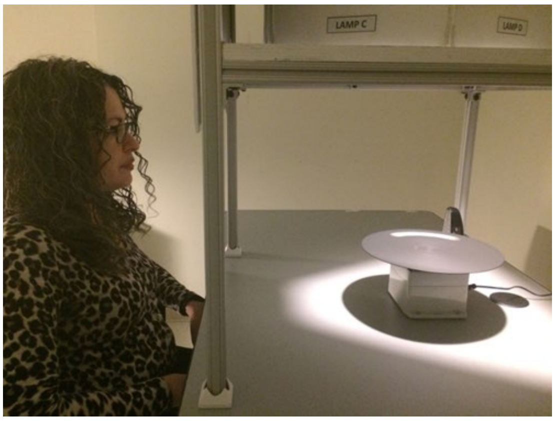


Interim Report:

Visual Perception under Energy-Efficient Light Sources - Detection of the Stroboscopic Effect Under Low Levels of SVM

Jennifer A. Veitch
National Research
Council of Canada

and

Christophe Martinsons
Centre Scientifique et
Technique du Bâtiment (France)

11 December 2018

Dr. Jennifer A. Veitch
Principal Research Officer
NRC Construction, Bldg M-24
National Research Council of Canada
1200 Montreal Road
Ottawa, ON, K1A 0R6,
Canada

https://www.nrc-cnrc.gc.ca/eng/

Dr. Christophe Martinsons
Head of division
Division of Lighting and Electromagnetic Fields
Centre Scientifique et Technique du Bâtiment
24, rue Joseph Fourier
38400 Saint Martin d'Hères
France

http://www.cstb.fr/en/

Acknowledgements

The authors are grateful to the governments of Australia, Canada, Denmark, France, Sweden and the United Kingdom for their direct contributions to fund this study, and to the larger IEA 4E SSL Annex for their support and expert assistance with the design and implementation of this project. Additional support was provided by the National Research Council of Canada High Performance Buildings program (project A1-014733).

The authors would like to thank Carsten Dam-Hansen from the Technical University of Denmark and Steve Coyne of Light Naturally (Australia) for their guidance and input in the project design and review of the findings. The authors also wish to express their deep gratitude to Peter Erwin, founder of Der LichtPeter laboratory and internet site, for sharing his knowledge and extensive database of lamps and luminaires as well as kindly providing some of the LED lamps used by the CSTB in this research. They also thank Naomi Miller of Pacific Northwest National Laboratory in the USA for her assistance in the selection of the LED lamps used by the NRC.

At the National Research Council of Canada, contributing staff included Chantal Arsenault, Erhan Dikel, Steven Kruithof, Sandra Mancini, Travis Moore, Sasa Muradori, and Mark Vuotari. We thank Prof. Chris Herdman of the Carleton University Applied Cognitive Engineering Laboratory for his assistance with data collection. At the Centre Scientifique et Technique du Bâtiment in France, contributing staff included Nicolas Picard, Jan Jagla and Simon Bailhache.

Disclaimer

About the IEA 4E Solid State Lighting Annex

The SSL Annex was established in 2010 under the framework of the International Energy Agency's Energy Efficient End-use Equipment (4E) Implementing Agreement to provide advice to its member countries seeking to implement quality assurance programmes for SSL lighting. This international collaboration currently consists of the governments of Australia, Canada, Denmark, France, the Republic of Korea, Sweden, the United Kingdom and the United States of America. Information on the 4E SSL Annex is available from: http://ssl.iea-4e.org/

Copyright:

Table of Contents

Executive Summary

Temporal light modulation (TLM, known colloquially as "flicker") of light sources has visual, neurobiological, and performance and cognition effects on viewers. This study aims to address an important gap in the scientific literature on the measurement of levels of TLM of LED light sources that may affect human health and productivity.

The outcomes of this research are intended to assist decision makers to develop new lighting policy measures and draft regulation. In particular this interim report has been issued to provide information in advance of an upcoming decision by European Union (EU) member countries on European Commission lighting policy Ecodesign regulation. A more detailed final report will be released early in 2019.

The IEC and the CIE have identified two metrics which may be measured and used to characterize lighting systems' TLM:

- P_{st}^{LM}, short term flicker metric for visible flicker at frequencies below 80 Hz) ($P_{st}^{LM} \leq 1.0$ is the proposed limit in draft EU regulation), and
- SVM, Stroboscopic Visibility Measure, for the higher frequency stroboscopic effect (SVM ≤ 1.6 is the proposed limit in draft EU regulation)

Although the development of these metrics and their associated measurement protocols continues, there is a parallel discussion concerning the appropriate levels. At a meeting in November 2018, Experts of the IEA 4E Solid-State Lighting Annex noted that the proposed European limits might be too high to provide adequate protection from adverse human impacts.

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. 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 on the market today, some of which have very high SVM values.

At present there is no public information concerning the occurrence of the stroboscopic effect under lamps of varying SVM values. To provide further information on levels of SVM, the IEA 4E SSL Annex commissioned this study 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 similar to previous research from which the metric was developed, with a target sample size of ~50 people across two sites (in France and Canada).

The decision to place a limit on any metric involves two choices: (1) The acceptable frequency of the outcome occurring; and (2) the acceptable proportion of the population who might experience this outcome.[7] These choices are value judgements that research can inform, but cannot answer. For the purposes of the draft lighting regulation that European Member States are discussing in December 2018, the following guidance can be drawn from this work:

- An SVM>2.0 caused virtually all of the participants to perceive stroboscopic effects of the disk all of the time.
- The proposed limit of SVM=1.6 is higher than the SVM for magnetic-ballasted T12 lamps,[8] which are known to cause headaches and eyestrain and to disrupt eye movements.[1,9-11]
- The most sensitive 25% of the people detected stroboscopic effects with the disk 90% or more of the time at SVM=1.4 (75th percentile overall). The EU-28 population includes ~101 million people between the ages of 0-30. Based on the data presented here, the proposed limit of SVM=1.6 would mean that nearly all of the time, one quarter of these 101 million young people could perceive the stroboscopic effect.

