Technical University of Berlin

Fakultät Elektrotechnik und Informatik
Computational Psychology

Experimental Monitor Calibration with Unstable Gamma Function Due to Temperature Dependency

Bachelor Thesis

Submitted by: Konstantinos Zamanis 412949

Supervisor: Dr. Guillermo Aguilar
guillermo.aguilar@mail.tu-berlin.de

2nd Supervisor: Prof. Guillermo Gallego
guillermo.gallego@tu-berlin.de

Date of Submission: January 2025

Contents

Declaration of Authorship

I hereby declare that the thesis submitted is my own, unaided work, completed without any external help. Only the sources and resources listed were used. All passages taken from the sources and aids used, either unchanged or paraphrased, have been marked as such.

Where generative AI tools were used, I have indicated the product name, manufacturer, the software version used, as well as the respective purpose (e.g., checking and improving language in the texts, systematic research). I am fully responsible for the selection, adoption, and all results of the AI-generated output I use.

I have taken note of the Principles for Ensuring Good Research Practice at TU Berlin dated 15 February 2023.

I further declare that I have not submitted the thesis in the same or similar form to any other examination authority.

Berlin, 20/01/2025

Signature ____________________

Note of Thanks

This bachelor thesis focuses on the calibration of an experimental monitor used in vision science, a project that provided both technical challenges and valuable learning opportunities. I would like to express my gratitude to Dr. Guillermo for entrusting me with this task and allowing me to independently explore solutions without micromanagement. Allthough in the beginning I felt that I was in deep waters I gained greater confidence when the steps that I had planed in my mind started to make sense and lead me to results. This approach not only enhanced my understanding of calibration and vision science but also significantly impacted my personal and professional growth and critical way of thinking.

Through this process, I gained a deeper appreciation for the complexities involved in experimental setups and the critical role of calibration in ensuring valid results. Moreover, this experience taught me to approach new tasks with confidence and resilience. I learned to embrace imperfection and present my work, even when it was not flawless, as part of a continuous learning process. This thesis represents not only a technical achievement but also a testament to the importance of autonomy and self-reliance in research.

Abstract

This study focuses on calibrating an experimental monitor used in the Faculty of Computational Psychology for vision science experiments. The monitor's performance is significantly influenced by temperature dependency, a challenge that has been acknowledged but not systematically measured or resolved. Intrigued by earlier findings of Jonas Schmiegel, this research aims to better understand, quantify and address this issue.

The study characterizes the monitor's temperature-dependent behavior, determines its impact on luminance output, and evaluates the possibility of implementing a calibration strategy to mitigate these effects. A systematic methodology was employed, involving repeated measurements of the monitor's luminance output and temperature under controlled conditions. Detailed descriptions of the experimental setup and procedures ensure replicability.

The results reveal that the monitor is highly inconsistent across different temperatures and days, making conventional Look-Up Table (LUT) calibration methods ineffective. These findings highlight the need for dynamic and adaptive calibration approaches to address temperature dependency in experimental monitors, that also make up for the monitors inconsistencies.

This thesis uncovers potential obstacles to calibrating a monitor and makes clear what future researchers need to keep in mind when dealing with calibration strategies. It contributes to understanding the limitations of existing calibration techniques and lays the groundwork for future studies to develop more robust solutions.

Zusammenfassung

Diese Studie konzentriert sich auf die Kalibrierung eines experimentellen Monitors, der in der Fakultät für Computational Psychology für visuelle Experimente eingesetzt wird. Die Leistung des Monitors wird erheblich von Temperaturabhängigkeiten beeinflusst – eine Herausforderung, die zwar anerkannt, aber bisher nicht systematisch gemessen oder gelöst wurde. Angeregt durch frühere Erkenntnisse von Jonas Schmiegel zielt diese Forschung darauf ab, das Problem besser zu verstehen, zu quantifizieren und anzugehen.

Die Studie charakterisiert das temperaturabhängige Verhalten des Monitors, bestimmt dessen Einfluss auf die Luminanzausgabe und bewertet die Möglichkeit, eine Kalibrierungsstrategie umzusetzen, um diese Effekte zu mildern. Eine systematische Methodik wurde angewandt, die wiederholte Messungen der Luminanzausgabe und der Temperatur des Monitors unter kontrollierten Bedingungen umfasst. Detaillierte Beschreibungen des experimentellen Aufbaus und der Verfahren gewährleisten die Reproduzierbarkeit.

Die Ergebnisse zeigen, dass der Monitor bei unterschiedlichen Temperaturen und an verschiedenen Tagen sehr inkonsistent ist, was herkömmliche Kalibrierungsmethoden mit Look-Up-Tabellen (LUT) unwirksam macht. Diese Ergebnisse unterstreichen die Notwendigkeit dynamischer und adaptiver Kalibrierungsansätze, die nicht nur die Temperaturabhängigkeit berücksichtigen, sondern auch die Inkonsistenzen des Monitors ausgleichen.

Diese Arbeit deckt potenzielle Hindernisse bei der Kalibrierung eines Monitors auf und macht deutlich, worauf zukünftige Forscher bei der Entwicklung von Kalibrierungsstrategien achten müssen. Sie trägt zum Verständnis der Grenzen bestehender Kalibrierungstechniken bei und legt die Grundlage für zukünftige Studien zur Entwicklung robusterer Lösungen.

1 Introduction

1.1 Background and Motivation

Vision science studies how living beings see and perceive the visual world. This field encompasses a variety of topics, including the physics of light and the psychology of visual perception. In many of our team's experiments, participants are presented with visual stimuli on monitors while their responses are recorded. Achieving robust and admissible results in such experiments requires adherence to fundamental principles of experimental design, including careful consideration of participant demographics, task design, and environmental control. This thesis focuses on the latter, specifically addressing the challenges of monitor calibration.

A controlled environment is essential for ensuring accurate and reliable results in vision science experiments. Variations in monitor settings, such as brightness, contrast, or color representation, can significantly influence participants' perceptions and responses, potentially introducing inconsistencies in the data (Lu & Dosher, 2013). For instance, if two participants complete the same experiment on monitors with differing brightness levels, one might perceive the stimuli more easily, creating a disparity in their results. This makes it challenging to distinguish whether observed differences are due to variations in visual perception or technical inconsistencies in the displays.

To address this, proper monitor calibration is critical to ensure that visual stimuli are presented consistently across all participants and experimental conditions. By eliminating variations in display settings, researchers can obtain data that more accurately reflect the intended experimental parameters.

