



Article

# Differences in Contrast Reproduction between Electronic Devices for Visual Assessment: Clinical Implications

Ainhoa Molina-Martín <sup>1,\*</sup>, David P. Piñero <sup>1</sup>, María B. Coco-Martín <sup>2</sup>, Luis Leal-Vega <sup>2</sup> and Dolores de Fez <sup>1</sup>

<sup>1</sup> Group of Optics and Visual Perception, Department of Optics, Pharmacology and Anatomy, University of Alicante, 03690 Alicante, Spain; david.pinyero@gcloud.ua.es (D.P.P.); dolores.fez@ua.es (D.d.F.)

<sup>2</sup> Group of Applied Clinical Neurosciences and Advanced Data Analysis, Department of Medicine, Dermatology and Toxicology, University of Valladolid, 47002 Valladolid, Spain; mbegococom@gmail.com (M.B.C.-M.); luis.leal.vega.1213@gmail.com (L.L.-V.)

\* Correspondence: ainhoa.molina@ua.es

**Abstract:** The easy access to electronic devices for users has resulted in the development of a vast range of programs and applications for visual evaluation and diagnosis that can be downloaded to any device. Some of them are based on tasks and stimuli that depend on luminance. The aim of the present study was to evaluate differences in luminance reproduction between electronic devices and their implications for contrast reproduction. A total of 20 Galaxy Tab A devices with 8-bit graphics processing units were evaluated. Characterization of every screen was performed obtaining the response curve for the achromatic stimulus. Mean, maximum and minimum luminance, standard deviation and coefficient of variation were obtained to assess differences between devices. Variation of luminance with increasing digital level was observed in all devices following a gamma distribution. Comparison between devices for mean results showed that some of them differed by as much as 45 cd/m<sup>2</sup>. The coefficient of variation varied from ~5 to 9%. Mean percentage of differences in luminance between devices reached 30%. In conclusion, differences in luminance reproduction between devices were present, even considering devices from the same manufacturing batch. It cannot be assumed that the characterization of one device can be extrapolated to other devices. Every device used for research purposes should be individually characterized to ensure the correct reproduction. For clinical purposes, limitations should be considered by visual specialists.

**Keywords:** contrast sensitivity; luminance reproduction; luminance characterization; display; tablet; electronic device



**Citation:** Molina-Martín, A.; Piñero, D.P.; Coco-Martín, M.B.; Leal-Vega, L.; de Fez, D. Differences in Contrast Reproduction between Electronic Devices for Visual Assessment: Clinical Implications. *Technologies* **2021**, *9*, 68. <https://doi.org/10.3390/technologies9030068>

Academic Editor: Ulas Sunar

Received: 26 July 2021

Accepted: 13 September 2021

Published: 15 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, there has been a technological transition in our clinics from the more conventional printable charts and optotypes to electronic devices. These devices allow for higher versatility on patients' visual exams, unifying in the same instrument different tests and procedures while saving costs and space [1]. From a clinical perspective, compared to conventional charts and printable content, screens give us the chance to create dynamic and more complex stimuli but also a better control over the exam conditions. Since they have their own illumination, the images and reproduced stimuli will not be so affected by the environmental conditions, especially if measurements are taken in a darkened room where the device is the only source of light [2].

Handheld electronic devices are becoming increasingly popular in clinical practice as diagnostic tools for specialists due to their portability and a greater capacity to adapt visual exams to different contexts, such as subjects with special needs or reduced mobility [3–9]. The use of these devices has been also proposed for the remote monitoring of the patients by vision specialists, which in the context of COVID-19 could lead to a notable reduction of risks [10]. Patients now can have access to a device controlled by the vision specialist, and many possibilities in visual training have been created for home therapies [11]. These

advantages and the easy access to this technology for users has resulted in the development of a vast range of programs and applications for visual examination that can be downloaded to any device with the purpose of diagnosis or patient monitoring by the specialists as well as self-monitoring or training by the patients [12]. These remote applications are supposed to be more variable than clinical assessment in a controlled environment, but most of the variables can be reduced if specific instructions are provided to the final user.

There is a wide variety of displays used for medical purposes, such as diagnostic tools for physicians, referred to as primary displays, that follow the most stringent requirements to guarantee adequate performance, since diagnosis or treatment can be influenced by the correct reproduction of a given image or stimulus [13,14]. In the case of programs that can be downloaded to different devices, however, do we really know what they are reproducing? The type of display or the characteristics of the graphics card can widely alter the reproduction of a given stimulus, even when environmental and ergonomic conditions are under control [15,16]. Differences in reproduction should be taken into consideration by the vision specialists as a potential clinical limitation but also by the manufacturers and designers as a technological challenge [17].

When using electronic devices to evaluate the visual function, and specifically psychophysical measurements based on luminance changes, two important issues must be considered by clinicians: the luminance of that particular device, since it will determine the adaptation of the subject and the main mechanism that mediates the response, and the variations of luminance, since they will determine the incremental threshold and therefore the ability of the eye to differentiate between stimuli of different luminance [18,19]. The RGB space of electronic devices is based on the level of switching of its primaries (digital to analog converter or DAC values). Changes in DAC values do provide proportional changes in luminance, and it is necessary to know the relation between these two spaces to provide a correct stimulus reproduction. Characterization is an essential step before using a specific device to perform visual task examinations, even more so when these tasks are dependent on luminance or chromatic changes, as occurs with psychophysical measurements like contrast sensitivity [7,8] or color vision deficiency tests [9]. The process of characterization allows us to know how a specific stimulus is reproduced in a specific electronic device, but there is an additional problem that must be considered by designers: the cross-reproduction errors. Once it is known that one device correctly reproduces a specific stimulus, could it be extrapolated to other devices? The same designed stimulus could lead to a different real stimulus in other devices.