- The 75th percentile detection rate dropped to 63% when the SVM was ~0.9, meaning that only one-quarter of the sample could detect the stroboscopic effect more than 63% of the time for these lamps. A reduction of the limit value to SVM=0.9 would mean that only 25% of the population of young people (or 25.4 million people) would detect the stroboscopic effect more than 63% of the time.
- At SVM levels of 0.4 and below, the stroboscopic detection rate for the most sensitive quarter of the people dropped to 10%.

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 referred to as 'flicker'). TLM may have visual, neurobiological, and performance and cognition effects on viewers.[1-3] The visual perception effects are collectively known as temporal light artefacts (TLA), comprising flicker, the stroboscopic effect, and the phantom array effect. Definitions of these phenomena, proposed TLM frequency ranges for their occurrence, and guidance for possible TLM metrics to predict the artefacts may be found in CIE Technical Note TN006:2016.[4] The IEC has published informative technical reports for a metric to address flicker [P_{st}^{LM} [5]] and the stroboscopic effect [SVM[6]].

Visual perception effects such as the stroboscopic effect occur very quickly, with very short exposures.[1,2] TLM can also cause ill effects on a longer time scale, such as disruptions to eye movements, headaches, and eyestrain.[1-3] 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.[7]

Regulators in some regions are considering the importance of establishing limits on TLM from lighting products under the precautionary principle, to prevent large numbers of long-life light sources coming into use that could adversely affect the health and well-being of the public. At present (December 2018), the most advanced of these proposals is a draft European Commission lighting policy[*] under the Ecodesign Directive that would set maximum limits on two metrics to limit two visual effects: flicker, for which it is proposed that the limit on P_{st}^{LM} would be less than or equal to 1.0, and the stroboscopic effect, for which it is proposed that SVM would be less than or equal to 1.6.

Stroboscopic Visibility Measure (SVM) is a visibility measure for which, by definition, a value of 1 means that the average person would detect the phenomenon 50% of the time; thus, a light source having an SVM value of 1 would mean that the average person can detect the stroboscopic effect 50% of the time when that light source is the sole source of illumination. This means that half of the population will detect the stroboscopic effect more than 50% of the time, and half will detect it less than 50% of the time. As noted by the CIE and by the National Electrical Manufacturers' Association (NEMA), the visibility threshold (i.e., SVM=1) is not a guarantee of acceptability of the visible phenomenon[4, 8] At SVM=1.6 (the proposed level in the draft European lighting regulation), the average (median) person would detect stroboscopic motion considerably more than 50% of the time, and half of the population would have an even higher probability of detection, with some people likely to detect it all of the time. The regulators who drafted the proposed policy suggested this value based on the one published standard that contained a proposed value.[8]

At the IEA 4E Solid State Lighting Annex's Expert Meeting in Ottawa, Canada (held 30 October to 1 November 2018), the Experts noted that the proposed European limits might be too high to provide adequate protection from adverse human health impacts. In particular, the value of SVM of 1.6 is slightly higher than the TLM characteristic of a magnetically-ballasted T12 fluorescent lamp (SVM of 1.3), which prior research established as a cause of disrupted visual performance[9, 10] and a likely cause of headache and eyestrain.[11] The Experts noted that there is limited published data to support establishing more stringent limits on this TLM metric based on any of the health and behavioural consequences of TLM, and no publicly-available data even for SVM that would provide population-

[*] Draft European Lighting Regulation: https://ec.europa.eu/info/law/better-regulation/initiative/1551/publication/310970/attachment/090166e5be38e163_en
Draft European Lighting Annexes: https://ec.europa.eu/info/law/better-regulation/initiative/1551/publication/310970/attachment/090166e5be38e164_en

based evidence for limits that could reduce the proportion of the population who might detect such effects.

This document provides the interim findings (i.e., further testing will be conducted and a full report published in 2019) on a project designed to provide some of the necessary data to support policy-makers' determination of a suitable limit for SVM. To increase the sample size in the time available and to ensure applicability of the results around the world, the experiment was conducted at two sites in parallel, using the same experimental design and methods. The two subject test sites were in Ottawa, Canada, led by the Jennifer Veitch at the National Research Council of Canada (NRCC) and Grenoble, France, led by Christophe Martinsons at the Centre Scientifique et Technique du Bâtiment (CSTB). Additional support was provided by other experts affiliated with the IEA 4E SSL Annex, particularly Carsten Dam-Hansen (Technical University of Denmark, Roskilde, Denmark) and Steve Coyne (Light Naturally, Brisbane, Australia).

The objectives of the project are as follows:

- 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[12, 13], with a target sample size of ~50 people across the two sites;
- Examine the population frequency of pattern glare sensitivity [PGS][14], which is known to predict sensitivity to headache and disrupted eye movements in response to TLM (see Appendix A, section A.3.2);
- If possible, establish preliminary information about the visibility of the stroboscopic effect by individuals high in PGS; and
- Collect preliminary information about the how people judge the acceptability of their perceptions under the five SVM levels, in terms of comfort, pleasantness, and annoyingness.