1.2 Gamma function

In order to understand monitor calibration it is essential to comprehend what the Gamma function is. The gamma function is a fundamental function that describes the non-linear relationship between a pixel's numerical value (input) and its corresponding output luminance (output), accounting for the non-linear response of human vision to changes in luminance. This relationship is mathematically represented as a power-law function:

$$Output = Input^{\gamma} \tag{1}$$

In this equation, γ (gamma) is the monitor's gamma value, which in the practice varies between devices. It defines how the input pixel values are transformed into luminance output. An example of this transformation can be observed in Figure 1

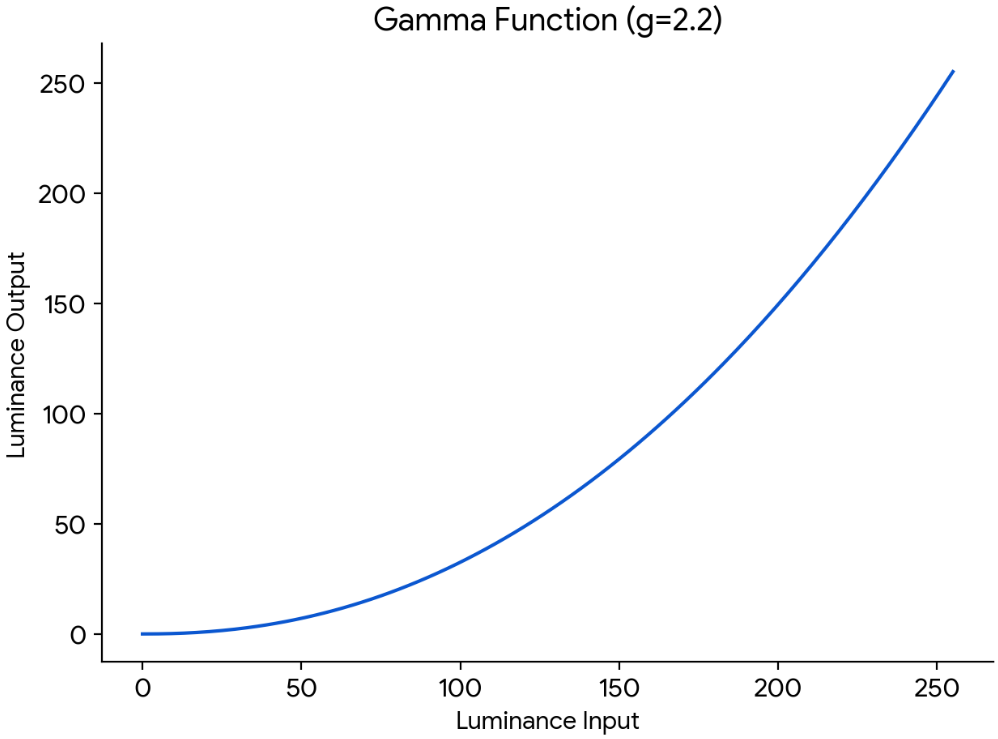

Figure 1: The x-axis represents the luminance input, and the y-axis represents the luminance output. The plot shows the gamma function with a γ value of 2.2.

Monitors apply this gamma correction to better match the human eye's sensitivity, which is more attuned to relative changes in darker areas than in brighter areas.

While a camera sensor detects twice the photons as twice the brightness, the human eye perceives only a fraction of this increase as being brighter. This biological adaptation allows our vision to operate effectively over a wide range of luminance levels, from dimly lit environments to bright outdoor settings.

The gamma function bridges this gap by encoding luminance in a way that aligns more closely with human perception. It ensures that variations in brightness are distributed in a way that is visually meaningful, emphasizing the darker tones where our eyes are more sensitive while compressing the brighter tones where our sensitivity decreases. Without gamma correction, displayed images would appear unnatural, failing to account for the unique characteristics of human vision.

Images of Gamma 1.0 (on the left) comparing to Gamma 2.2 (on the right)

Figure 2: The figure shows the same picture rendered with different gamma functions. On the left, the γ value is 1, whereas on the right, the γ value is 2.2. (Harding, 2021)

Now from a technical aspect, the gamma function is essential in visual science experiments, as it ensures that the luminance levels of displayed stimuli are consistent with the experimental design's intended values. Proper gamma correction improves the reliability and interpretability of experimental results by aligning the stimuli with human perceptual responses, minimizing unintended variability caused by device-specific characteristics.

Theoretically, a consistent gamma value of $\gamma = 1$ would result in a linear relationship between input luminance (Lin) and output luminance (Lout), as depicted in Figure 2. In this ideal scenario, we do not take into consideration the fact that temperature fluctuations do not affect the monitor's output. However, practical observations deviate from this theoretical expectation.

It is well-established that temperature can influence a monitor's gamma value (Schmiegel, 2023). Unfortunately, a standardized approach to quantify this influence is lacking. Consequently, each monitor requires individual characterization to accurately assess its temperature-dependent behavior. This initial step is essential to address the challenges posed by temperature variations.

1.3 How Monitors Work in Vision Research

The process of displaying stimuli on a monitor involves a pipeline starting from software-defined stimuli to the eventual visual output. In our lab, stimuli are defined as numerical arrays (e.g., using NumPy) normalized to values between 0 and 1 instead of the usual 0-255 values. These arrays are passed to rendering software, which sends the data to the graphics card. The graphics card then transmits the input values to the monitor, which converts them into visible luminance.

However, monitors do not inherently provide a linear relationship between input values and luminance output. This expansive power function, expressed as $y = x^{\gamma}$, is a characteristic of individual monitors and varies between devices as mentioned earlier. This non-linearity complicates the creation of precise stimuli, as the actual luminance emitted may not correspond directly to the intended values, but this has not been a problem so far.

To give a solution to this problem, linearize this gamma function after it has been established. Linearization involves applying an inverse function to the input values before they are processed by the monitor. This correction ensures that the relationship between input values and luminance output becomes linear, simplifying stimulus creation and ensuring accuracy. The corrected function is stored as a lookup table (LUT), which maps input values to adjusted outputs. These corrections are applied pixel-wise during rendering, enabling consistent luminance levels across the screen.

In our lab, the linearization process is facilitated by specialized tools and software. For example, photometric measurements are taken across a wide range of input values, and the inverse of the luminance response is calculated with high precision by the aforementioned software. These measurements are then used to generate the LUT, which is subsequently validated to confirm the linearization's accuracy. The visualization of this process can be seen on Figure 3

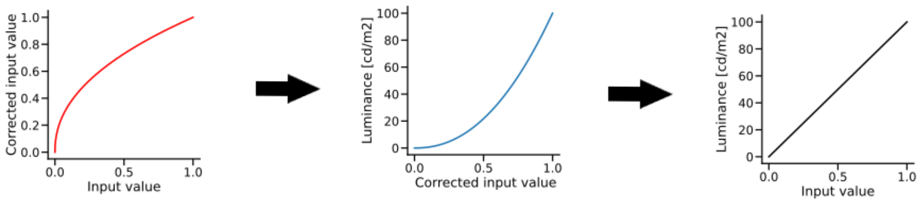

Figure 3: These three graphs describe the linearization process. In the graph on the left, the input value is on the x-axis, and the corrected value is on the y-axis. In the graph in the middle, the correct input value is on the x-axis, and the luminance in cd/m^2 is on the y-axis. Finally, in the graph on the right, we have the linearized graph with the input value on the x-axis and the luminance in cd/m^2 on the y-axis.