The problem of reproduction has been widely studied by engineers and physicians [20–24], and many methods have been proposed for screen characterization [25–28]. The problem of cross-reproduction between different devices has been also studied by many authors [15,20,22]. Even comparison between devices with the same type of screen has demonstrated that the reproduction of an achromatic or a chromatic stimulus in different devices cannot be simply performed [16,29]. The aim of the present study was to evaluate differences in luminance reproduction between devices from the same model, year of fabrication and manufacturing batch to determine if a global characterization based on a generic device is a feasible solution to ensure the correct cross-reproduction for an achromatic stimulus, or otherwise, individual luminance characterization should be performed on each device. Specifically, implications for reproduction of contrast sensitivity measurements are discussed.

## 2. Materials and Methods

The process of characterization should consider the spatial uniformity and the temporal stabilization before beginning the colorimetric calibration of the device. Once these factors were studied and controlled, the colorimeter was used to measure the screen luminance for a predefined set of DAC values normalized from 0 to 1.

## 2.1. Devices

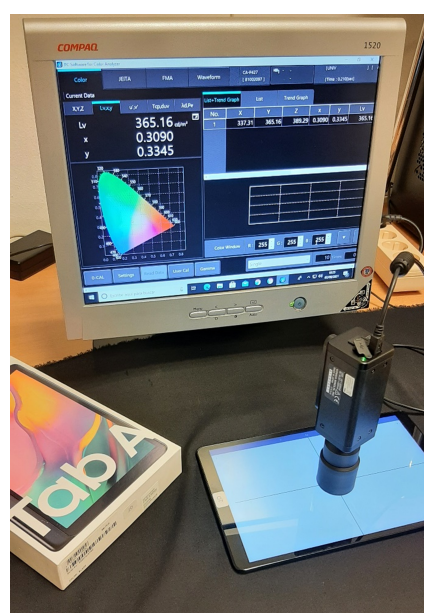
Twenty Samsung Galaxy Tab A (SM-T510, version 2019) devices (Samsung Electronics Co Ltd., Seoul, Korea) with the Android operating system were evaluated. The size of the screen was 10.1" (255.4 mm), and the resolution was  $1920 \times 1200$  pixels. A thin film transistor-liquid crystal display is used to provide up to 16 million colors with an 8-bit graphics processing unit. All devices were manufactured in 2019 and belonged to the same manufacturing batch, with correlative serial numbers.

## 2.2. Measurements

Every device was set to the highest screen luminance. Automatic brightness adjustment and screensavers were disabled to avoid changes during examination. The devices were plugged into an outlet throughout to avoid possible automatic adjustments in brightness as a function of battery charge level. Measurements of luminance in  $\text{cd}/\text{m}^2$  were obtained experimentally by using the CA-P427 Display Color Analyzer with the CA-S40 software (Konica Minolta, INC., Tokyo, Japan). This probe performs an automatic zero-point calibration and has a precision of 1.5% for luminance values over  $0.1 \text{ cd}/\text{m}^2$ . Measurements were obtained in tristimulus values (CIEXYZ 1931), and the second value (Y) was considered for luminance results. Three consecutive measurements were obtained, and the average was considered for the analysis. All measurements were performed in a darkened room.

### 2.2.1. Spatial Location

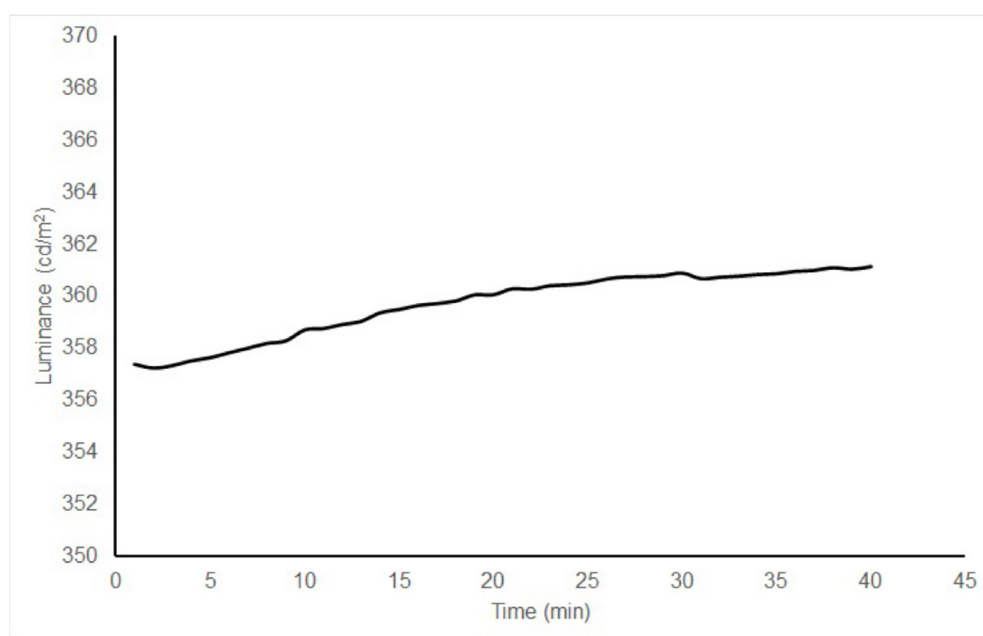
The matrix of pixels from the screen is manufactured with a uniform distribution over space, but due to fabrication processes and individual differences in the components, it has been demonstrated that the distribution of luminance on the screen is not spatially uniform [17,24,30]. This heterogeneity will affect the luminance results if the same region of the screen is not evaluated in every measurement. To avoid variability in luminance results due to the spatial position of the photodetector, the same region of the screen was evaluated in every device using a reference stimulus with a central cross. The colorimeter was attached to the center of the screen perpendicularly to avoid variations in luminance due to the viewing angle [17,24,31]. The measurement procedure followed is represented in Figure 1.