2. Results

2.1 Sample and lighting details

This interim report is based on data from 36 people, 18 tested in France and 18 tested in Canada. Table 1 summarises their characteristics. The samples in the two sites were comparable. All participants were university students. The limited age range was chosen because there is evidence that younger people may be more sensitive to TLM,[15] and all had self-reported normal or corrected-to-normal vision, and normal hearing. Five participants in total showed responses to pattern glare that would indicate a risk of visual stress in everyday life.[14]

Table 1. Demographic characteristics of participants, by site

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

LED lamps were selected and purchased from the market to match the five levels of SVM, and hence represent SSL products available in the market in North America and Europe. Although the specific

[†] The high level of SVM > 2 was included to provide a validation of the test method by including a level known to be above the proposed limit value, and for which it should be the case that stroboscopic effect detection rates would be high.

lamps used by the laboratories were different in the two countries, the five experimental conditions were matched in terms of SVM for methodological consistency. Pst^{LM} is a metric that characterizes TLM in the range of 0-80 Hz, which is the range in which viewers can report seeing the temporal variation in light output (flicker). The light sources were chosen to keep Pst^{LM} very low to avoid experimental confounding. Details are available in Appendix A. Table 2 summarizes the key information about the lamps used at each site, showing that they are similar enough to be expected to produce comparable visual perceptions. Lamps were tested and operated with a clean sinusoidal power supply. At both sites the illuminance on the surface of the rotating disc was between 330-350 lux, the light source correlated colour temperatures (CCT) were in the range of 2700-3000K, and the general colour rendering index, R_a, was between 80-85.

Table 2. Light source characteristics, by site

			SVM		
Condition	**1**	**2**	**3**	**4**	**5**
CSTB	0	0.43	0.96	1.47	3.09
NRC	0.04	0.42	0.91	1.38	2.80
			Pst^{LM}		
Condition	**1**	**2**	**3**	**4**	**5**
CSTB	0.39	0.05	0.08	0.26	0.38
NRC	0.05	0.07	0.08	0.06	0.33

2.2 Stroboscopic effect – Rotating disc

Participants each completed 10 trials for each of the five lamps. For each trial, the individual looked at a rotating black disc on which was a white spot. 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. (See Appendix A, section A.3.3). For each trial the individual 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 rate, with theoretical minimum=0 and maximum=1.0). The SVM metric was developed from this task. It is intended that the median performance for an SVM=1 light source should be a detection rate of 0.50.

We used analysis of variance (ANOVA) to test for differences in the detection rate, both for the whole sample (N=36) and for each site separately. The results were the same for both the NRC and the CSTB data, and are reported here for the combined data set as well as for each separately (Table 3). For both samples and overall, the only paired comparison that did not show a statistically significant difference was the two lowest levels: Conditions 1 (SVM ~0) and 2 (SVM ~0.4) delivered equivalent results, of mean detection 10% of the time. The detection rates for each of the higher SVM conditions progressively increased and all the increases were statistically significant. For condition 3 (SVM ~0.90), mean detection was 27%; for condition 4 (SVM ~1.4), mean detection was 54%, and for condition 5 (SVM ~ 3.0), mean detection was 99%.

Figure 1 shows the 75th percentile detection rate for the stroboscopic effect across the two sites and the average of both sites. This means that for each SVM level, the bars are show the detection rates for 75% of the sample; 25% of the sample had a detection rate higher than the top of the bar. The choice of population statistic has no hard-and-fast rule; it reflects a judgement about what proportion of the population might be placed at risk. We chose to focus on the 75th percentile because of our relatively small sample; with a larger sample one would have a more precise estimate of the true value for the population as a whole and could focus on a smaller portion of the whole without basing a conclusion on a tiny number of people.

Figure 1. This chart shows the detection rate for the stroboscopic effect on the rotating disc for 75% of the participants, for each lamp type, for each site separately and for the whole sample combined ("All"). One-quarter of the participants had a higher detection rate than the height of the bar.

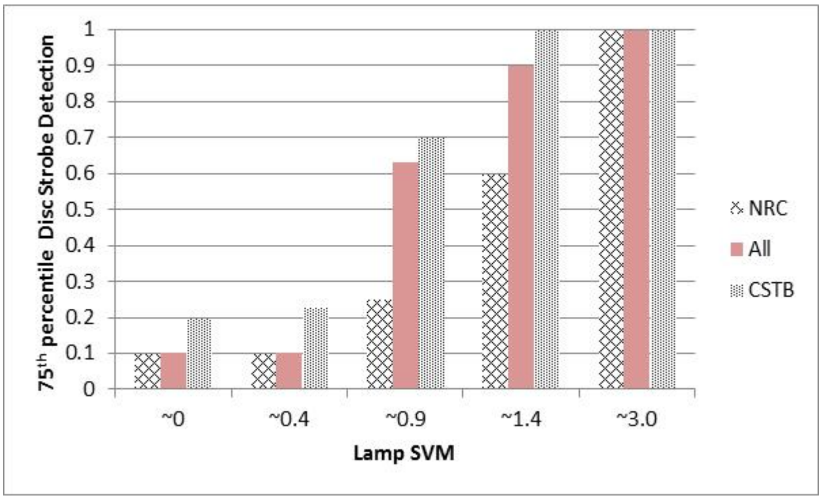


Table 3. The statistical test results (ANOVA) for the full sample and for the two sites separately have consistent strong linear trends of higher mean rotating disc stroboscopic effect detection with higher SVM. The sample distributions show how individual scores are distributed (see text).