1.4 Challenges with monitors

While GPUs and monitors both apply gamma correction, the monitor's gamma value often varies between devices and is typically unknown. This creates challenges in achieving a linear relationship between input pixel values and output luminance, which is vital for ensuring the intended luminance levels are accurately represented. Misalignments in this relationship can distort visual experiments and reduce the reliability of findings. So far to address this, calibration tools and procedures are employed to measure and adjust the monitor's gamma response. Recently, new findings have established that temperature can influence a monitor's gamma value, there unfortunately no standardized approach to quantify this effect across different hardware (Schmiegel, 2023). Consequently, each monitor must be individually characterized to accurately assess its temperature-dependent behavior. This characterization is a crucial first step in addressing the challenges posed by temperature variations. Before conducting experiments, it is essential to characterize the monitor's performance by taking into account all influencing parameters. Preliminary measurements in our lab, in Figure 4 indicate that, even when the input value remains stable, the output luminance can drop by up to 35% over a three-hour period. This highlights the importance of ongoing calibration to mitigate the effects of luminance drift and maintain experimental precision.

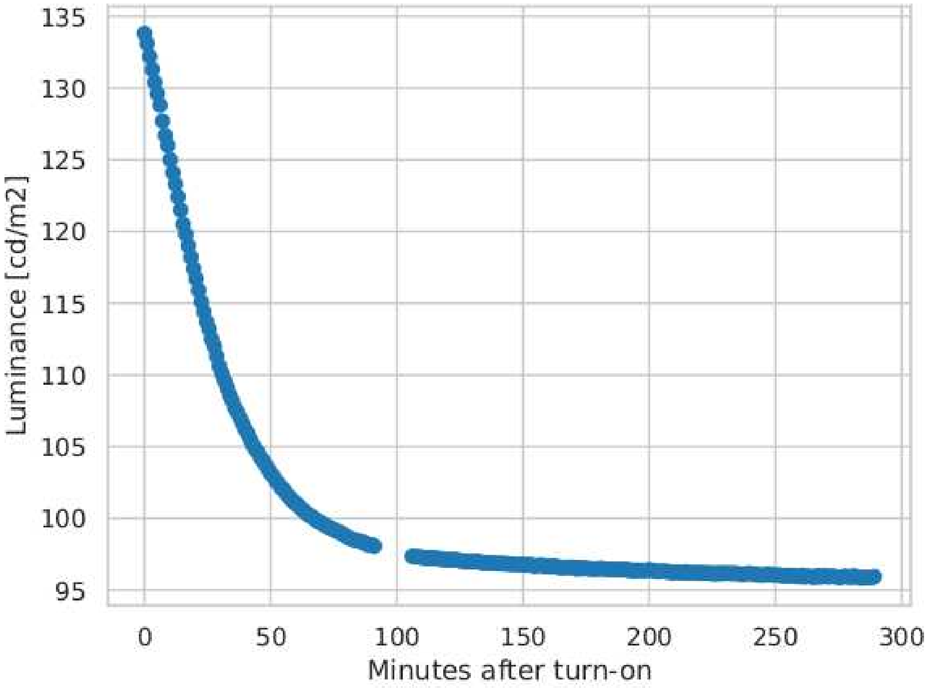

Figure 4: The graph shows the decline in luminance output over time with a fixed input value. On the x-axis, we have the minutes after turn-on, and on the y-axis, we have the luminance in cd/m^2.

1.5 Previous work

More specifically, after the research of our colleagues' team in Tübingen, new findings have indicated that two parameters—the internal temperature of the monitor and the ambient temperature—influence the monitor calibration process (Schmiegel, 2023). Jonas Schmiegel's thesis explored this temperature dependency using monitors similar to those employed by our team. Specifically, his research utilized two VIEWPixx /3D Lite monitors, which is the same monitor that our team uses in our lab. Despite manufacturer recommendations to warm up monitors before experiments, the effect of temperature on luminance output performance had not been thoroughly quantified until Schmiegel's work.

Schmiegel measured luminance discrepancies caused by monitor warm-up using a photometer. This device, aligned and focused on the center of the screen at a distance of 30 cm and an angle of approximately 6.5 degrees, captured luminance measurements in candela per square meter. These measurements were recorded alongside internal temperature data, obtained via the monitors' embedded thermometers (Schmiegel,2023).

As preliminary measurements revealed, our monitor behavior displayed a dependency on temperature. However, this dependency had not been systematically quantified or compared to the equipment used by our colleagues. Jonas Schmiegel addressed this issue effectively by creating a graph with input intensity values (Lin) on the x-axis and the corresponding measured luminance (Lout) on the y-axis. This graph was constructed using repeated measurements for all 33 input values during the monitor's warm-up phase, resulting in multiple data points for each input intensity.

To emphasize the temperature dependency, Schmiegel employed a color-coding approach where each data point was shaded based on the monitor's internal temperature at the time of measurement. In his case, the temperature range was represented from deep blue for the minimum temperature (28°C) to dark red for the maximum temperature (42°C)(Schmiegel,2023).

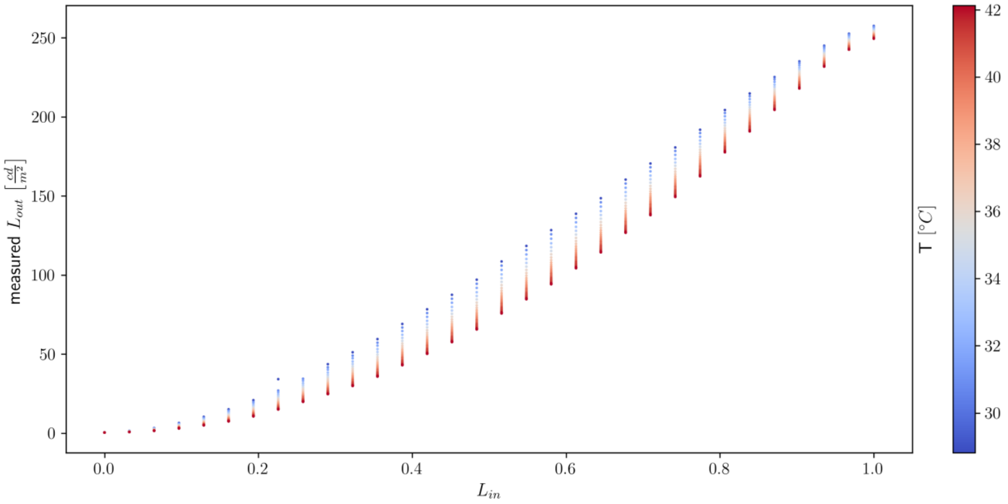

Figure 5: Variation of absolute luminance as well as gamma dependency with increasing temperature during warm-up. Adapted from *Characterisation of a High Luminance Resolution Display for Psychophysical Experiments* by J. Schmiegel (2023).

This visualization clearly demonstrated a pattern: at lower temperatures, the luminance output (Lout) was higher, while as the monitor warmed, (Lout) decreased for the same input intensity. The most significant discrepancies were observed in the mid-range of the input spectrum, particularly for input values between 0.4375 and 0.625. In contrast, the variation in luminance output was much smaller at the extreme ends of the input range, indicating that the temperature dependency was less pronounced in these regions. These findings underscore the non-linear and temperature-sensitive nature of the monitor's gamma response, further validating the need for real-time temperature compensation to ensure consistent luminance calibration. Schmiegel's data revealed significant fluctuations in luminance output for the same input values, with variations of up to 40% depending on the monitor's temperature.

More specifically, with an input value near 0.6, Schmiegel observed luminance output ranging from approximately 105 cd/m^2 to 140 cd/m^2 as the monitor temperature increased from 30°C to 42°C. This highlights the critical importance of calibration in ensuring reliable experimental results. My research builds on this by replicating Schmiegel's measurements with our monitor, aiming to better characterize its performance and assess its comparability with Schmiegel's findings.

