International Commission on Illumination
Commission Internationale de l'Eclairage
Internationale Beleuchtungskommission

PP04

DETECTION OF THE STROBOSCOPIC EFFECT UNDER LOW LEVELS OF THE STROBOSCOPIC VISIBILITY MEASURE

Jennifer Veitch et al.

DOI 10.25039/x46.2019.PP04

from

CIE x046:2019

Proceedings
of the
29th CIE SESSION
Washington D.C., USA, June 14 – 22, 2019

(DOI 10.25039/x46.2019)

CIE Central Bureau
Babenbergerstrasse 9
A-1010 Vienna
Austria
Tel.: +43 1 714 3187
e-mail: ciecb@cie.co.at
www.cie.co.at

DETECTION OF THE STROBOSCOPIC EFFECT UNDER LOW LEVELS OF THE STROBOSCOPIC VISIBILITY MEASURE

Veitch, J.A.[1], Martinsons, C.[2]
[1] National Research Council of Canada, Ottawa, CANADA,
[2] Centre Scientifique et Technique du Bâtiment, Saint Martin d'Hères, FRANCE

jennifer.veitch@nrc-cnrc.gc.ca

DOI 10.25039/x46.2019.PP04

Abstract

Cyclic variations in lighting system luminous flux, known as temporal light modulation (TLM), may have visual, neurobiological, and performance and cognition effects on viewers. Researchers have derived a Stroboscopic Visibility Measure (SVM) to characterize the TLM signal in a manner that is thought to predict the visibility of the stroboscopic effect. An SVM of 1 means that the average person would detect the phenomenon 50% of the time. If there exists a broad range of visual sensitivity in the general population, setting a limit based on the average person might lead to an unacceptable risk of an adverse consequence for the most sensitive individuals. There is an absence of published data concerning the relationship of SVM to stroboscopic visibility among the general population. This experiment was conducted to provide such data, using commercially available lamps to provide a range of SVM conditions (SVM: 0; 0.4-0.6; 1.0; 1.6; and >2.0).

Keywords: temporal light artefacts; temporal light modulation; population; stroboscopic effect

1 Introduction

The introduction of solid-state lighting to the marketplace has brought renewed concern about possible adverse consequences of exposure to temporal cyclic or transient variations in lighting system luminous flux, known as temporal light modulation (TLM) (or more commonly, but incorrectly, referred to as 'flicker'). TLM may have visual, neurobiological, and performance and cognition effects on viewers (Wilkins et al., 2010, IEEE Power Electronics Society (IEEE), 2015). Visual perception effects occur very quickly, with very short exposures (IEEE, 2015). TLM can also cause ill effects on a longer time scale, such as disruptions to eye movements, headaches, and eyestrain (Wilkins, 2016, IEEE, 2015). There is not yet expert consensus about all of the possible health and behavioural effects of TLM, and no single metric to predict their occurrence. This remains an active area for research and standardization (CIE, 2017).

The visual perception effects are collectively known as temporal light artefacts (TLA), comprising flicker, the stroboscopic effect, and the phantom array effect (CIE, 2016). This paper concerns the stroboscopic effect, and its occurrence under non-directional lamp (A-lamp) replacement light sources varying in the degree of TLM. The experiment was designed to fill a gap in the literature by providing additional population data concerning the visibility of the stroboscopic effect, focusing particularly on young adults because this age group is thought to be more sensitive to TLM (Brundrett, 1974).

Recent years have brought several papers on the stroboscopic effect under solid-state lighting (SSL). Bullough and colleagues (2011) examined visibility and acceptability of flicker, and the stroboscopic effect, across a range of TLM frequencies, modulation depths, and duty cycles, and data from ten participants viewing each of nine conditions once, for less than two minutes. The stroboscopic effect was detectable at frequencies up to 300 Hz for a hand waving under the lamp, but acceptability of what had been seen was high for all frequencies at and above 120 Hz (controlling for modulation depth), and when modulation depth was 33% (but not 50% or 100%). A subsequent experiment (Bullough et al., 2012) also with ten participants but changing the task to a wand waved under the light source, found that 100 Hz TLM was detected 80% or more of the time for modulation depths equal to or greater than

25%. The conditions were rated as just acceptable, on average, when the modulation depth was 25%, but unacceptable at higher values of modulation depth. One limitation of these experiments is that participants controlled the movement that was viewed.