	Statistics					ANOVA Test Statistics – Linear Trend			
Condition	**1**	**2**	**3**	**4**	**5**	***F***	***df***	***p***	$\eta^2_{partial}$
SVM	~0	~0.4	~0.9	~1.4	~3.0				*(effect size)*
Both sites	.10	.10	.27	.54	.99	809.30	1,35	.000	.96
Means (StDev)	(.17)	(.16)	(.33)	(.37)	(.07)				
Sample distribution									
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	0.12	0.11	0.34	0.65	1.00	355.97	1,17	.000	.95
Means (StDev)	(0.20)	(0.17)	(0.35)	(0.38)	(0.00)				
Sample distribution									
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						495.57	1,17	.000	.97
Means (StDev)	0.08 (0.13)	0.09 (0.16)	0.21 (0.31)	0.44 (0.32)	0.97 (0.10)				
Sample distribution									
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				

2.3 Stroboscopic effect - Metronome

In addition to the horizontal rotating disc detection task discussed in section 2.2, the authors developed a new vertical task for this experiment. Participants were told to look at a black dot on a mechanical metronome moving at 150 bpm at CSTB (where the dominant frequency was 100 Hz) and 180 bpm at NRC (where the dominant frequency was 120 Hz). The participants were asked whether they saw a series of black dots (stroboscopic effect) or if they saw a black blur (no stroboscopic effect detected). The metronome location differed slightly for the NRC and CSTB. At NRC it was farther back than the rotating disk, and slightly to the side, with the white wall of the room behind. At CSTB the line of sight placed the desk surface behind the moving metronome arm. The answers were summed to provide a detection proportion, as for the horizontal disc detection task.

The analysis of variance for the metronome detection rates showed an overall effect (Table 4), but the separate analyses for the two sites showed a strong difference, with nearly no detection at CSTB and a moderate detection rate at NRC. This is evident also from the median values shown in Figure 2.

Figure 2. This chart shows the detection rate of the stroboscopic effect on the metronome for 75% of the participants, for each lamp type, for each site separately and for the whole sample combined ("All"). One-quarter of the participants had a higher detection rate than the height of the bar.

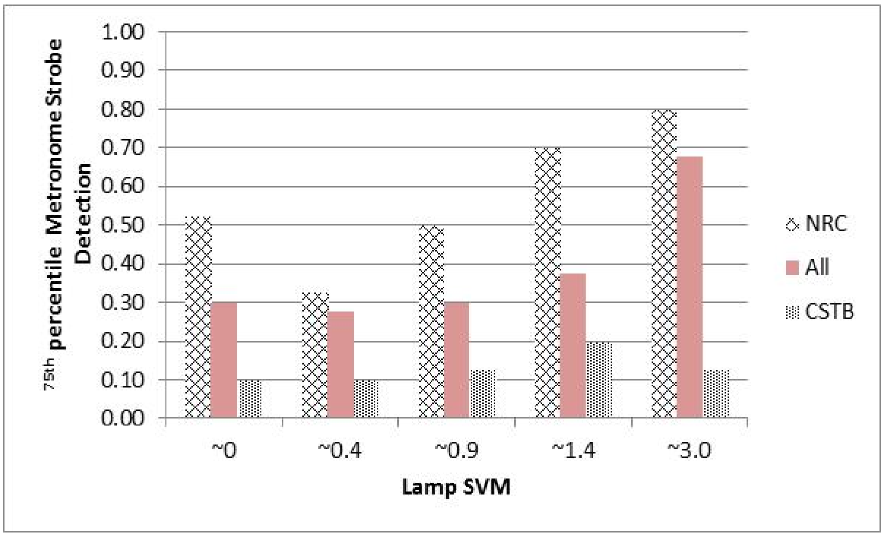


Table 4. The statistical test results (ANOVA) for the full sample and for the two sites separately do not show a clear effect of SVM on metronome stroboscopic effect detection, except at NRC. The sample distributions show how individual scores are distributed (see text).

	Means (Standard Deviations)					ANOVA Test Statistics – Linear Trend			
Condition	**1**	**2**	**3**	**4**	**5**	***F***	***df***	***p***	$\eta^2_{partial}$
SVM	~0	~0.4	~0.9	~1.4	~3.0				(effect size)
Both sites	0.17	0.14	0.16	0.21	0.30	9.04	1,35	.01	.21
Means (StDev)	(0.25)	(0.22)	(0.24)	(0.29)	(0.36)				
Sample distribution									
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	0.05	0.06	0.06	0.09	0.14	2.34	1,17	.14	.12
Means (StDev)	(0.10)	(0.10)	(0.10)	(0.15)	(0.27)				
Sample distribution									
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	0.29	0.22	0.27	0.33	0.47	6.94	1,17	.02	.29
Means (StDev)	(0.30)	(0.27)	(0.28)	(0.34)	(0.38)				
Sample distribution									
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				

2.4 Acceptability ratings

Participants rated the acceptability of each light source in terms of comfort, pleasantness, and annoyingness on the tenth block of trials. The two labs 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. Note that the questions were asked in English at NRC and in French at CSTB, the translation having been performed in-house at CSTB. There was no statistically significant difference in the ratings by SVM condition, neither overall nor by site (Figure 3 and Table 5). For Figure 3 we have shown the lowest 25% of the sample to highlight that for some individuals none of the conditions were acceptable, but most (75%) of the people had higher acceptability ratings than the bars shown.

Figure 3. This chart shows the acceptability rating for the lowest 25% of the participants, for each lamp type, for each site separately and for the whole sample combined ("All"). This is the bottom of the distribution. These values did not change regardless of the SVM. Seventy-five percent of the participants had acceptability ratings higher than the bar.