**Figure 1.** Measurement procedure. The screen of the tablet shows the reference stimulus for the colorimeter positioning. All measurements were performed in a darkened room.

### 2.2.2. Temporal Stabilization

Luminance of electronic displays depends on the intensity level of the matrix of pixels, that is the number of DAC values that the graphics card can reproduce, but also on the time they were emitting light. When switching on the display, there is a warming up period, and after a few minutes we can consider that the display to be stabilized [17,20]. Temporal stabilization was achieved in one randomly selected tablet measuring the luminance of the screen every minute during a period of 40 min, and results are represented in Figure 2. The image shown on the screen for temporal stabilization purposes was a uniform achromatic stimulus with the maximum DAC value in the three primaries ( $R = G = B = 1$ ). Twenty minutes after switching on the device, the variation in luminance over time was less than  $1 \text{ cd/m}^2$  (that is a percentage of variation of 0.3% in luminance), so this period was established as necessary to provide stable results. Screens of every tablet were then turned on at least 20 min before the study measurements.



**Figure 2.** Luminance response curve over time for a randomly selected tablet. Luminance increases more rapidly in the initial minutes and reaches a variation of less than  $1 \text{ cd/m}^2$  20 min after switching on the device.

### 2.2.3. Characterization Method

Different mathematical methods have been proposed to carry out the process of characterization [25–28]. One of the most popular methods (developed for cathode ray tube displays) is “gain–offset–gamma” which has demonstrated a perfect mathematical adjustment (potential adjustment) of the results for CRT displays [25–27], whereas other types such as TFT or LCD followed a sigmoid adjustment [22]. In the present study, the gain–offset–gamma method was used for characterization.

### 2.2.4. Achromatic Characterization

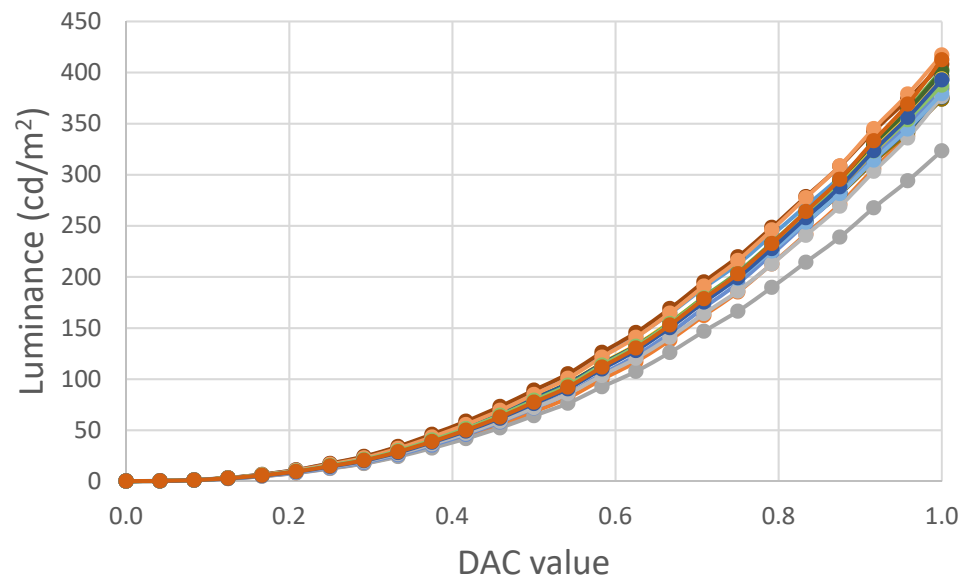
The problem of screen reproduction is critical when using the devices for luminance-based measurements but is even more critical when analyzing the response to a chromatic stimulus as demonstrated in previous studies [15,16]. In this study, we focused only on luminance characterization of the devices, and therefore achromatic stimuli were measured. For study purposes, a discrete set of stimuli (uniform images) was created with 25 different DAC values in constant steps from 0 to 1 and the same level for each primary color ( $R = G = B$ ).

### 2.3. Statistical Analysis

Statistical analysis was performed using the SPSS program v.19.0.0 (SPSS Inc., Chicago, IL, USA) and MATLAB R2019b (Mathworks, Inc., Natick, MA, USA) for Windows. Analysis of the data samples was evaluated by means of the Shapiro–Wilk method, and accordingly parametrical or non-parametrical statistical tests were applied [32]. Mean, maximum and minimum luminance results for each device by DAC values were used as outcomes. The residuals, which represent the error between the real reproduced stimulus and the target stimulus due to the adjustment, were also obtained to assess reproducibility. In addition, standard deviation and coefficient of variation were obtained to assess differences between devices.