The ultimate goal of my research is to determine whether the discrepancies observed due to temperature fluctuations can be compensated for. By replicating measurements, analyzing warming rates, and developing correction methods, this thesis seeks to linearize the luminance input-output relationship to minimize variability and improve the reliability of experimental results.

1.6 Research question

The ultimate aim of this research is to address a specific challenge in monitor calibration: the extent to which the luminance dependency of the gamma function, influenced by internal and ambient temperature, can be compensated for in experimental monitors. By replicating measurements, analyzing warming rates, and developing correction methods, this study seeks to linearize the luminance input-output relationship to minimize variability and ensure consistency across varying conditions, specifically removing the monitor's temperature as an influencing parameter. By achieving this, the research contributes to improving the reliability and accuracy of experimental results in vision science.

2 Methods

This section outlines the equipment and software used to conduct the measurements, as well as the detailed process employed to capture luminance and temperature data. It describes how stimuli were displayed on the monitor, how the photometer and sensors were used to record measurements, and how the data was structured to be analyzed. By explaining the methodology in detail and providing pseudocode for the measurement script, this section ensures the process is replicable regardless of the specific experimental setup, enabling future researchers to reproduce the findings accurately.

2.1 Equipment and software used

For hardware, I utilized the same monitor as our colleagues in Tübingen, the VIEWPixx monitor (Vpixx Technologies, Inc., Saint-Bruno, QC, Canada). Measurements were conducted using a Minolta LS-100 photometer (Konica Minolta, Tokyo, Japan). Additionally, a NODEMCU microcontroller paired with a DHT22 (AM2303) sensor was used to measure the ambient temperature during the experiments.

To standardize the luminance input, I developed a Python script that scaled intensity values by dividing them by 32, resulting in a normalized range from 0 to 1 in increments of 0.03125. This range formed the basis for conducting all measurements.

To ensure reproducibility and transparency, I have included all plots and Python code used for the measurements in my GitLab repository, which provides detailed instructions. The repository is accessible via the following link: GitLab Repository.

2.2 Measurement process

The Python script was designed to display stimuli with predefined luminance values on the monitor while capturing the corresponding luminance output using the Minolta photometer. Stimuli were created as a rectangle occupying half the width and half the height of the monitor. The photometer was placed approximately half a meter from the monitor and adjusted to a height where it pointed precisely to the center of the screen, as it can be seen in Figure 6. To ensure accurate alignment, another Python script displayed a dot at the center of the rectangle to guide the photometer's positioning.

Figure 6: Measurement setup, where the photometer is aligned to the center of the monitor

Stimuli were presented in a randomized order to mitigate hysteresis effects. Initial measurements revealed that the monitor's temperature stabilized after approximately 90 minutes, so all subsequent measurements were planned to last around this duration. I had 32 predefined stimulus values, and during each iteration (cycle), after the order of the values would get randomized, the program would go through all of them sequentially. To ensure accurate timing, I adjusted the program's delay between displaying each image and capturing the corresponding luminance with the photometer, which required approximately 2.6 seconds per value. Typically, 60 cycles were run per session, ensuring that the entire measurement process for each day lasted the planned 90 minutes.

The Python script also recorded the data into a file for each measurement day. These files contained the following columns: Timestamp, Stim value, Temperature, Temperature2, Measured Luminance, and Cycle, providing all the necessary information for subsequent analysis of the data. A pseudocode representation of the measurement process can be found in Algorithm 1.

The script allowed for automated repetition of measurements over any desired time frame. To ensure reliable measurement results, it was crucial to account for the monitor's warm-up time. By analyzing multiple measurements taken on different days, we could evaluate the consistency of this warm-up rate and its impact on the monitor's luminance output.

Algorithm 1: Pseudocode for the measurement process.

```
Data: Stimulus values, photometer, Arduino sensor
Result: Measurement data for luminance and temperature
Set background lighting to 0;
Initialize photometer;
Initialize Arduino sensor;
Define stim_vals = [0, 0.03125, 0.0625, ..., 1.0];
window_height = monitor_height / 2;
window_width = monitor_width / 2;
Open file measurement_data.csv in write mode;
Write header row [Timestamp, Stim value, Temperature,
 Temperature2, Measured Luminance, Cycle];
start_time = current time;
cycles = 60;
for each cycle in cycles do
    Shuffle stim_vals;
    for each stim_value in stim_vals do
        Display rectangle with stim_value, window_width, and
         window_height;
        Display fixation point at center of rectangle;
        measured_luminance = read luminance from photometer;
        internal_temperature = read monitor temperature;
        ambient_temperature = read temperature from Arduino;
        timestamp = current time - start_time;
        Write row [timestamp, stim_value, internal_temperature,
         ambient_temperature, measured_luminance, cycle] to file;
        Wait for 2.6 seconds;
    end
end
Close file measurement_data.csv;
```

Additionally, the entire measurement setup was surrounded by special curtains to completely block external light, minimizing interference from ambient illumination. Using this setup, I replicated Jonas Schmiegel's graph, demonstrating the variation in absolute luminance and gamma dependency during the monitor's warm-up period as its internal temperature increased.

Simultaneously, I recorded the internal temperature of the monitor and introduced additional measurements of ambient temperature and relative humidity using a custom-built Arduino sensor. These external measurements enabled a comprehensive characterization of the monitor's performance and its comparison to findings from Tübingen.

3 Results

The results of this study reveal significant variability in monitor performance across multiple days, particularly in warm-up rates and luminance output. These findings highlight the challenges associated with achieving consistent calibration.

3.1 Warm-up rate variations

Figure 7 depicts the warm-up behavior of the monitor across four different days, with time in minutes on the x-axis and temperature in degrees Celsius on the y-axis. Although it was anticipated that the monitor's warm-up pattern would be consistent, the results indicated otherwise. The data reveal significant variability both in starting temperatures and the time required to reach the maximum operating temperature.

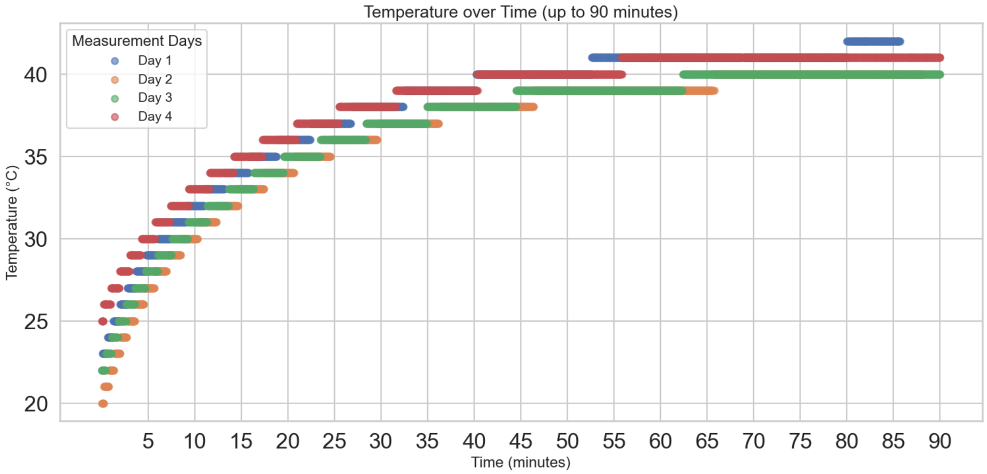

Figure 7: This plot shows the warmup behavior of the monitor across four different days, with time in minutes on the x-axis and temperature in degrees Celsius on the y-axis.