Perz and colleagues conducted a length series of experiments, the outcome of which was the proposed Stroboscopic Visibility Measure (SVM) to serve as a visibility measure for the stroboscopic effect (Perz et al., 2015, Perz et al., 2018). SVM is calculated from the TLM waveform for the light source (Perz et al., 2015, CIE, 2016). Its normalization to the visibility threshold means that, by definition, a value of 1 means that the average person would detect the phenomenon 50% of the time).

As noted by the CIE (2016) and by the National Electrical Manufacturers Association (NEMA) (2017), the visibility threshold (i.e., SVM=1) is not a guarantee of acceptability of the visible phenomenon. NEMA proposed SVM=1.6 as a limit value for indoor lighting, although noting that there was limited data on which to base a limit. Perz (2019) estimated that the average detection probability is 0.95 when SVM=1.5. More importantly, individuals who are more sensitive to TLM will detect the stroboscopic effect even at much lower levels of SVM. Additional data on these individuals will help society to provide conditions that are suitable for the full range of population diversity.

The most relevant paper to the question of acceptability is that of Perz and colleagues (2017), who reported a series of experiments as one data set with a combined range of SVM conditions from 0 to 4.9. An unknown number of participants performed various office-work-type tasks for variable amounts of time under varying sets of these conditions, and then rated the acceptability of the room conditions. Based on these data. Perz et al. constructed a logistic function to predict the percentage of annoyed people in relation to SVM, concluding that 20% of the population would be annoyed at an SVM of 1.5, and ~12% at an SVM of 1. It is problematic that there is no information available about the sample size or composition, because it is impossible to assess how generalizable the results might be. Moreover, by combining the results from studies in this way, the function combines within-subject and repeated-measures variation, which adds to the difficulty of interpreting the outcomes.

As far as we have been able to find, the published literature does not address the population rates for stroboscopic motion detection under varying levels of SVM. The metric's development has included several experiments, each with samples ranging from 20 to 35 individuals, the data from which show considerable individual variability in responses to TLM (Vogels et al., 2011, Perz et al., 2015). Perz and colleagues found that the relationship between the visibility threshold and TLM frequency held for individuals and for the average across individuals, and developed the metric based on the averaged response. Data have not been published showing the effect of varying levels of SVM on stroboscopic visibility in the population at large. Independent researchers were unable to assess the proportion of the population that detects the stroboscopic effect when SVM=1. Put another way, at what value of SVM does the stroboscopic effect become almost undetectable? This experiment was designed to provide answers to these questions.

2 Method

We used a repeated-measures design to test the visibility of the stroboscopic effect for five levels of SVM (0, 0.4-0.6; 1; 1.6; and >2) using an experimental method as similar as feasible to the published work from which the metric was developed (Vogels et al., 2011, Perz et al., 2015, Perz et al., 2018). To increase the statistical power, as well as to support the needs of sensitive populations, participants were young adults (Brundrett, 1974) and also completed a measure of pattern glare sensitivity [PGS] (Wilkins and Evans, 2012), which predicts sensitivity to headache and disrupted eye movements in response to TLM. We also collected preliminary information about how people judge the acceptability of their perceptions under the five SVM levels, in terms of comfort, pleasantness, and annoyingness. Data were collected in parallel in Canada and in France. All texts for participants were jointly developed in English, and translated in-house to French by CSTB.

2.1 Participants

Participants were university students. In Canada, the experiment was reviewed and approved by the NRC Research Ethics Board. In France, it was exempt from review and complied with the EU and French General Data Protection Regulation (GDPR). In both countries, participants completed an informed consent procedure at the start of the experimental session. In Canada, participants either received an honorarium of $20 for their participation or were awarded 1% bonus credit for a Psychology undergraduate course. In France, participants received a 15 € gift card for their participation. Table 1 reports the characteristics of the two samples.