### 3. Results

Luminance measurements from each device for the evaluated DAC values are represented in Figure 3. As can be seen in this figure, a variation of luminance with increasing DAC values was observed in all devices, and this variation followed a gamma function  $L = \alpha * DAC^\gamma$ . Mean luminance, standard deviation, coefficient of variation and the residuals for each DAC value are represented in Table 1. Luminance was found to be normally distributed until DAC value = 0.63 (mean luminance of approximately 130  $cd/m^2$ ); therefore, all data were analyzed by means of non-parametrical statistical tests.



**Figure 3.** Luminance measurements from each device for the evaluated DAC values. A potential curve fit between DAC values and luminance  $L = \alpha * DAC^\gamma$  was obtained in all cases.

**Table 1.** Mean luminance (L mean), standard deviation (SD), coefficient of variation (CV) and the residuals for each DAC value.

DAC	L Mean ( $cd/m^2$ )	SD	CV (%)	Residuals ( $cd/m^2$ ) [min, max]
0.00	0.33	0.02	7.33	0.33 [0.27, 0.37]
0.04	0.53	0.04	7.04	0.29 [0.20, 0.34]
0.08	1.33	0.10	7.35	0.11 [−0.11, 0.26]
0.13	3.2	0.3	7.88	0.01 [−0.51, 0.40]
0.17	6.0	0.5	8.28	−0.10 [−1.12, 0.71]
0.21	9.9	0.9	8.60	−0.38 [−2.10, 1.03]
0.25	15.3	1.4	8.80	−0.36 [−3.06, 1.97]
0.29	21.1	1.8	8.85	−1.25 [−5.03, 2.06]
0.33	30	3	8.78	−0.98 [−6.28, 3.69]
0.38	40	3	8.59	−0.28 [−7.46, 5.96]

Table 1. Cont.

DAC	L Mean (cd/m <sup>2</sup> )	SD	CV (%)	Residuals (cd/m <sup>2</sup> ) [min, max]
0.42	51	4	8.32	−0.13 [−9.36, 7.62]
0.46	64	5	7.99	0.02 [−11.50, 9.36]
0.50	78	6	7.66	0.18 [−13.87, 11.06]
0.54	93	6	7.32	−1.09 [−17.61, 11.15]
0.58	112	8	6.95	0.56 [−19.27, 14.40]
0.63	130	9	6.67	−0.71 [−23.55, 14.44]
0.67	153	10	6.37	0.25 [−26.22, 16.75]
0.71	177	11	6.10	2.00 [−28.48, 19.82]
0.75	201	12	5.90	0.64 [−33.71, 19.45]
0.79	229	13	5.73	1.86 [−37.16, 21.49]
0.83	258	14	5.58	2.55 [−41.20, 22.89]
0.88	287	16	5.47	1.17 [−47.36, 22.40]
0.92	322	17	5.37	2.78 [−51.22, 26.59]
0.96	353	19	5.33	−0.14 [−59.45, 26.36]
1.00	390	21	5.36	−0.75 [−66.89, 29.05]

Although all the devices followed a gamma function, there were variations in  $\alpha$  (maximum luminance) and  $\gamma$  (gamma exponent) parameters depending on the device (Table 2). A mean value of 394.19 cd/m<sup>2</sup> (from 328.77 to 423.78) was found for  $\alpha$  and 2.34 (from 2.22 to 2.46) for  $\gamma$ . A coefficient of regression (r-squared) >0.99 was obtained in all cases along with a root mean square error (RMSE) of 0.86.

Table 2. Obtained  $\alpha$  and  $\gamma$  values for each device.

Tablet	$\alpha$	$\gamma$
1	400.386	2.370
2	376.357	2.462
3	328.774	2.360
4	390.069	2.271
5	402.086	2.217
6	403.001	2.352
7	402.569	2.372
8	417.678	2.237
9	423.778	2.351
10	381.056	2.242
11	383.146	2.249
12	401.442	2.369
13	392.583	2.454
14	422.536	2.322
15	371.392	2.379
16	398.357	2.362
17	384.525	2.300
18	396.740	2.327
19	398.209	2.397
20	409.112	2.412

The comparison between the devices for mean results showed that some of them have assumable variations in maximum luminance for a specific DAC value of less than 1 cd/m<sup>2</sup>, whereas other devices differed by as much as 45 cd/m<sup>2</sup>. Specifically, tablet number 3 (corresponding to the lower gray fit from Figure 3) showed a significant decrease in luminance performance in comparison to the other devices. Standard deviation of the measurements increased with the DAC value, and differences were even to the extent of  $\pm 21$  cd/m<sup>2</sup> (for the maximum luminance values). Coefficient of variation varied from ~5 to 9% and was lower for high luminance values. Excluding device number 3 from the analysis, the maximum standard deviation value was reduced to  $\pm 14$  cd/m<sup>2</sup>, and coefficient of

variation varied from ~4 to 8%. Mean percentage differences in luminance between devices reached 30% (22% excluding device number 3).

#### 4. Discussion

The use of clinical tests on electronic devices in clinical practice leads to a need to attend to aspects that are irrelevant to conventional tests such as the type of screen, its color reproducibility or the use of the test on different devices. These aspects are critical when evaluating a subject using a particular device but even more so when measuring different subjects with different devices and trying to compare these results. Since the COVID-19 situation has forced more remote interaction with patients, these procedures are increasing in utility in clinical practice for diagnostic purposes as well as training solutions for home-based therapies. Some of the tests are based on more simple stimuli such as optotypes, where the spatial resolution of the screen is considered, but others are based on more complex stimuli for which variables such as luminance have to be considered.

