EMPA. The authors would like to thank Forrest Webler who contributed to the running of the studies that were conducted as an application of the proposed methodology.

References

ÅKERSTEDT, T., & Gillberg, M. (1990). Subjective and Objective Sleepiness in the Active Individual. *International Journal of Neuroscience*, *52*(1–2), 29–37. https://doi.org/10.3109/00207459008994241

AMUNDADOTTIR, M. L. (2016). *Light-driven model for identifying indicators of non-visual health potential in the built environment*. Ecole polytechnique fédérale de Lausanne, Lausanne, Switzerland.

BADIA, P., Myers, B., Boecker, M., Culpepper, J., & Harsh, J. R. (1991). Bright light effects on body temperature, alertness, EEG and behavior. *Physiology & Behavior*, *50*(3), 583–588. https://doi.org/10.1016/0031-9384(91)90549-4

BERSON, D. M., Dunn, F. A., & Takao, M. (2002). Phototransduction by Retinal Ganglion Cells That Set the Circadian Clock. *Science*, *295*(5557), 1070–1073. https://doi.org/10.1126/science.1067262

BUYSSE, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

CAJOCHEN, C. (2007). Alerting effects of light. *Sleep Medicine Reviews*, *11*(6), 453–464. https://doi.org/10.1016/j.smrv.2007.07.009

CARSKADON, M. A., & Dement, W. C. (1977). Sleepiness and Sleep State on a 90-Min Schedule. *Psychophysiology*, *14*(2), 127–133. https://doi.org/10.1111/j.1469-8986.1977.tb03362.x

DINGES, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers*, *17*(6), 652–655. https://doi.org/10.3758/BF03200977

HATTAR, S., Liao, H. W., Takao, M., Berson, D. M., & Yau, K. W. (2002). Melanopsin-containing retinal ganglion cells: architecture, projections, and intrinsic photosensitivity. *Science*, *295*(5557), 1065–1070. https://doi.org/10.1126/science.1069609

HODDES, E., & Dement, W. (n.d.). ZarconeV (1972). The development and use of the Stanford Sleepiness Scale (SSS). *Psychophysiology*, *9*, 150.

HORNE, J. A., & Östberg, O. (1976). A Self-Assesment Questionnaire to Determine Morningness-Eveningness in Human Circadian Rhythms. *International Journal of Chronobiology*, *4*, 97–110.

HUIBERTS, L. M., Smolders, K. C. H. J., & Kort, Y. A. W. D. (2017). Seasonal and time-of-day variations in acute non-image forming effects of illuminance level on performance, physiology, and subjective well-being. *Chronobiology International*, *34*(7), 827–844. https://doi.org/10.1080/07420528.2017.1324471

HUIBERTS, Laura M., Smolders, K. C. H. J., & de Kort, Y. A. W. (2016). Non-image forming effects of illuminance level: Exploring parallel effects on physiological arousal and task performance. *Physiology & Behavior*, *164, Part A*, 129–139. https://doi.org/10.1016/j.physbeh.2016.05.035

JOHN, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research, 2nd ed.* (pp. 102–138). New York, NY, US: Guilford Press.

JOHNS, M. W. (1991). A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale. *Sleep*, *14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540

KAIDA, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., Fukasawa, K., 2006. Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology* 117, 1574–1581. https://doi.org/10.1016/j.clinph.2006.03.011

KAY, M., Rector, K., Consolvo, S., Greenstein, B., Wobbrock, J., Watson, N., & Kientz, J. (2013). PVT-Touch: Adapting a Reaction Time Test for Touchscreen Devices. *Proceedings of the ICTs for Improving Patients Rehabilitation Research Techniques*. Presented at the ICTs for improving Patients Rehabilitation Research Techniques, Venice, Italy. https://doi.org/10.4108/icst.pervasivehealth.2013.252078

KROENKE, K., Spitzer, R. L., & Williams, J. B. (2001). The Phq-9. *Journal of General Internal Medicine*, *16*(9), 606–613.

LOK, R., Smolders, K. C. H. J., Beersma, D. G. M., & Kort, Y. A. W. de. (2018). Light, alertness, and alerting effects of white light: a literature overview. *Journal of Biological Rhythms*, *33*(6), 589–601. https://doi.org/10.1177/0748730418796443

MACKWORTH, N. H. (1948). The Breakdown of Vigilance during Prolonged Visual Search. *Quarterly Journal of Experimental Psychology*, *1*(1), 6–21. https://doi.org/10.1080/17470214808416738

MAIEROVA, L., Borisuit, A., Scartezzini, J.-L., Jaeggi, S. M., Schmidt, C., & Münch, M. (2016). Diurnal variations of hormonal secretion, alertness and cognition in extreme chronotypes under different lighting conditions. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep33591

MONK, T. H. (1989). A visual analogue scale technique to measure global vigor and affect. *Psychiatry Research*, *27*(1), 89–99.