For instance, all four measurements began at different initial temperatures—23°C on Day 1, 21°C on Day 2, 22°C on Day 3, and 25°C on Day 4—potentially influenced by variations in ambient temperature. Before each measurement, care was taken to ensure that the monitor was not used beforehand to maintain a controlled environment. Despite these precautions, the monitor had unpredictable behavior.

Notably, in the first five minutes after power-on, the temperature readings showed variability: 29°C on Day 1, 27°C on Day 2, 28°C on Day 3, and 30°C on Day 4. Furthermore, the time taken to reach the maximum temperature of 42°C differed significantly across days, with Day 4 reaching this point fastest at 57 minutes, while Day 1 took 80 minutes.

These results confirmed the challenge of standardizing warm-up behavior and validated the decision to employ real-time temperature measurements rather than relying on a pre-determined warm-up period for the calibration techniques.

Unfortunately, I was unable to find a correlation that would help in the analysis of the measurements between ambient temperature and temperature or between ambient temperature and measured luminance. The first relationship can be observed in Figure 8 and as it can be observed, while the temperature of the monitor rises from 25 to 41°C the ambient temperature only went from 23 to 24.1°C, so I decided to disregard it in my analysis.

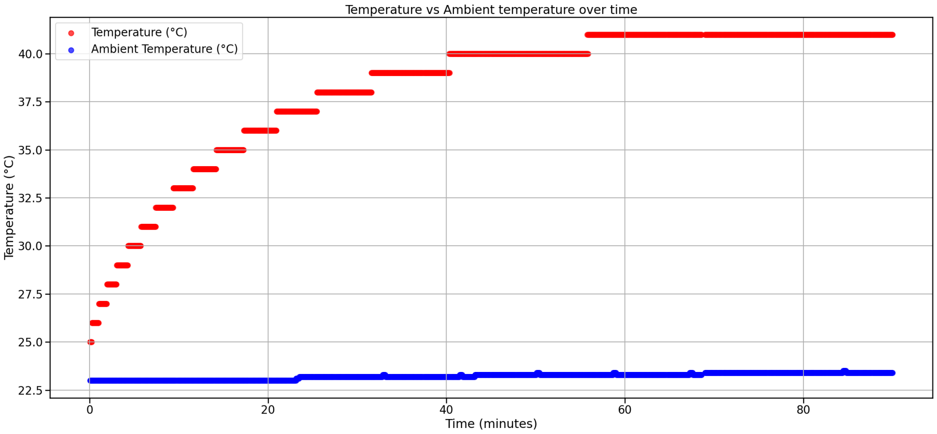

Figure 8: Temperature vs ambient temperature over time, with Temperature in °C on the y-axis and Time in minutes on the x-axis, with Monitor temperature being the red line and ambient temperature the blue line.

3.2 Behavioral trends of the monitor across input values

In this section, the behavior of the monitor over time is analyzed in greater detail. Observations drawn from our colleagues' graphs are corroborated here. It is evident that as the monitor operates at cooler temperatures, it exhibits higher luminance values. Conversely, as the monitor approaches its warmed-up state—defined as reaching 42°C—luminance decreases and eventually stabilizes.

To illustrate this behavior, Figure 9 presents the overall trend of the monitor's luminance output across 16 input values, sampled at every second value from the predefined list. This general graph highlights the consistent pattern of decreasing luminance as the monitor warms up over time, demonstrating the global behavior across a wide range of input values.

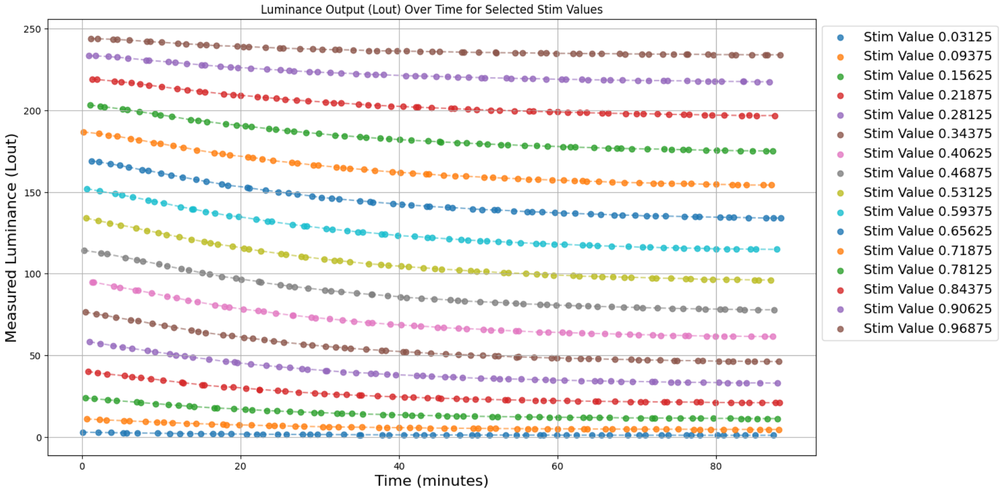

Figure 9: The plot shows measured luminance on the y-axis and time in minutes on the x-axis, representing data for 16 values.

To go deeper into this trend, Figure 10 isolates the graph to focus on 6 particular input values from the middle and extremes of the spectrum. This targeted view highlights more significant behaviors, such as the greater variability in the mid-range input values compared to the relatively stable behavior observed at the extremes.