Quality analyses performed by the manufacturers are based on spatial characteristics and light emission of the final product, and although these analyses allow the detection of manufacturing defects, they cannot guarantee a minimum performance for visual examination purposes. These analyses are normally performed in some isolated devices to ensure a correct production process, but defects may pass undetected, as probably occurred in the case of device number 3 from our study. Additionally, a range of luminance can be considered as acceptable for the quality analyses performed by the manufacturers, but this variability may not be acceptable when evaluating the interaction of these devices with vision. The human eye has a photodetection system in photopic conditions, the cones, which allow for detection of small luminance increments based on Webber's fraction, that is approximately  $1 \text{ cd/m}^2$  for luminance values of  $100 \text{ cd/m}^2$  [18]. In the present study, considering an average maximum luminance of  $\sim 400 \text{ cd/m}^2$ , changes of  $4 \text{ cd/m}^2$  in luminance reproduction errors can be considered as clinically significant.

##### 4.1. Reproduction Errors

When designing a visual test in an electronic device, it cannot be assumed that changes in DAC values will produce proportional changes in the luminance and chromatic characteristics of the stimulus. Higher DAC values will provide higher luminance and lower DAC values lower luminance, but the mean value in DAC will not provide the mean luminance. This is since the relation between DAC values and luminance is not linear, and specifically, in the case of the model Samsung Galaxy Tab A, followed a gamma function: A change in DAC values for lower luminance is smaller than for higher luminance. In other words, if the same increment of luminance is intended, it does not necessarily lead to the same variation in DAC values depending on dark or light colors. For this reason, variations in visual stimuli should be performed in a language understandable by the visual system before translating these variations to the language of the device to program the instructions. This process only can be performed if the method of this translation is previously known, that is, if the equivalence between DAC values and luminance (in the case of an achromatic stimulus), which is obtained with the process of characterization of the device.

Regarding the characterization, in the present study, the reproduction of the model Samsung Galaxy Tab A was evaluated by measuring a discrete set of 25 stimuli with different DAC values (previously and specifically designed) and obtaining the luminance curve response. The gain–offset–gamma methodology has some limitations, since it is not as accurate as the look-up tables method used previously [20]. The use of this method assumes the constancy of chromaticity of the primaries (that is each RGB channel increasing in luminance without changing the color), the independency of the three channels (that is varying luminance or color in one channel has no effect on the others) and the additivity law of luminance (that is luminance from the stimulus can be directly summed). These assumptions cannot be made for all types of displays and are the cause of the lack of precision between the real reproduced color and the color predicted by the mathematical method [33]. Addi-

tionally, the discrete set of stimuli might not be sufficient to achieve a perfect adjustment. In the present study, results showed a curve adjustment  $>0.99$  in all devices. Moreover, residuals from Table 1, which represent the error between the real reproduced stimulus and the target stimulus due to the adjustment, were of less than  $1 \text{ cd/m}^2$  in most cases (maximum  $2.78 \text{ cd/m}^2$ ). These results suggest that the reproducibility errors due to the gain–offset–gamma methodology are not clinically relevant. Alternatively, an advantage of choosing this methodology was the time saved in the process of characterization.

In this work, the luminance reproduction of 20 devices model Samsung Tab A, an economic low-medium range device, was studied. In the past, de Fez et al. [15] compared the reproduction of four devices and found that the high range devices provide better reproducibility results than the lower range ones. Compared to our results, higher range devices seemed to provide better reproducibility results [29], probably due to the higher resolution and quality of the screen, but this is not a guarantee. It is not possible to affirm that other high range devices will provide more acceptable results than low-medium range devices, and future studies evaluating different devices should be conducted to answer this question.

#### 4.2. Cross-Reproduction Errors

When visual stimuli are created for one specific device, the equivalent luminance will differ when translating this design to other devices. As example, the mean DAC value of 0.5 will not reproduce the same luminance in a device with a maximum luminance of  $100 \text{ cd/m}^2$  as a device emitting  $400 \text{ cd/m}^2$ , even considering the same gamma curve. The detection threshold of one patient in one device may not be the threshold in another. Likewise, the colors from the confusion lines in an anomalous subject could be sufficiently different in another device to produce a correct answer.

Comparisons between different devices with different screen characteristics have been previously reported [15,20,22], including even comparisons between devices with the same screen (with the same characteristics) [16,29]. In the present study, not only were the compared devices the same, but they were also from the same year of production (2019) and the same fabrication batch (with correlative serial numbers). This ensures that all components were manufactured in the same production process but does not guarantee that the individual components are. Since display quality depends on the quality of the components, and the origin of these components cannot be tracked, we have to assume that manufacturing differences in display characteristics will always be present.

Manufacturers supply calibrated devices, assuming that the changes in luminance with the DAC values are constant, that is following a linear progression, or at best, following a potential progression but assuming a standard gamma exponent (for example,  $\gamma = 2$  in a cathode ray tube or  $\gamma = 1.8$  in Apple devices) [34]. It could be assumed that although the problem of reproduction is present, differences in devices from the same model will be anecdotal, and once we know how a model of a device works, we can extrapolate this performance to other devices of the same model. However, this is not the case as demonstrated by the results of the present paper. Even considering devices from the same manufacturing batch with correlative serial numbers, the variations in  $\alpha$  and  $\gamma$  and consequently the luminance performance were different between devices.