Table 1 – Demographic characteristics of participants, by site

	Sex		Age		Pattern Glare
	Male	Female	18 to 29	30 to 39	>=2
CSTB (France)	11	7	17	1	2
NRC (Canada)	7	11	16	2	3

2.2 Apparatus and Lighting Conditions

At both sites, testing occurred in a small windowless room. The rooms were minimally furnished with a desk and chairs. During the session trials all the illumination came from the custom desktop luminaires described below. During the instructions and demographic questionnaires, a desk lamp with the low-SVM condition was used on a separate desk from the one holding the apparatus.

The teams at each site each independently designed and constructed their own apparatus to deliver the target light level (~300 lx) from a lamp suspended ~1 m above a table, in a device designed to permit the different lamps to be rotated into place by the experimenter as needed and in any order. Figure 1 shows images of the apparatus in each country.

Figure 1 – The apparatus in each country, CSTB on the left and NRC in the middle. The right panel shows a close-up of the visual task, as set up at NRC.

In joint planning, the research team together determined the target values for the five SVM conditions: 0, 0.4-0.6; 1; 1.6; and >2. We further agreed that the lamps would be ~3000 K lamps having 800 lumen rated output and P_{st}^{LM} values <1. This latter value was chosen in order to avoid any possible visible flicker (< 80 Hz TLM). Each site then obtained a wide selection of lamps and measured their output following the procedures outlined in established protocols (International Electrotechnical Commission (IEC), 2017, 2018, CIE, 2018, Ohno, 2014). Table 2 displays the characteristics of the five lamps used at each site, as measured in an otherwise dark room, horizontally in the location of the stroboscopic task.

Table 2 – Characteristics of the light sources at each site

	Labelled		Measured								
Lamp	CCT [K]	Lum. flux [lm]	Illum [lx]	CCT [K]	R_a	Duv	Dom. Frequ. [Hz]	Mod. [%]	FI [%]	P_{st}^{LM}	SVM
CSTB1	2700	806	344	2756	83	0.00001	100	2.1	0.6	0.39	0.00
CSTB2	2700	810	330	2810	82	0.0009	100	11.8	3.7	0.05	0.43
CSTB3	2600	720	318	2559	90	0.0022	100	27.8	7.9	0.08	0.96
CSTB4	2700	810	312	2641	81	0.0016	100	40.2	12.3	0.26	1.47
CSTB5	2700	600	324	2799	80	0.0022	100	79.4	26.9	0.38	3.09
NRC1	2700	800	341	2872	83	-0.0008	120	4.7	0.43	0.05	0.04
NRC2	2700	800	319	3018	84	-0.0016	120	14.0	3.79	0.07	0.42
NRC3	2700	800	354	2717	83	-0.0001	120	32.0	8.47	0.08	0.91
NRC4	3000	800	334	3094	83	-0.0023	120	55.6	13.25	0.06	1.38
NRC5	3000	800	335	3027	83	-0.0003	120	91.5	29.99	0.33	2.80
Note. Dom. freq. = dominant frequency. Mod % = modulation depth (%). FI – IES flicker index.											

2.3 Dependent measures

2.3.1 Stroboscopic Effect

There were two probes for the stroboscopic effect. The first of these used a white dot on a rotating horizontal black disk, as was used by prior researchers (Vogels et al., 2011, Perz et al., 2015). The dot on the disk rotates at a speed of 4 m/s, which is the typical speed of hands moving in an office context. The participant was asked to look at the disk and to report whether or not they saw individual dots (stroboscopic effect). The rotating disks used in this experiment were designed, assembled, and programmed at NRC using a programmable DC motor, two being shipped to CSTB with a suitable power cable for operation in France. The reflectance of the black surface was ρ=6.96%. and the white dot was ρ=90.85%, making the luminance ratio 13.05:1, as similar as possible to the original Vogels paper (see Figure 1).