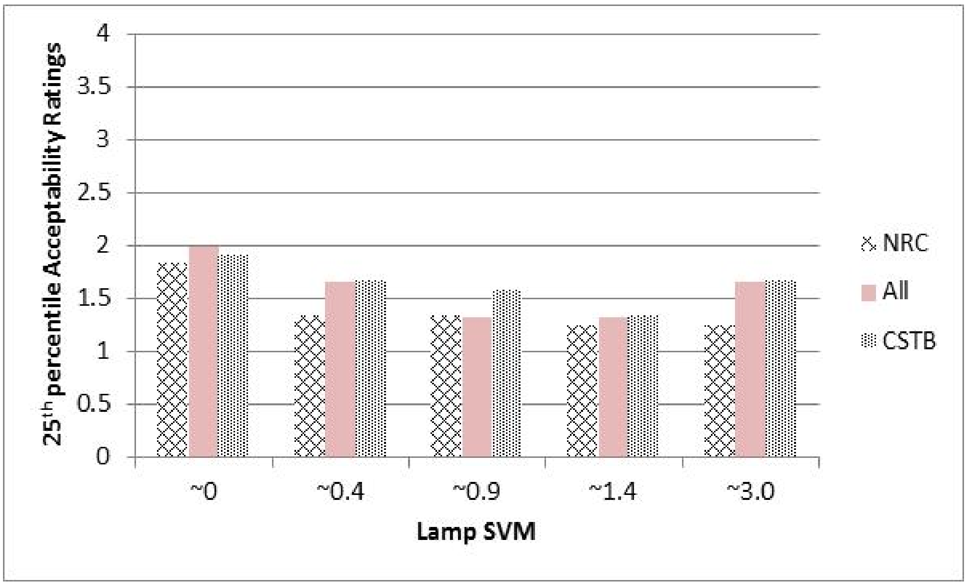


Table 5. The statistical test results (ANOVA) for the full sample and for the two sites separately do not show a clear effect of SVM on acceptability ratings. The sample distributions show how individual scores are distributed (see text).

	Means (Standard Deviations)					Test Statistics – Linear Trend			
Condition SVM	**1** ~0	**2** ~0.4	**3** ~0.9	**4** ~1.4	**5** ~3.0	***F***	***df***	***p***	$\eta^2_{partial}$ (effect size)
Both sites	2.30	2.19	2.08	2.14	2.19	0.46	1,35	.50	.01
Means (StDev)	(0.70)	(0.86)	(0.82)	(0.94)	(0.89)				
Sample distribution									
25th percentile	2.00	1.67	1.33	1.33	1.67				
50th percentile	2.33	2.00	2.33	2.33	2.00				
75th percentile	2.67	3.00	2.67	2.92	3.00				
CSTB	2.24	2.37	2.24	2.37	2.41	0.56	1,17	.46	.03
Means (StDev)	(0.65)	(0.87)	(0.83)	(1.01)	(0.90)				
Sample distribution									
25th percentile	1.92	1.67	1.58	1.33	1.67				
50th percentile	2.33	2.67	2.33	2.50	2.33				
75th percentile	2.67	3.00	2.67	3.33	3.33				
NRC	2.35	2.00	1.93	1.91	1.96	1.89	1,17	.19	.10
Means (StDev)	(0.76)	(0.83)	(0.80)	(0.82)	(0.85)				
Sample distribution									
25th percentile	1.83	1.33	1.33	1.25	1.25				
50th percentile	2.33	1.83	1.83	2.00	1.83				
75th percentile	3.00	2.75	2.42	2.42	2.67				

3. Summary and Conclusions

In the preparations for this experiment, both laboratories acquired a variety of commercially available LED replacement lamps that are available on the North American and European markets today. In the laboratory 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. There was no relationship found between price paid for the lamp and the SVM value measured. In the absence of consensus concerning the best metric with which to characterise TLM, and with no requirement for package labelling or product technical specifications to report TLM characteristics (noting that it is uncertain as to whether consumers would understand this metric if declared), there is no way for a consumer to know what the TLM performance of a product will be prior to its use.

Each laboratory selected five lamps for this experiment, based on their SVM characteristics, taking care to seek similar performance at each chosen level to permit the data to be combined. The primary visual perception task in this experiment was carefully chosen to replicate the rotating disc task with which the SVM was developed.[12, 13] 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 in a strong linear trend with increasing SVM. Moreover, the 75th percentile – the most sensitive part of the sample – detected the effect 63% of the time overall when SVM was ~0.90.

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. 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.

The following are the limitations of this experiment that were identified and 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;
- Only 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 levels, leaving gaps 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.[7] 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 on the market 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, no evidence exists.[8]

The decision to place a limit on any metric involves two choices: (1) The acceptable frequency of the outcome occurring; and (2) the acceptable proportion of the population who might experience this outcome.[7] These choices are value judgements that research can inform but cannot answer. Although this research is on-going and it is expected that further subjects will be studied using the same apparatus in early 2019 with more rigorous analysis to follow, for the purposes of the draft lighting regulation that European Member States are discussing in December 2018, the following guidance can be drawn from this work:

- An SVM>2.0 caused virtually all of the participants to perceive stroboscopic effects of the disk all of the time.
- The proposed limit of SVM=1.6 is higher than the SVM for magnetic-ballasted T12 lamps,[8] which are known to cause headaches and eyestrain and to disrupt eye movements.[1,9-11]
- The most sensitive 25% of the people detected stroboscopic effects with the disk 90% or more of the time at SVM=1.4 (75th percentile overall). The EU-28 population includes ~101 million people between the ages of 0-30. Based on the data presented here, the proposed limit of SVM=1.6 would mean that nearly all of the time, one quarter of these 101 million young people could perceive the stroboscopic effect.
- The 75th percentile detection rate dropped to 63% when the SVM was ~0.9, meaning that only one-quarter of the sample could detect the stroboscopic effect more than 63% of the

time for these lamps. A reduction of the limit value to SVM=0.9 would mean that only 25% of the population of young people (or 25.4 million people) would detect the stroboscopic effect more than 63% of the time.