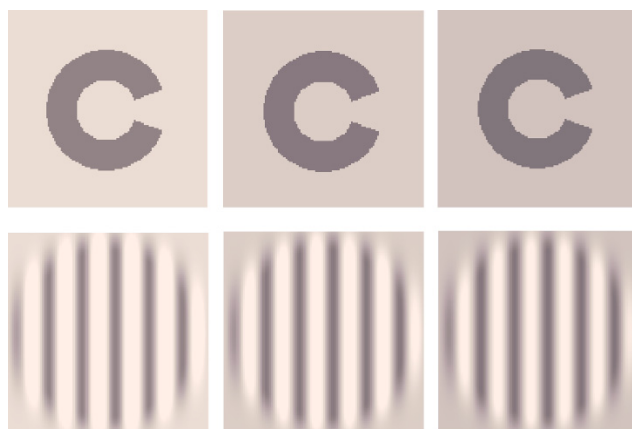
In this study, we focused only on the luminance description of the devices, but previous papers demonstrated that this problem is present also for chromatic representations. De Fez et al. [16] compared in the past results obtained from three iPads of the same model but different years of manufacturing and demonstrated that although the color reproduction errors on each screen were in the range of the minimum appreciable difference by the visual system (1 CIEDE2000 unit), the colorimetric design valid for a given device may not be correct when displayed in another device, showing color differences between 4 and 6 units [16].



#### 4.3. Implications for Contrast Sensitivity

Differences found in luminance reproduction will directly affect contrast sensitivity measurements since contrast is determined by the maximum and minimum luminance of the stimulus. Even if the same model is evaluated, differences between devices have shown a high variability in the maximum reproducible luminance  $\alpha$  and the exponent  $\gamma$ , and this implies that even when considering the same subject, results obtained for different devices will probably differ.

Apart from the reproduction errors of the device, when using this device to create a stimulus to be reproduced in other devices, the cross-reproduction errors should be evaluated. Considering as an example a Pelli–Robson test, that is a stimulus with luminance in the background (maximum luminance) and luminance in the letters (minimum luminance), the reproduction for every device will affect results. If we consider tablet number 9 (which is the device with the highest luminance values of our sample) to reproduce a contrast of 0.50 (with a mean luminance of 85 cd/m<sup>2</sup> as recommended in the literature), a maximum luminance of 127.5 cd/m<sup>2</sup> in the background would be needed and a minimum of 42.5 cd/m<sup>2</sup> in the letters (DAC values for this device are 0.60 and 0.38, respectively). Using these same DAC values to create the same stimulus in other devices, such as tablet number 15 (which is the device with the lowest luminance values and similar gamma exponent values of our sample), the contrast reproduced would be 0.51. This change in contrast would not be clinically significant (target 0.50), but looking carefully at the luminance of the stimulus, the mean values would change from 85 to 73.4 cd/m<sup>2</sup>, the maximum luminance values would change from 127.5 to 110.0 cd/m<sup>2</sup> and the minimum luminance values would change from 42.5 to 36.2 cd/m<sup>2</sup>. On the other hand, considering the specific case of Tab number 3, for the same considered DAC values, a contrast of 0.49 would be obtained (again similar to the original target contrast of 0.50), but a maximum luminance of 98.5 cd/m<sup>2</sup> and a minimum luminance of 33.5 cd/m<sup>2</sup> would be obtained. This implies that even reproducing the same contrast between devices, the luminance of the stimulus and therefore the adaptive status of the subject will differ [18]. A graphical example of the implications of these variations in luminance in the reproduction of a stimulus can be observed in Figure 4. In all cases, the DAC values of the stimulus are the same, but not the final performance of the stimulus depending on the device used to reproduce it.



**Figure 4.** Graphic representation of a stimulus in three different devices with same DAC values. Note that the final performance of the stimulus varies with the device. These images were created for a characterized screen, and their performance is dependent on screen characteristics.

The same calculations can be performed when analyzing not only the maximum luminance  $\alpha$  but also the variations in the gamma exponent  $\gamma$ . In the example of tablet number 9 (with a gamma exponent of 2.35), if a stimulus for that tablet is designed to be reproduced in a tablet with different  $\alpha$  and  $\gamma$ , for example tablet number 5 (with a gamma

exponent of 2.22), results would also differ: The mean luminance changes to values of  $87.5 \text{ cd/m}^2$  (reference value  $85 \text{ cd/m}^2$ ) and contrast values to 0.48 (reference value 0.50).

Bodduluri et al. [29]. also studied the colorimetric characterization of 15 iPad mini retina devices to evaluate if results derived from one device could be extrapolated to others. These authors conclude that colorimetric characterization of one device can be used in the other devices as no significant variation in gamma function was found between devices (gamma exponent from 2.12 to 2.21). On the other hand, results from luminance measurements reveal that some of the devices reproduced values up to  $439 \text{ cd/m}^2$ , whereas others showed results of  $360 \text{ cd/m}^2$  [29]. These differences represent approximately a 20% of variation in total luminance between tablets, which is in our opinion a considerable percentage. In the present study, similar maximum luminance variability was found between devices compared to these authors, and also in the gamma exponent. It is true that the proper variation in luminance between tablets does not mean that the contrast is going to be wrongly reproduced, but it will affect the adaptive status of the subject, and therefore their response. As shown in the previous example, variations in gamma exponent between devices mainly affect the reproduced contrast between devices, while variations in mean luminance of the stimulus affect the adaptation status.