We also added a vertical task. Participants were asked to look at black a dot on the end of the arm of a mechanical metronome (operating at 180 bpm in Canada and 150 bpm in France) and to report whether or not they saw individual dots or a blur. The metronomes were identical in the two countries, but each team made and attached its own black dot. Figure 1 shows the metronomes as well as the disc apparatus.

2.3.2 Acceptability ratings

Participants reported on the stroboscopic effect in a series of trials described below. On the last trial, they also rated the comfort, pleasantness, and annoyingness of that condition (three separate questions), each on a 5-point Likert scale from 0= not at all to 4 = extremely. Similar ratings had been obtained by Vogels et al.(2011).

2.3.3 Pattern glare sensitivity

Prior to the stroboscopic visibility trials, participants were asked to complete the Wilkins and Evans Pattern Glare Test (Wilkins and Evans, 2012). Scores on this test have been shown to be correlated to the propensity for headache associated with visual stimuli. The test consists of three plates of square wave patterns. After viewing each pattern, participants were asked three questions. The number of sensations indicated is summed (maximum 8). Sensitivity is the difference between the scores for Pattern 2 and Pattern 3. (Pattern 1 is a probe for response bias.) Normative scores have been established (Evans and Stevenson, 2008), to which scores were compared.

2.4 Procedure

When participants arrived, a lamp at the lowest SVM level provided the room light. Participants received information about the study and signed the consent form in this condition. They also completed a short paper-based questionnaire to record demographic information (age, sex, education, eye colour (Whissell, 1980), and visual corrections).

For the visual perception trials, the participant was asked to rotate away from the desk that held the task (facing the opposite wall) and to close his or her eyes while the researcher set up each trial. Setting up involved moving the light box to reveal one or another light source. Light sources were presented in blocks of five with the conditions in random orders in each block. The random orders of presentation were listed on a pre-printed data sheet for that session. The experimenter asked the participant to turn around, and to look at first the rotating disk to answer the question “Do you see white dots?” with an answer “yes” or “no”. Next, the same question was asked for the metronome. After this second question, the participant turned away and closed his or her eyes while the next trial was set up.

In the final block of five trials, the participant was asked to rate the appearance of the condition on the three scales described above after the metronome question.

At the conclusion of the session the participant was provided the debriefing information sheet and asked not to share the information with other potential participants.

3 Results

3.1 Stroboscopic visibility: horizontal task

The stroboscopic effect occurs when one sees the white spot as a distinct circle that jumps from one location to another. When there is no stroboscopic effect the white spot looks like an undifferentiated blur. For each trial the participant answered whether or not they saw the stroboscopic effect when looking at the rotating disc. For each participant, we summed the number of positive responses and calculated the proportion of “yes” responses out of the 10 trials (detection score, with theoretical minimum=0 and maximum=1.0). The SVM was developed from this task. It is intended that the average performance for an SVM=1 light source should be a detection score of 0.50. Table 3 shows the descriptive statistics for this task.

The data were not normally distributed, so we used the sign test to analyse planned comparisons between levels in sequence (condition 1 vs. condition 2; condition 2 vs. condition 3; condition 3 vs. condition 4; condition 4 vs. condition 5). This test computes the difference between the two conditions for each individual and compares the median value of the difference against the null hypothesis that it equals zero (as it would be if there was no systematic difference between the conditions). For the combined data (N=36), the comparison between condition 1 and condition 2 was not significant, but the comparisons between conditions 2 and 3, 3 and 4, and 4 and 5 were all statistically significant. The same pattern occurred for the CSTB data alone. For the NRC data, both the comparisons between conditions 1 and 2, and 2 and 3, were not statistically significant, but levels 3 and 4, and 4 and 5, did show differences.

Table 3 – Means, standard deviations, and quartile values for stroboscopic visibility scores for the rotating disc, for the full sample combined and separately for CSTB and NRC.