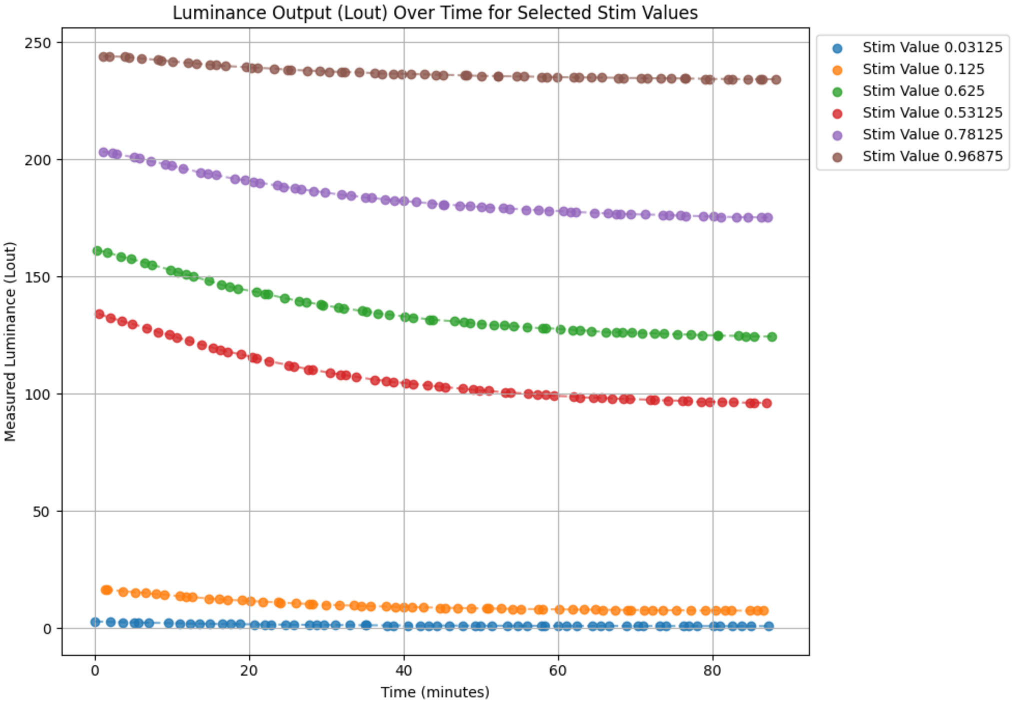

Figure 10: The plot shows measured luminance on the y-axis and time in minutes on the x-axis, representing data for 5 values.

A particularly notable observation is that the impact of the monitor's temperature is significantly more pronounced for mid-range input values compared to extreme values. For instance, with an input value of 0.625, luminance values range from 162 to 124, whereas for the input value 0.03125, the range is from 1.8 to 0.6. While the percentage difference is larger for the latter case, the absolute difference is minimal and not perceivable to the human eye. As shown in Figure 7, a more explicit example of minimal variance can be observed. When focusing solely on the extreme data points, the resulting graph exhibits a near-linear trend, suggesting a weak relationship between the variables.

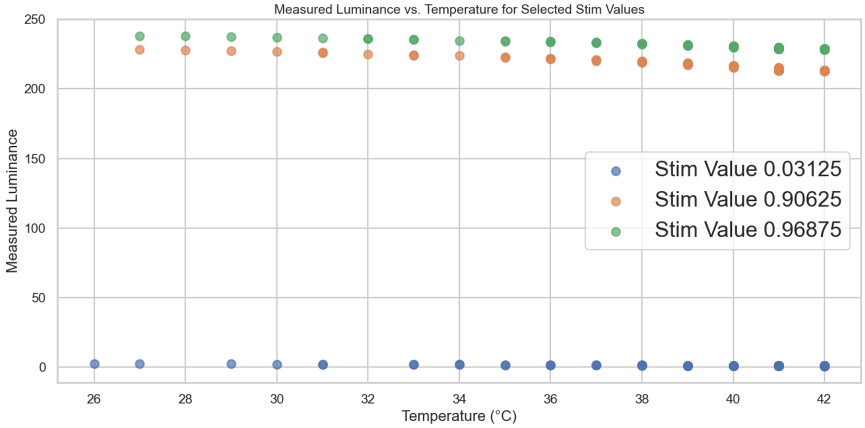

Figure 11: The plot shows luminance output on the y-axis versus temperature on the x-axis for extreme Stim values, highlighting less significant variance.

3.3 Recreation of Jonas Schmiegel's graph

In this section, I replicated the graph utilized in Jonas Schmiegel's work in Figure 12, which plots stimulus values on the x-axis and luminance output Lout on the y-axis, with data points color-coded according to the monitor's temperature at the time of measurement. This approach serves as a basis for comparing the performance and behavior of our monitors, which share the same model.

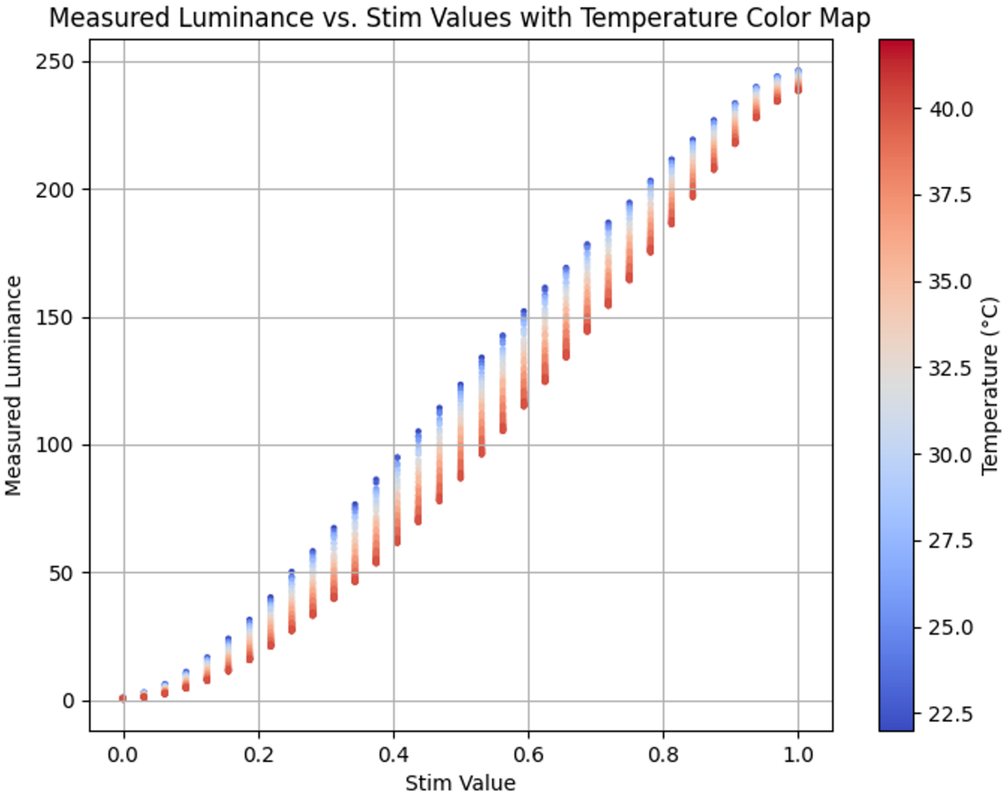

Figure 12: Replication of Jonas Schmiegel's graph on the VIEWPixx monitor, with Stim values on the x-axis and measured luminance on the y-axis. The points on the graph are colored from deep blue to dark red, corresponding to the temperature of the monitor at the time of measurement.

During these measurements, the monitor's minimum temperature was 22.5°C, stabilizing after reaching its maximum temperature of 42°C. The graph reveals a similar behavior to that observed in Schmiegel's results. The mid-range stimulus values exhibit the greatest temperature dependency, as indicated by the wide spread of luminance output for the same input values. In contrast, as the stimulus values approach the extremes, the points cluster more closely together, indicating reduced variability in luminance output.

Additionally, the luminance output for the same input was consistently higher when the monitor was colder, decreasing as the monitor warmed up. A notable difference between our monitor and the one used by Schmiegel is that our monitor reached its maximum temperature (42°C) more rapidly, resulting in a greater proportion of data points being associated with higher temperatures (depicted in red on the graph) and fewer points in the ”colder” temperature range.

3.4 Consistency analysis across days

To evaluate the monitor’s consistency, measurements were repeated on different days. However, the results deviated significantly from expectations. While some level of consistency was anticipated, the monitor exhibited arbitrary and unpredictable behavior. As discussed in Section 3.1, the warm-up rate varied significantly across sessions, but this factor was addressed by modifying the methodology to incorporate real-time temperature readings before applying the LUT file.