- At SVM levels of 0.4 and below, the stroboscopic detection rate for the most sensitive quarter of the people dropped to 10%.

References

1. Wilkins AJ, Veitch JA, Lehman B. 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; 2010. p. 171-8.

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

3. Wilkins AJ. A physiological basis for visual discomfort: Application in lighting design. *Lighting Research and Technology*. 2016;48(1):44-54.

4. Commission Internationale de l'Eclairage (CIE). *Visual aspects of time-modulated lighting systems – Definitions and measurement models*. TN 006-2016 Vienna, Austria: CIE, 2016.

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

6. International Electrotechnical Commission (IEC). *Equipment for general lighting purposes - Objective test method for stroboscopic effects of lighting equipment*. IEC TR 63158:2018 Geneva, Switzerland: IEC, 2018.

7. Commission Internationale de l'Eclairage (CIE). *Final report CIE stakeholder workshop for temporal light modulation standards for lighting systems* CIE TN 008:2017 Vienna, Austria: CIE, 2017.

8. National Electrical Manufacturers Association (NEMA) Lighting Systems Division. *Standard for temporal light artifacts: Test methods and guidance for acceptance criteria*. NEMA 77-2017 Rosslyn, VA: NEMA, 2017.

9. Veitch JA, McColl SL. Modulation of fluorescent light: Flicker rate and light source effects on visual performance and visual comfort. *Lighting Research and Technology*. 1995;27(4):243-56.

10. Wilkins AJ. Intermittent illumination from visual display units and fluorescent lighting affects movement of the eyes across text. *Human Factors*. 1986;28(1):75-81.

11. Wilkins AJ, Nimmo-Smith I, Slater AI, Bedocs L. Fluorescent lighting, headaches and eyestrain. *Lighting Research and Technology*. 1989;21(1):11-8.

12. Vogels IM, Sekulovski D, Perz M. Visible artefacts of LEDs. *Proceedings of the 27th Session of the Commission Internationale de l'Eclairage, Sun City, South Africa*. Vienna, Austria: CIE; 2011. p. 42-51.

13. Perz M, Vogels IMLC, Sekulovski D, Wang LL, Tu Y, Heynderickx IEJ. Modeling the visibility of the stroboscopic effect occurring in temporally modulated light systems. *Lighting Research and Technology*. [Article]. 2015;47(3):281-300.

14. Wilkins AJ, Evans BJW. *I.O.O. Pattern Glare Test*. London, UK: i.O.O Sales Ltd.; 2012.

15. Brundrett GW. Human sensitivity to flicker. *Lighting Research and Technology*. 1974;6(3):127-43.

16. Evans BJW, Stevenson SJ. The Pattern Glare Test: a review and determination of normative values. *Ophthalmic and Physiological Optics*. 2008;28(4):295-309.

17. Whissell CM. Sex, eye-color, and aesthetic ratings of colors. *Perceptual and Motor Skills*. 1980;51(3, Pt 1):1012-4.

Appendix A - Method

In Canada, this research protocol was reviewed and approved by the National Research Council of Canada Research Ethics Board (Protocol 2018-129) and by the Carleton University Research Ethics Board-B (CUREB-B Clearance # 109982).

This research complies with the EU and French General Data Protection Regulation (GDPR).

A.1 Research Design and Hypothesis

This is a repeated-measures experiment with 5 levels of the independent variable light source. The five light sources were commercially available products chosen because they are known to exhibit the chosen levels of the metric SVM (International Electrotechnical Commission (IEC), 2018; Perz, et al., 2015). They were chosen based on the following criteria: about 800 lm output; ~2600-3000 K correlated colour temperature; P_{st}^{LM} << 1, and having the target SVM values (0, 0.4-0.6; 1; 1.6; and >2). These values were chosen to provide a range of data for conditions at and below the proposed regulatory limit being considered in Europe (SVM<1.6) (National Electrical Manufacturers Association (NEMA) Lighting Systems Division, 2017), and one condition considerably higher as a manipulation check. To the extent possible, chromaticity coordinates were matched, although this was dependent on the existence of alternatives at any given SVM value.

Participants performed 50 trials in total, 10 blocks of five light source conditions. Within each block the order of presentation was random.

Data were collected in parallel using this same procedure by teams in Canada (NRC) and in France (CSTB). The results presented in this report are based on a combination of the data in a mixed design for analysis (site being a between-groups variable, and light source a repeated measures variable).

The hypotheses to be tested were:

> H1: The visibility of the stroboscopic effect is 50% for SVM=1. (This is the definition of the metric.)
>
> H2: The visibility of the stroboscopic effect increases with increasing SVM.
>
> H3: Comfort and pleasantness should drop with increasing SVM, and annoyingness should increase with increasing SVM.

A.2 *Setting and Lighting Conditions in Canada and France*

Testing occurred in a dedicated small windowless room. The room was minimally furnished with a desk and chairs. During the session trials all the illumination will come from the custom desktop luminaire 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 (see photos).

The custom luminaire consisted of an aluminum frame supporting a light box. The central light box had six chambers, in each of which was a standard E27 socket. Five locations were used, each with one of the five lamps (described below). All lamps were powered continuously during the session to maintain constant temperature and light output.

At NRC, the light box drum could be rotated such that one chamber was located over an aperture that allowed light to fall onto the desk surface, while the other lamps were blocked by the plywood base of the light box. The chambers in the light box were painted black (NRC) or covered with black adhesive velvet sheets, but some were modified with white reflective plastic to increase the illuminance on the desk below when that chamber was in use. There was no diffuser over the lamps, but the participant was shielded from any view of the aperture by a cover on the frame. At CSTB, each chamber of the light box had an individual shutter. The selection of the chamber was also done by rotating the device.