Condition	1	2	3	4	5
Nominal SVM	~0	~0,4	~0,9	~1,4	~3,0
Combined *M (SD)*	0,10 (0,17)	0,10 (0,16)	0,27 (0,33)	0,54 (0,37)	0,99 (0,07)
25th percentile	0,00	0,00	0,00	0,20	1,00
50th percentile	0,00	0,00	0,10	0,50	1,00
75th percentile	0,10	0,10	0,63	0,90	1,00
CSTB: *M (SD)*	0,12 (0,20)	0,11 (0,17)	0,34 (0,35)	0,65 (0,38)	1,00 (0,00)
25th percentile	0,00	0,00	0,00	0,35	1,00
50th percentile	0,00	0,00	0,20	0,80	1,00
75th percentile	0,20	0,23	0,70	1,00	1,00
NRC: *M (SD)*	0,08 (0,13)	0,09 (0,16)	0,21 (0,31)	0,44 (0,32)	0,97 (0,10)
25th percentile	0,00	0,00	0,00	0,20	1,00
50th percentile	0,00	0,00	0,10	0,40	1,00
75th percentile	0,10	0,10	0,25	0,60	1,00

3.2 Stroboscopic visibility: vertical task

We treated the data for the stroboscopic visibility scores for the metronome in the same way as we had for the rotating disc. The descriptive statistics are shown in Table 4. The stroboscopic effect was largely not visible to participants at CSTB, but was somewhat visible at NRC. The sign test results showed no statistically significant differences between SVM conditions for the CSTB data. For the NRC data, and overall, the scores were different only between conditions 4 and 5.

Table 4 – Means, standard deviations, and quartile values for stroboscopic visibility scores for the metronome, for the full sample combined and separately for CSTB and NRC.

Condition	1	2	3	4	5
Nominal SVM	~0	~0,4	~0,9	~1,4	~3,0
Combined *M (SD)*	0,17 (0,25)	0,14 (0,22)	0,16 (0,24)	0,21 (0,29)	0,30 (0,36)
25th percentile	0,00	0,00	0,00	0,00	0,00
50th percentile	0,00	0,00	0,00	0,10	0,10
75th percentile	0,30	0,28	0,30	0,38	0,68
CSTB: *M (SD)*	0,05 (0,10)	0,06 (0,10)	0,06 (0,10)	0,09 (0,15)	0,14 (0,27)
25th percentile	0,00	0,00	0,00	0,00	0,00
50th percentile	0,00	0,00	0,00	0,00	0,00
75th percentile	0,10	0,10	0,13	0,20	0,13
NRC: *M (SD)*	0,29 (0,30)	0,22 (0,27)	0,27 (0,28)	0,33 (0,34)	0,47 (0,38)
25th percentile	0,00	0,00	0,00	0,00	0,00
50th percentile	0,20	0,10	0,25	0,20	0,50
75th percentile	0,53	0,33	0,50	0,70	0,80

3.3 Acceptability ratings

Participants rated the acceptability of each light source in terms of comfort, pleasantness, and annoyingness on the tenth block of trials. We reverse-coded annoyingness so that for all of the ratings, a higher value indicated a more acceptable condition, and then averaged the three scores for each light source to give a single score for acceptability between 0 and 4. This scale showed acceptable internal consistency, calculated based on responses to condition 5 (Cronbach's alpha = 0.65). Note that the questions were asked in English at NRC and in French at CSTB, the translation having been performed in-house at CSTB. Table 5 displays the descriptive statistics for the acceptability ratings. Note that the mean values follow what one might consider to be the expected pattern of lower scores for higher SVM for the overall sample and for the NRC data alone, but not for the CSTB data.

The acceptability ratings were normally distributed. We examined the same planned comparisons as had been performed for the stroboscopic detection tasks, both overall and for each site separately. Of these, there was one statistically significant contrast, between the acceptability of condition 1 and condition 2 in the NRC sample (F (1,17) = 9.76, p=.006, $\eta 2partial$ = 0.37, Cohen's d= 0.74). This was a medium-to-large effect (Cohen, 1988).