#### 4.4. Possible Solutions

Even if the same model is evaluated, differences between devices have shown a variability in the mean luminance and contrast reproduction, and this implies that comparison of contrast sensitivity results from different patients with different devices cannot be simply performed. The correct solution for avoiding reproduction and cross-reproduction errors is to characterize every device, since differences found between devices demonstrate that global characterization is not a feasible solution to guarantee the correct luminance and contrast reproduction, even considering devices from the same manufacturing batch. If reproduction and cross-reproduction errors are not considered by the manufacturers and designers, and only variations in RGB space (belonging to that particular device) are considered, vision specialists should consider that these measurements cannot be comparable between devices, and the obtained results will be relative to that particular screen, not an absolute measurement.

Colorimetric characterization requires very specific instrumentation and knowledge, which is unavailable for the final user, and for this reason the implication of visual specialists is imperative. Visual specialists could provide visual tests but also the characterized devices to ensure correct performance of the patient. Future solutions could involve the development of hardware with integrated photometers for self-calibrating, [35] but the cost is still high, and this solution is not affordable in handheld devices.

Following the results obtained in the present study, contrast reproduction between devices is not as variable as the luminance values, and it is possible that clinically, the subject's variability in contrast sensitivity measurements is higher than the device variability, but future studies should answer this question. Although reproduction and cross-reproduction errors currently have no simple solution, vision specialists should be aware of these limitations when designing and performing visual tests in electronic devices.

## 5. Conclusions

Considering the studied devices from the same model, luminance differences between devices can exceed in most cases the incremental threshold of the eye. Every device will reproduce the same software order with different luminance. We cannot assume that characterization of one device can be extrapolated to other devices even if they are from the same manufacturing batch. Every device used for research purposes should be individually characterized. For clinical purposes, limitations should be considered by visual specialists.

**Author Contributions:** Conceptualization, methodology, writing—original draft preparation, A.M.-M., D.P.P., D.d.F.; writing—review and editing, A.M.-M., D.P.P., D.d.F., M.B.C.-M., L.L.-V.; All authors have read and agreed to the published version of the manuscript.

**Funding:** The author David P. Piñero has been supported by the Ministry of Economy, Industry and Competitiveness of Spain within the program Ramón y Cajal, RYC-2016-20471. The study was developed with the support of the project OPTiTRAIN (IDI-20180123), co-financed by Centre for Industrial Technological Development (CDTI) and the European Regional Development Fund (ERDF).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hogarty, D.T.; Hogarty, J.P.; Hewitt, A.W. Smartphone use in ophthalmology: What is their place in clinical practice? *Surv. Ophthalmol.* **2020**, *65*, 250–262. [[CrossRef](#)] [[PubMed](#)]
2. Lakowski, R. Theory and practice of colour vision testing: A review. *Br. J. Ind. Med.* **1969**, *26*, 173–189.
3. O’Neill, S.; McAndrew, D.J. The Validity of Visual Acuity Assessment Using Mobile Technology Devices in the Primary Care Setting. *Aust. Fam. Physician* **2016**, *45*, 212–215.
4. Zhang, Z.-T.; Zhang, S.-C.; Huang, X.-G.; Liang, L.-Y. A Pilot Trial of the iPad Tablet Computer as a Portable Device for Visual Acuity Testing. *J. Telemed. Telecare* **2013**, *19*, 55–59. [[CrossRef](#)]
5. Black, J.M.; Jacobs, R.; Phillips, G.; Chen, L.; Tan, E.; Tran, A.; Thompson, B. An Assessment of the iPad as a Testing Platform for Distance Visual Acuity in Adults. *BMJ Open* **2013**, *3*, e002730. [[CrossRef](#)]
6. Vingrys, A.J.; Healey, J.K.; Liew, S.; Saharinen, V.; Tran, M.; Wu, W.; Kong, G.Y. Validation of a Tablet as a Tangent Perimeter. *Transl. Vis. Sci. Technol.* **2016**, *5*, 3. [[CrossRef](#)]
7. Kingsnorth, A.; Drew, T.; Grewal, B.; Wolffsohn, J.S. Mobile App Aston Contrast Sensitivity Test. *Clin. Exp. Optom.* **2016**, *99*, 350–355. [[CrossRef](#)]
8. Rodríguez-Vallejo, M.; Remón, L.; Monsoriu, J.A.; Furlan, W.D. Designing a new test for contrast sensitivity function measurement with iPad. *J. Optom.* **2015**, *8*, 101–108. [[CrossRef](#)] [[PubMed](#)]
9. de Fez, D.; Luque, M.J.; Matea, L.; Piñero, D.P.; Camps, V.J. New iPad-based Test for the Detection of Color Vision Deficiencies. *Graefes Arch. Clin. Exp. Ophthalmol.* **2018**, *256*, 2349–2360. [[CrossRef](#)] [[PubMed](#)]
10. Keilty, M.; Houston, K.E.; Collins, C.; Trehan, R.; Chen, Y.T.; Merabet, L.; Watts, A.; Pundlik, S.; Luo, G. Inpatient Virtual Vision Clinic Improves Access to Vision Rehabilitation Before and During the COVID-19 Pandemic. *Arch. Rehabil. Res. Clin. Transl.* **2021**, *3*, 100100.
11. Aslam, T.M.; Parry, N.R.; Murray, I.J.; Salleh, M.; Col, C.D.; Mirza, N.; Czanner, G.; Tahir, H.J. Development and testing of an automated computer tablet-based method for self-testing of high and low contrast near visual acuity in ophthalmic patients. *Graefes Arch. Clin. Exp. Ophthalmol.* **2016**, *254*, 891–899. [[CrossRef](#)] [[PubMed](#)]
12. Aruliyothi, L.; Janakiraman, A.; Malligarjun, B.; Babu, B.M. Smartphone applications in ophthalmology: A quantitative analysis. *Indian J. Ophthalmol.* **2021**, *69*, 548–553. [[CrossRef](#)] [[PubMed](#)]
13. American Association of Physicists in Medicine. *Display Quality Assurance—RPT 270*; American Association of Physicists in Medicine: Alexandria, VA, USA, 2019.
14. Badano, A.; Revie, C.; Casertano, A.; Cheng, W.-C.; Green, P.; Kimpe, T.; Krupinski, E.; Sisson, C.; Skrvøseth, S.O.; Treanor, D.; et al. Consistency and standardization of color in medical imaging: A consensus report. *J. Digit. Imaging* **2015**, *28*, 41–52. [[CrossRef](#)]
15. de Fez, D.; Luque, M.J.; García-Domene, M.C.; Camps, V.; Piñero, D. Colorimetric Characterization of Mobile Devices for Vision Applications. *Optom. Vis. Sci.* **2015**, *93*, 85–93. [[CrossRef](#)]
16. de Fez, D.; Luque, M.J.; García-Domene, M.C.; Caballero, M.T.; Camps, V.J. Can Applications Designed to Evaluate Visual Function Be Used in Different iPads? *Optom. Vis. Sci.* **2018**, *95*, 1054–1063. [[CrossRef](#)]
17. Aslam, T.M.; Murray, I.J.; Lai, M.Y.; Linton, E.; Tahir, H.J.; Parry, N.R. An assessment of a modern touch-screen tablet computer with reference to core physical characteristics necessary for clinical vision testing. *J. R. Soc. Interface* **2013**, *10*, 20130239. [[CrossRef](#)]
18. Norton, T.T.; Corliss, D.A.; Bailey, J.E. *The Psychophysical Measurement of Visual Function*; Butterworth-Heinemann: Oxford, UK, 2002.
19. Krauskopf, J. Discrimination and detection of changes in luminance. *Vis. Res.* **1980**, *20*, 671–677. [[CrossRef](#)]
20. Gibson, J.E.; Fairchild, M.D. *Colorimetric Characterization of Three Computer Displays (LCD and CRT)*; Munsell Color Science Laboratory Technical Report; Rochester Institute of Technology: Rochester, NY, USA, 2000.
21. Fairchild, M.; Wyble, D. *Colorimetric Characterization of the Apple Studio Display (flat panel LCD)*; Rochester Institute of Technology: Rochester, NY, USA, 1998; p. 22.
22. Seime, L.; Hardeberg, J.Y. Colorimetric characterization of LCD and DLP projection displays. *J. Soc. Inf. Disp.* **2003**, *11*, 349. [[CrossRef](#)]
23. Day, E.A.; Taplin, L.; Berns, R.S. Colorimetric characterization of a computer-controlled liquid crystal display. *Color Res. Appl.* **2004**, *29*, 365–373. [[CrossRef](#)]
24. Jung, H.; Kim, H.J.; Kang, W.S.; Yoo, S.K.; Fujioka, K.; Hasegawa, M.; Samei, E. Assessment of flat panel LCD primary class display performance based on AAPM TG 18 acceptance protocol. *Med. Phys.* **2004**, *31*, 2155–2164. [[CrossRef](#)]