It was critical for the monitor to demonstrate consistent behavior under identical conditions, meaning the same input values and the same internal monitor temperature. The graph in Figure 13 compares the overall measurements across two different days, with data from each day slightly shifted for clarity to prevent overlap.

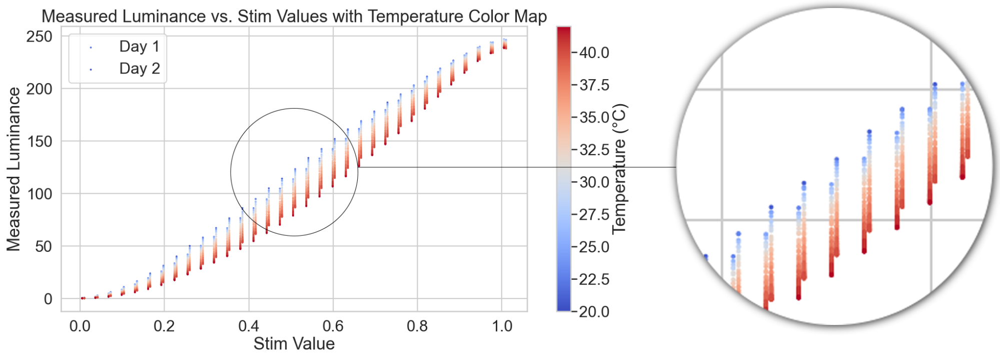

Figure 13: Comparison of measurements on two different days, with Stim values on the x-axis and measured luminance on the y-axis. The points on the graph are colored from deep blue to dark red, corresponding to the temperature of the monitor at the time of measurement.

Across the entire range of stimulus values, the results were inconsistent. Even at the extremes, where the impact of temperature variation was smaller, the data from different days displayed discrepancies. As expected, the largest inconsistencies occurred in the mid-range stimulus values, where the temperature dependency had the greatest influence. In the graph in Figure 14, which plots Temperature (°C) on the x-axis and Measured Luminance on the y-axis, the measurements from four datasets collected on different days are displayed. While the same reverse exponential trend is evident across all datasets, significant differences in the actual luminance values are observed. These discrepancies highlight the inconsistent behavior of the monitor, even under similar conditions.

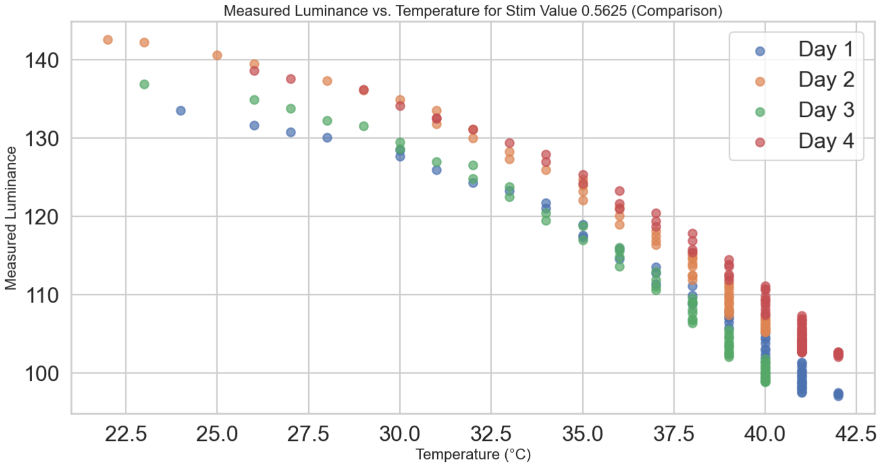

Figure 14: Temperature on the x-axis versus measured luminance on the y-axis across multiple days, with a Stim value of 0.5625. The plot highlights significant variance in the measurements, even at the same temperature.

To provide a contrasting example, the following graph in Figure 15 demonstrates a case of small variance, where the luminance difference across multiple days was at most 0.4 cd/m^2.

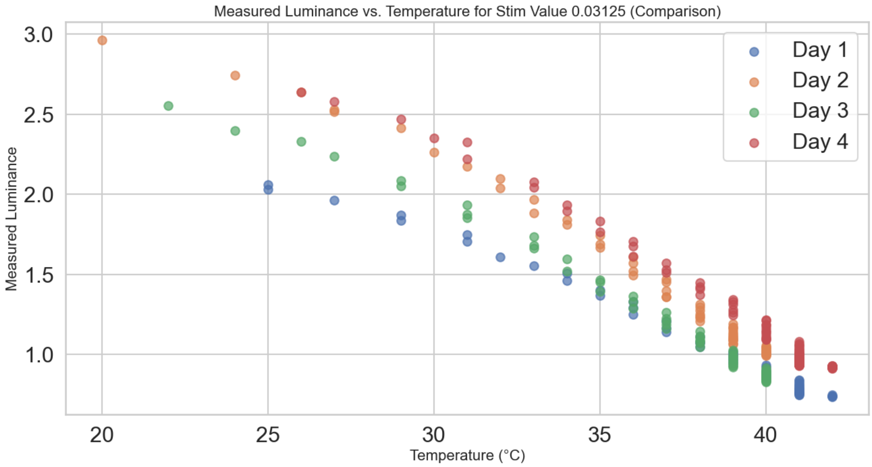

Figure 15: Temperature on the x-axis versus measured luminance on the y-axis across multiple days, with a Stim value of 0.03125. The plot shows low variance in the measurements, with a maximum difference of 0.4 cd/m^2 at 40°C.

The results demonstrate how the monitor's behavior changes with different stimulus values. At a mid-range stimulus value of 0.5625, the mean luminance across datasets varies significantly: 103.30, 106.08, 100.01, and 109.25 cd/m^2. The overall standard deviation is 3.71 cd/m^2, indicating considerable spread in the measurements. This variation suggests inconsistent monitor performance, likely due to the stronger influence of temperature at this range.

In contrast, at a low stimulus value of 0.03125, the mean luminance values are much closer: 0.89, 1.02, 0.87, and 1.15 cd/m^2. The overall standard deviation is only 0.12 cd/m^2, reflecting highly consistent measurements with minimal discrepancies. This stability suggests that the monitor is less affected by temperature at lower stimulus values.

In summary, the high standard deviation at 0.5625 indicates greater variability and inconsistency, while the low standard deviation at 0.03125 reflects stable and reliable behavior. These results highlight the significant impact of stimulus value and temperature on the monitor's response.

4 Discussion

This section addresses key insights gained from the findings and summary of results, comparing them to previous research, examines the study's limitations, and proposes directions for future research.

4.1 Summary of results

The results clearly demonstrate that the monitor's performance is significantly affected by temperature dependency. The output luminance (Lout) was consistently higher when the monitor was colder and decreased as the monitor warmed up. This temperature effect was most pronounced at mid-range stimulus (stim) values, while it had less impact at extreme stim values. This highlights the critical role of real-time temperature monitoring, as failing to account for these variations led to noticeable luminance fluctuations.