Table 5 – Means, standard deviations, and quartile values for acceptability ratings, for the full sample combined and separately for CSTB and NRC.

Condition	1	2	3	4	5
Nominal SVM	~0	~0.4	~0.9	~1.4	~3.0
Combined *M (SD)*	2.30 (0.70)	2.19 (0.86)	2.08 (0.82)	2.14 (0.94)	2.19 (0.89)
25^{th} percentile	2.00	1.67	1.33	1.33	1.67
50^{th} percentile	2.33	2.00	2.33	2.33	2.00
75^{th} percentile	2.67	3.00	2.67	2.92	3.00
CSTB: *M (SD)*	2.24 (0.65)	2.37 (0.87)	2.24 (0.83)	2.37 (1.01)	2.41 (0.90)
25^{th} percentile	1.92	1.67	1.58	1.33	1.67
50^{th} percentile	2.33	2.67	2.33	2.50	2.33
75^{th} percentile	2.67	3.00	2.67	3.33	3.33
NRC: *M (SD)*	2.35 (0.76)	2.00 (0.83)	1.93 (0.80)	1.91 (0.82)	1.96 (0.85)
25^{th} percentile	1.83	1.33	1.33	1.25	1.25
50^{th} percentile	2.33	1.83	1.83	2.00	1.83
75^{th} percentile	3.00	2.75	2.42	2.42	2.67

4 Discussion

In the preparations for this experiment, both laboratories acquired a variety of LED lamps that are commercially available on the North American and European markets today. The lamps were measured under clean sinusoidal power supply conditions and found to exhibit a wide range of TLM characteristics, from nearly none to very high SVM. Each laboratory selected five lamps for this experiment, based on their SVM characteristics, taking care to seek similar performance at each chosen condition to permit the data to be combined. In this manner we obtained a larger – although not large -- data set in six weeks than either laboratory would have been able to collect on its own.

The primary visual perception task in this experiment was carefully chosen to replicate the rotating disc task with which the SVM was developed (Vogels et al., 2011, Perz et al., 2015). The metric is defined such that the median person ought to detect the stroboscopic effect 50% of the time if SVM=1; that is, half of the people will detect the stroboscopic effect more than 50% of the time, and half will detect it less than 50% of the time. The results of this experiment showed that, under these experimental conditions, the median detection was lower than expected; for our SVM=0.9 condition, median overall detection rate was 0.27 (27%), although it was (as expected) slightly higher for the CSTB lamp (SVM=0.96) than the NRC lamp (SVM=0.91). Stroboscopic visibility increased with higher SVM values. Moreover, the 75th percentile detected the effect 63% of the time overall when SVM was ~0.90; this means that 25% of the sample had a higher detection rate than 63%.

The detection rates for the vertical plane stroboscopic effect were different between CSTB and NRC, which might reflect subtle differences in the location of the metronomes in relation to the light sources, backgrounds, or other small methodological differences. For the NRC site, the vertical plane stroboscopic visibility approached the horizontal plane rotating disc visibility.

The judgements of acceptability did not show any consistent relation to SVM in the total sample, although in the NRC sample the best acceptability did occur with the SVM=0 lamp. This might have been because the viewing conditions were brief and were unusual. There was no surrounding room (ambient) lighting, only the direct light on the desk surface; this might have made it difficult for participants to make these judgements especially during a short exposure.

The following are limitations of this experiment that could constitute the basis of the differences which were observed between the two sites:

- A small sample size (N=18) at each site;
- Not enough very sensitive people (5 in total) participating to permit focused analysis;
- Young participants, although this could provide guidance on the upper limit of detection, as older people might be less sensitive;
- Similarly, a limited range of eye colour and ethnicity in the sample might have excluded some sensitive individuals;
- Short viewing times;
- Only 5 SVM conditions, leaving gaps between the levels where information is lacking;
- Non-immersive surroundings, which had been the setting for prior research; perhaps SVM becomes more annoying as one moves around the space; and,
- Only one visual perception outcome investigated, the stroboscopic effect; thus the data do not inform concerning possible effects of SVM level on detection of the phantom array, nor on complex phenomena like eyestrain, headache, reading or cognitive performance.