25. Vrhel, M.J.; Trussell, H.J. Color device calibration: A mathematical formulation. *IEEE Trans. Image Process.* **1999**, *8*, 1796–1806. [[CrossRef](#)] [[PubMed](#)]
26. Berns, R.S. Methods for characterizing CRT displays. *Displays* **1996**, *16*, 173–182. [[CrossRef](#)]
27. International Commission on Illumination. *The Relationship between Digital and Colorimetric Data for Computer-Controlled CRT Displays*; CIE Central Bureau: Vienna, Austria, 1996.
28. Thomas, J.-B.; Hardeberg, J.Y.; Foucherot, I.; Gouton, P. The PLVC display color characterization model revisited. *Color Res. Appl.* **2008**, *33*, 449–460. [[CrossRef](#)]
29. Bodduluri, L.; Boon, M.Y.; Dain, S.J. Evaluation of tablet computers for visual function assessment. *Behav. Res. Methods* **2017**, *49*, 548–558. [[CrossRef](#)] [[PubMed](#)]
30. García-Domene, M.C.; Luque, M.J.; de Fez, D. The problem of spatial homogeneity in an LCoS projector. *Optik* **2019**, *198*, 163252. [[CrossRef](#)]
31. Castro, J.J.; Pozo, A.M.; Rubiño, M. Color dependence with horizontal viewing angle and colorimetric characterization of two displays using different backlighting. Proc SPIE 8785. In Proceedings of the 8th Iberoamerican Optics Meeting and 11th Latin American Meeting on Optics, Lasers, and Applications, Porto, Portugal, 18 November 2013.
32. Sheskin, D.J. *Handbook of Parametric and Nonparametric Statistical Procedures*, 3rd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2003.
33. Barco, L.D.J.; Diaz, J.A.; Jimenez, J.R.; Rubino, M. Considerations on the calibration of color displays assuming constant channel chromaticity. *Color Res. Appl.* **1995**, *20*, 377–387. [[CrossRef](#)]
34. Artal, P. (Ed.) *Handbook of Visual Optics: Instrumentation and Vision Correction*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2017; Volume 2.
35. Han, H.K.; Jones, P.R. Plug and play perimetry: Evaluating the use of a self-calibrating digital display for screen-based threshold perimetry. *Displays* **2019**, *60*, 30–38. [[CrossRef](#)]