MÜNCH, M., Nowozin, C., Regente, J., Bes, F., De Zeeuw, J., Hädel, S., … Kunz, D. (2016). Blue-Enriched Morning Light as a Countermeasure to Light at the Wrong Time: Effects on Cognition, Sleepiness, Sleep, and Circadian Phase. *Neuropsychobiology*, *74*(4), 207–218. https://doi.org/10.1159/000477093

OKEN, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology*, *117*(9), 1885–1901. https://doi.org/10.1016/j.clinph.2006.01.017

PACHITO, D. V., Eckeli, A. L., Desouky, A. S., Corbett, M. A., Partonen, T., Wilson Rajaratnam, S. M., & Riera, R. (2016). Workplace lighting for improving mood and alertness in daytime workers. In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. Retrieved from http://doi.wiley.com/10.1002/14651858.CD012243

PHIPPS-NELSON, J., Redman, J. R., Dijk, D.-J. J., & Rajaratnam, S. M. (2003). Daytime exposure to bright light, as compared to dim light, decreases sleepiness and improves psychomotor vigilance performance. *Sleep*, *26*(6), 695–700.

PHIPPS-NELSON, J., Redman, J. R., Schlangen, L. J. M., & Rajaratnam, S. M. W. (2009). BLUE LIGHT Exposure Reduces Objective Measures of Sleepiness during Prolonged Nighttime Performance Testing. *Chronobiology International*, *26*(5), 891–912. https://doi.org/10.1080/07420520903044364

PROVENCIO, I., Rodriguez, I. R., Jiang, G., Hayes, W. P., Moreira, E. F., & Rollag, M. D. (2000). A Novel Human Opsin in the Inner Retina. *The Journal of Neuroscience*, *20*(2), 600–605. https://doi.org/10.1523/JNEUROSCI.20-02-00600.2000

RAHMAN, S. A., Flynn-Evans, E. E., Aeschbach, D., Brainard, G. C., Czeisler, C. A., & Lockley, S. W. (2014). Diurnal Spectral Sensitivity of the Acute Alerting Effects of Light. *SLEEP*. https://doi.org/10.5665/sleep.3396

REVELL, V. L., Arendt, J., Fogg, L. F., & Skene, D. J. (2006). Alerting effects of light are sensitive to very short wavelengths. *Neurosci Lett*, *399*(1–2), 96–100. https://doi.org/10.1016/j.neulet.2006.01.032

ROSENTHAL, L., Roehrs, T. A., & Roth, T. (1993). The sleep-wake activity inventory: A self-report measure of daytime sleepiness. *Biological Psychiatry*, *34*(11), 810–820. https://doi.org/10.1016/0006-3223(93)90070-T

RÜGER, M., Gordijn, M. C., Beersma, D. G., de Vries, B., & Daan, S. (2006). Time-of-day-dependent effects of bright light exposure on human psychophysiology: comparison of daytime and nighttime exposure. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *290*(5). https://doi.org/10.1152/ajpregu.00121.2005

RYAN, R. M., & Frederick, C. (1997). On Energy, Personality, and Health: Subjective Vitality as a Dynamic Reflection of Well-Being. *Journal of Personality*, *65*(3), 529–565. https://doi.org/10.1111/j.1467-6494.1997.tb00326.x

SHAPIRO, C. M., Auch, C., Reimer, M., Kayumov, L., Heslegrave, R., Huterer, N., … Devins, G. M. (2006). A new approach to the construct of alertness. *Journal of Psychosomatic Research*, *60*(6), 595–603. https://doi.org/10.1016/j.jpsychores.2006.04.012

SHAPIRO, C. M., Flanigan, M., Fleming, J. A. E., Morehouse, R., Moscovitch, A., Plamondon, J., … Devins, G. M. (2002). Development of an adjective checklist to measure five FACES of fatigue and sleepiness Data from a national survey of insomniacs. *Journal of Psychosomatic Research*, *7*.

SMOLDERS, K. C. H. J., de Kort, Y. A. W., & Cluitmans, P. J. M. (2012). A higher illuminance induces alertness even during office hours: Findings on subjective measures, task performance and heart rate measures. *Physiology & Behavior*, *107*(1), 7–16. https://doi.org/10.1016/j.physbeh.2012.04.028

SMOLDERS, K. C. H. J., & de Kort, Y. A. W. (2014). Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal. *Journal of Environmental Psychology*, *39*, 77–91. https://doi.org/10.1016/j.jenvp.2013.12.010

SOTO MAGÁN, V.E., Webler, F.S., Andersen, M. (2018). *Perceived and yet not seen: non-visual effects in daylit spaces*. Academy of Neuroscience for Architecture, Salk Institute, La Jolla, CA.

SOUMAN, J. L., Tinga, A. M., te Pas, S. F., van Ee, R., & Vlaskamp, B. N. S. (2018). Acute alerting effects of light: A systematic literature review. *Behavioural Brain Research*, *337*, 228–239. https://doi.org/10.1016/j.bbr.2017.09.016

TE KULVE, M., Schlangen, L. J., Schellen, L., Frijns, A. J., & van Marken Lichtenbelt, W. D. (2017). The impact of morning light intensity and environmental temperature on body temperatures and alertness. *Physiology & Behavior*, *175*, 72–81.

WEISGERBER, D. M., Nikol, M., & Mistlberger, R. E. (2017). Driving home from the night shift: a bright light intervention study. *Sleep Medicine*, *30*, 171–179. https://doi.org/10.1016/j.sleep.2016.09.010