Decisions concerning standards and regulations are best made when based on a body of independently replicated evidence, and limits that are set necessarily reflect societal consensus about the balance of evidence and the tolerance for risk (CIE, 2017). Discussions concerning the best metrics to characterise TLM and suitable limits for them are expected to continue, but in the meantime our lit environment is being transformed by long life SSL products being sold today, some of which have very high SVM values. This experiment has been conducted to provide evidence for a limit on SVM where, at present, very limited evidence exists (NEMA, 2017, Perz et al., 2017).

The decision to place a limit on any metric based on its effects on people involves at least three choices: (1) The choice of an effect on which to focus; (2) the acceptable frequency of the outcome occurring; and (3) the acceptable proportion of the population who might experience this outcome (CIE, 2017). These choices are value judgements that research can inform but cannot answer.

The stroboscopic effect might not be the most important perceptual, behavioural, or health effect that TLM can cause. Some might argue that whether or not it is detected depends on there being moving objects, a condition that might not occur all of the time in common interior spaces. It has the merit, as a visual perception effect, of occurring rapidly, whereas other effects that might be of interest occur on a longer time scale or are difficult to study because of ethical concerns (e.g., headache incidence or severity). In recent years it has been the chosen effect by various researchers (Bullough et al., 2011, Bullough et al., 2012, Vogels et al., 2011, Perz et al., 2015), and as a consequence there exists a metric that could be used in regulations (Perz et al., 2015, CIE, 2016, IEC, 2018). Therefore, this experiment sought data concerning the occurrence of the stroboscopic effect in a sensitive population in relation to varying levels of that metric, to provide the information needed as part of the value judgement discussions.

Further data from a larger sample, which is expected to include a larger number of people with high pattern glare sensitivity, can provide deeper insights than are presented here. A larger sample of participants will capture more people from the most sensitive parts of the population, and could serve as the foundation for further discussions between all parties concerning suitable limit values for a TLM metric. To support that discussion, this research is on-going, with new results expected in mid-2019.

References

BRUNDRETT, G. W. 1974. Human sensitivity to flicker. *Lighting Research and Technology,* 6**,** 127-143.

BULLOUGH, J. D., SWEATER HICKCOX, K., KLEIN, T. R., LOK, A. & NARENDRAN, N. 2012. Detection and acceptability of stroboscopic effects from flicker. *Lighting Research and Technology,* 44**,** 477-483.

BULLOUGH, J. D., SWEATER HICKCOX, K., KLEIN, T. R. & NARENDRAN, N. 2011. Effects of flicker characteristics from solid-state lighting on detection, acceptability and comfort. *Lighting Research and Technology,* 43**,** 337-348.

COHEN, J. 1988. *Statistical power analysis for the behavioral sciences,* Hillsdale, NJ, Erlbaum.

CIE 2016. CIE TN006:2016*. Visual aspects of time-modulated lighting systems – Definitions and measurement models*. Vienna, Austria: CIE.

CIE 2017. CIE TN 008:2017. *Final report CIE stakeholder workshop for temporal light modulation standards for lighting systems*. Vienna, Austria: CIE.

CIE 2018. CIE 15:2018, *Colorimetry*. 3 ed. Vienna, Austria: CIE.

EVANS, B. J. W. & STEVENSON, S. J. 2008. The Pattern Glare Test: a review and determination of normative values. *Ophthalmic and Physiological Optics,* 28**,** 295-309.

IEEE POWER ELECTRONICS SOCIETY 2015. IEEE S1789-2015. *IEEE recommended practices for modulating current in high-brightness LEDs for mitigating health risks to viewers.* New York, NY: Institute for Electrical and Electronics Engineers, Inc. (IEEE).

INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC) 2017. IEC TR 61547-1:2017. *Equipment for general lighting purposes - EMC immunity requirements - Part 1: An objective light flickermeter and voltage fluctuation immunity test method.* Geneva, Switzerland: IEC.

INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC) 2018. IEC TR 63158:2018. *Equipment for general lighting purposes - Objective test method for stroboscopic effects of lighting equipment*. Geneva, Switzerland: IEC.

NATIONAL ELECTRICAL MANUFACTURERS ASSOCIATION (NEMA) LIGHTING SYSTEMS DIVISION 2017. NEMA 77:2017. *Standard for temporal light artifacts: Test methods and guidance for acceptance criteria*. Rosslyn, VA: NEMA.

OHNO, Y. 2014. Practical use and calculation of CCT and Duv. *LEUKOS,* 10**,** 47-55.

PERZ, M. 2019. Modelling visibility of temporal light artefacts. Ph.D., Technische Universiteit Eindhoven.

PERZ, M., SEKULOVSKI, D. & BEECKMAN, P. 2017. Acceptability criteria for the stroboscopic effect visiblity measure. *Proceedings of the Conference "Smarter Lighting for Better Life" at the CIE Midterm Meeting 2017, October 23-25, 2017, Jeju, South Korea.* Vienna, Austria: Commission Internationale de l'Eclairage.

PERZ, M., SEKULOVSKI, D., VOGELS, I. M. & HEYNDERICKX, I. E. J. 2018. Stroboscopic effect: Contrast threshold function and dependence on illumination level. *Journal of the Optical Society of America A: Optics and Image Science, and Vision,* 35**,** 309-319.

PERZ, M., VOGELS, I. M. L. C., SEKULOVSKI, D., WANG, L. L., TU, Y. & HEYNDERICKX, I. E. J. 2015. Modeling the visibility of the stroboscopic effect occurring in temporally modulated light systems. *Lighting Research and Technology,* 47**,** 281-300.

VOGELS, I. M., SEKULOVSKI, D. & PERZ, M. 2011. Visible artefacts of LEDs. *Proceedings of the 27th Session of the Commission Internationale de l'Eclairage, Sun City, South Africa.* Vienna, Austria: CIE.

WHISSELL, C. M. 1980. Sex, eye-color, and aesthetic ratings of colors. *Perceptual and Motor Skills,* 51**,** 1012-1014.

WILKINS, A. J. 2016. A physiological basis for visual discomfort: Application in lighting design. *Lighting Research and Technology,* 48**,** 44-54.

WILKINS, A. J. & EVANS, B. J. W. 2012. *I.O.O. Pattern Glare Test,* London, UK, i.O.O Sales Ltd.

WILKINS, A. J., VEITCH, J. A. & LEHMAN, B. 2010. LED lighting flicker and potential health concerns: IEEE Standard PAR1789 update. *Proceedings of the Energy Conversion Congress and Exposition (ECCE) 2010 IEEE, 12-16 Sept., 2010.* New York, NY: Institute of Electrical and Electronics Engineers.

Acknowledgements

This project was funded through the IEA 4E SSL Annex with direct contributions from the governments of Australia, Canada, Denmark, France, Sweden and the United Kingdom. Additional support was provided by the National Research Council of Canada through its High Performance Buildings program (project A1-014733) and by the Centre Scientifique et Technique du Bâtiment through its Health and Comfort Program.

The authors thank Nils Borg (IEA 4E SSL Annex), Steve Coyne (Light Naturally), Carsten Dam-Hansen (Technical University of Denmark), Peter Erwin (Der LichtPeter), Chris Herdman (Carleton University Applied Cognitive Engineering Laboratory), Naomi Miller (Pacific Northwest National Laboratory), and Michael Scholand (IEA 4E SSL Annex) for their assistance. At the NRC, contributing staff included Chantal Arsenault, Erhan Dikel, Steven Kruithof, Sandra Mancini, Travis Moore, Saša Muradori and Mark Vuotari. At the CSTB, contributing staff included Nicolas Picard, Jan Jagla and Simon Bailhache.