The height of the luminaire was adjusted once so that all the lamps delivered ~330 lx on the surface of the principal task, a rotating disk (see below). There were no sides to the frame, so that viewers had the full field of view available to them and light from the luminaire could provide ambient illumination for the rest of the room beyond the desk on which it sits. An uninterruptible power supply was used by NRC to ensure clean power for the luminaire during testing. A laboratory specification AC power supply was used by CSTB for the same purpose.

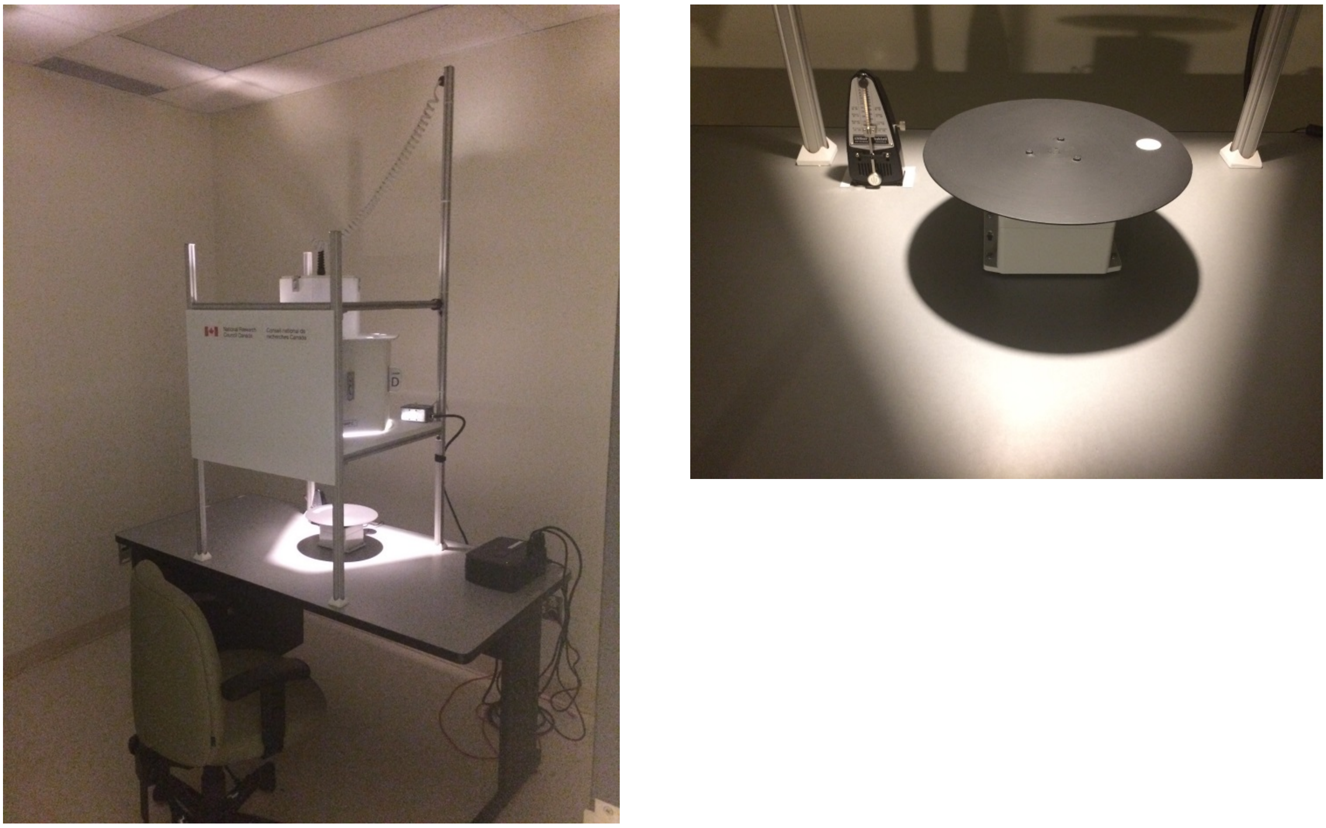


Figure A1. Images of the NRC apparatus installed in the test room. For the image on the left, the hallway outside provided fill light for the photo, but the door was closed during testing. The image on the right shows the desk surface as seen by the participant.

Figure A2. Images of the CSTB apparatus installed in the test room.

Table A1. NRC test lamp characteristics claimed CCT and luminous flux, and light characteristics measured horizontally at the location of the rotating disk under the light box for each of the five lighting conditions.

Condition	CCT (label) [K]	Lum. flux (claim) [lm]	Illum (meas) [lx]	CCT (meas) [K]	R_a (meas)	Duv (meas)	Dominant Frequency [Hz]	Modu -lation [%]	Flicker Index [%]	Pst^{LM}	SVM
1	2700	800	341	2872	83	-0.0008	120	4.7	0.43	0.05	0.04
2	2700	800	319	3018	84	-0.0016	120	14.0	3.79	0.07	0.42
3	2700	800	354	2717	83	-0.0001	120	32.0	8.47	0.08	0.91
4	3000	800	334	3094	83	-0.0023	120	55.6	13.25	0.06	1.38
5	3000	800	335	3027	83	-0.0003	120	91.5	29.99	0.33	2.80

Table A2. CSTB test lamp characteristics claimed CCT and luminous flux, and light characteristics as measured horizontally at the location of the rotating disk under the light box for each of the five lighting conditions.											
Condition	CCT (label) [K]	Lum. flux (claim) [lm]	Illum (meas) [lx]	CCT (meas) [K]	R_a (meas)	Duv (meas)	Dominant Frequency [Hz]	Modu-lation [%]	Flicker Index [%]	Pst^{LM}	SVM
1	2700	806	344	2756	83	0.00001	100	2.1	0.6	0.39	0.00
2	2700	810	330	2810	82	0.0009	100	11.8	3.7	0.05	0.43
3	2600	720	318	2559	90	0.0022	100	27.8	7.9	0.08	0.96
4	2700	810	312	2641	81	0.0016	100	40.2	12.3	0.26	1.47
5	2700	600	324	2799	80	0.0022	100	79.4	26.9	0.38	3.09

A.3 Dependent Variables

The same questions and tasks were used in both countries. They were originally written in English and translated to French at CSTB.