The warm-up period, ranging between 57 and 80 minutes to reach 42°C, proved unreliable as a fixed value, further complicating calibration efforts. Additionally, while the monitor showed more stability at lower input values, its overall performance remained inconsistent, even when setup and measurement parameters were carefully replicated across different days. This indicates that the monitor suffers from significant temperature dependency, making it essential to investigate other factors influencing luminance output and to develop more dynamic calibration strategies.

4.2 Impact on experiments

It is established that the monitor suffers from a temperature dependency, and with the new findings, we know that this is not the only influencing parameter. Previously, we assumed that allowing the monitor to warm up would stabilize its output. Although this happens to a great extent, the results are not always deterministic, as observed in Figure 14. Similar observations were made on different days: even when the monitor reached its maximum temperature and maintained it over time through repeated cycles and numerous measurements, the measured luminance, allthough limited to a small range, remained arbitrary.

In the worst-case scenario, we observed a variance in measured luminance with the same temperature and input value (input = 0.5625, temperature = 40°C) of up to 11 cd/m^2.

The human eye's ability to perceive brightness differences is explained by Weber's Law, which states that the smallest noticeable change in luminance is proportional to the current luminance level. In simple terms, our eyes detect changes in brightness based on the percentage difference relative to the background light level. This implies that a 10 cd/m^2 difference is more noticeable in a dark environment than in a well-lit one (Shen, 2002).

This creates a significant issue in our experiments, which are conducted in a completely blacked-out space. As a result, the human eye can perceive a 10 cd/m^2 difference in the same image during our experimental sessions.

To mitigate this effect, the monitor must be turned on for at least 90 minutes before any experiments are conducted. This serves as a temporary solution to minimize luminance variability until a robust calibration technique is developed.

4.3 Insights compared to previous research

There are several similarities and differences between the findings of this thesis and the work of Jonas Schmiegel. Both studies highlight that the gamma functions of the monitors suffer from significant temperature dependency, underscoring the critical influence of thermal conditions on luminance output.

Despite both studies focusing on the same monitor model (VIEWPixx /3D), the monitors themselves exhibited distinct behaviors. One key difference lies in the warm-up rate: Schmiegel reported that his monitor required up to 120 minutes to fully warm up, whereas in this study, the monitor consistently reached its maximum temperature within 90 minutes. Consequently, our monitor spent less time in the colder temperature range, resulting in a more rapid decrease in luminance output immediately after being turned on.

Another notable difference is the maximum temperature reached. Schmiegel's monitor achieved a peak temperature of 45°C, while our monitor only reached 42°C. This variation in thermal behavior further emphasizes the inherent differences between monitors of the same model and highlights the challenges of achieving consistent performance.

Finally, a critical difference is the inconsistency observed in our monitor. This inconsistency made it impossible to apply a Look-Up Table (LUT) for calibration purposes since the monitor operated with a different gamma function every day. Such variability meant that applying a fixed correction would not return meaningful results, making the LUT approach impractical for our experimental setup. As a result, a direct comparison of calibration efficiency between the two studies was not feasible. This introduces the limitations imposed by the unpredictable behavior of our monitor.

4.4 Study limitations

This study faced several limitations that influenced the findings. First, the ambient temperature varied from day to day and could not be fully controlled, either due to weather conditions or even the limitations of air conditioning, when compared to the research of our colleagues in Tübbingen.

Another significant limitation is the inconsistency between monitors of the same model. Even though the monitors were from the same manufacturer, they displayed different behaviors in terms of warm-up rates, maximum and minimum temperatures, and overall performance. Most notably, there was no consistency in the gamma function across different days, as discussed in Section 3.4. This was particularly clear with mid-range input values, where high deviations in luminance output were observed. This lack of predictability made it impossible to establish a standard for determining and predicting the monitor's gamma function at any given moment.

Finally, the number of measurements conducted may have also been a limiting factor. It is possible that a larger dataset could have led to different conclusions. However, based on the current measurements, the conclusion remains that the monitor is inconsistent and unpredictable, making calibration and standardization extremely challenging.

4.5 Future research directions and proposal

The results of this study highlight the need for improved calibration methods to address the significant impact of temperature dependency and the inconsistency observed in the monitor's performance. Future research should focus on developing strategies that can adapt dynamically to the monitor's current thermal state, rather than relying on assumptions about the starting temperature (TempMin) or determining when the monitor was turned on or last used.

A promising direction is the creation of Look-Up Tables (LUTs) constructed to specific temperature ranges. These LUTs would apply real-time corrections by measuring the monitor's temperature before applying the LUT to the input and presenting any stimuli. This approach would eliminate the need to assume a consistent starting temperature, which has been shown to vary depending on ambient conditions or prior usage. By linking luminance corrections directly to measured temperature values, this method could significantly reduce the variability caused by temperature changes.

Additionally, further research should explore how to address the inconsistencies observed in monitors of the same model. These differences in warm-up rate, maximum and minimum temperatures, and gamma function behavior make it challenging to apply universal calibration methods. Investigating how to standardize performance across monitors or account for their differences in calibration strategies is crucial.

Increasing the number of measurements is another important step for future research. A larger dataset could provide a more complete picture of the monitor's behavior across a wider range of conditions, helping to improve calibration models and improve their accuracy. Exploring advanced modeling techniques, such as machine learning and statistical regression, could also offer new ways to predict luminance behavior and account for variability.

Finally, a robust calibration process should address both temperature and luminance variability dynamically. Such a system would measure the current temperature, predict the expected luminance output, and apply adjustments in real-time to ensure consistency. By focusing on precision, future research can create the way for more reliable and efficient calibration methods, improving the accuracy of psychophysical experiments.

5 Conclusion

The findings of this study demonstrate that calibrating experimental monitors is a highly complex task. Initially, it was believed that temperature would be the primary influencing factor, and the plan was to determine the monitor's warm-up rate to establish a standard temperature at specific times after turning on. However, this approach proved unfeasible due to the inconsistent warm-up rates observed.

An alternative approach was developed, involving real-time temperature measurements before applying the Look-Up Table (LUT). Unfortunately, this method was also unsuccessful because the monitor's output remained inconsistent, even when key parameters such as input values and temperature were kept constant. Thus it would not be possible to create a standard for the LUT files and the application of them would be inefficient.

The primary limitation of using an uncalibrated monitor in experimental setups is the necessity of a warm-up period. Regardless of whether this period lasts 90 minutes or longer, it is essential for ensuring stable performance. Without this warm-up phase, the significant discrepancies between a cold and fully warmed-up monitor, as discussed earlier, could jeopardize the validity of experimental results.

List of Figures

References

Harding, S. (2021, May). *What is a monitor's gamma? a basic definition* [Last updated May 19, 2021]. https://www.example.com/what-is-monitors-gamma-definition

Lu, Z.-L., & Dosher, B. (2013, October). *Visual psychophysics*. MIT Press.

Schmiegel, J. (2023, November). *Characterisation of a high luminance resolution display for psychophysical experiments* [Master's thesis, University of Tübingen] [Submitted on 30.11.2023]. https://www.researchgate.net/publication/378306862_Characterisation_of_a_high_luminance_resolution_display_for_psychophysical_experiments

Shen, J. (2002). Weber's law and weberized tv restoration. https://ww3.math.ucla.edu/camreport/cam02-20.pdf