A.3.1 Demographics

At the start of the session (after signing the Agreement to Participate), participants answered demographic questions with paper and pen on the desk surface, under the lowest SVM condition. These questions are shown in Table A2.

Table A2. Demographic questions.						
What is your sex?	0 Male	1 Female	2 Prefer not to say			
How old are you?	0 18 to 29	1 30 to 39	2 0 to 49	3 50 to 59	4 60 and older	
What type of correction lenses do you have today?	0 None	1 Reading Glasses	2 Distance Glasses	3 Bi- or Trifocal Lenses	4 Gradual or Multi-focal Lenses	5 Contact Lenses
Have you ever been diagnosed with a hearing impairment?	0 No	1 Yes				
Do you use any form of hearing aid?	0 No	1 Yes				
What best describes your eye colour? We ask because there is evidence that eye colour predicts some visual perceptions although the mechanism for this is unknown.	0 Blue	1 Grey	2 Green	3 Light brown	4 Brown or black	
What is the highest level of education you have completed?	Secondary/ high school graduation certificate or less = 1	Diploma or certificate from a community college, institute of technology etc. = 2	Some university courses or a university certificate below the Bachelor level = 3	Undergraduate (Bachelor's) degree = 4	Graduate or professional degree = 5	

A.3.2 Pattern glare sensitivity

Participants were asked to complete the Wilkins and Evans Pattern Glare Test.[14] The test consists of three plates of square wave patterns, shown below in reduced size (Figure D2). Scores on this test have been shown to be correlated to the propensity for headache associated with visual stimuli. Normative scores have been established,[16] to which scores were compared. 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.) This was used as an individual difference variable to examine variation in response to TLM.

Figure D2: Pattern Glare Test

You will be asked to look at a striped pattern for 5 seconds. Please focus on the white square in the centre of the pattern.
If you find the pattern extremely uncomfortable to view, please avert your eyes until the pattern has been removed.
Did you experience the following when looking at the stripes?

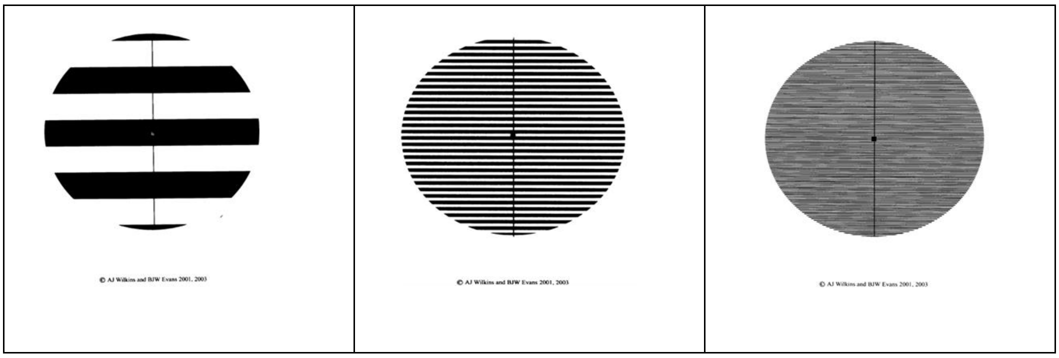
- Colours (if yes, which? Red / Blue / Green / Yellow)
- Bending of any lines
- Blur
- Shadowy shapes amongst the lines
- Flickering / Shimmering of the lines
- Fading of the lines
- Other
- None

Were any of the above

- Predominantly on the left of the image
- Predominately on the right of the image
- Roughly equal both sides

How uncomfortable was the pattern to look at?

- No problem
- Slightly uncomfortable
- Uncomfortable
- Very uncomfortable



A.3.3 Stroboscopic effect

Within each trial 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.[2, 13] 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) (see Figure A3). 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.

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. See figure A1 for the metronome used in Canada and figure A2 for France. The metronomes were identical in the two countries, but each team made and attached its own black dot.

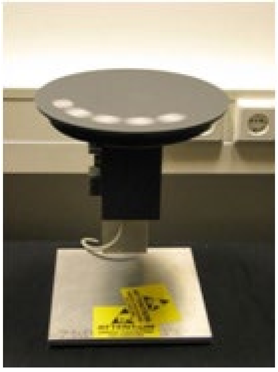


Figure A3. These images from Reference 12 demonstrate the stroboscopic effect. When one sees this effect, the moving disk looks like the image on the left. With no stroboscopic effect, it looks like the disk on the right.

A.3.4 Judgements of light sources.

On the last trial for each lamp (i.e., after having repeated the visibility task 10 times), they also were asked to rate the comfort, pleasantness, and annoyingness of that condition, each on a 5-point Likert scale[12]:

0	1	2	3	4
not at all	a little	moderately	very much	extremely

A.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[17], 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 will be 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.

Participation took approximately 50 minutes. 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.