November 16, 2016

REPORT #E16-298

Revising the TV Energy Use Test Procedure: Incorporating HDR and other Needed Changes

Prepared by:
Ecos Research LLC
2188 SW Park Place, Suite 302
Portland, OR 97205

Northwest Energy Efficiency Alliance
PHONE
503-688-5400
FAX
503-688-5447
EMAIL
info@neea.org

Table of Contents

1. Executive Summary iv
2. Introduction 1
3. Methodology 2
4. Findings 3
4.1 Suitability of current DOE/IEC TV test procedure 3
4.2 Recommended near term policy revisions 6
4.3 HDR Overview 7
4.3.1 Key HDR Concepts 7
4.3.2 Impact on TV Energy Consumption 9
4.4 Factors to Consider When Developing a New TV Test Clip 11
4.5 CLASP Collaboration 14
4.5.1 Scene Cut Frequency Analysis 15
4.5.2 Color Depth Evaluation 15
4.5.3 Real World Content Comparison 16
4.5.4 Metrics Assessment 19
5. Conclusions 20
6. References 23
7. Appendix A: Models Tested 26
8. Appendix B: Test Equipment 26
9. Appendix C: Test Clips 27
10. Appendix D: Detailed Power Plots 28
11. Appendix E: Technical Background 30
11.1 HDR Technical Background 30
11.1.1 Video Pipeline 30
11.1.2 Color Grading 32
11.1.3 Transfer Functions 33
11.1.4 Color Mapping 35
11.1.5 HDR Compatibility 36
11.2 Evaluation of Picture Level Metrics 38
11.2.1 Theoretical Analysis of Approaches 38
11.2.2 Empirical Analysis of Approaches 41

Table 1. Limitations of the current DOE/IEC TV Test Method 4
Table 2. TV Policy Initiatives in Progress 5
Table 3. Power Results Table 9

Table 4. HDR Setting Changes Table 9
Table 5. CLASP Test Results from Retail TV Testing 10
Table 6. Scene Cut Frequency Analysis 15
Table 7. 10-bit vs. 8-bit HDR CLASP Test Clip Power Results 16
Table 8. Summary of Key Study Findings 21
Table 9: TV Models Tested 26
Table 10: Test Equipment List 26
Table 11. Test Clip List 27
Table 12. Example Calculation of Average Light Output Power 34
Table 13. Compatibility of Existing HDR Standards 37
Table 14. Preliminary Framework for Determining Number of Needed TV Test Clips 38
Table 15. Line of Best Fit R^2 Values 43
Table 16. Average R2 Values for Video Clips Tested 43
Table 17. Average R^2 Values by TV Technology Type 44

Figure 1. Histogram of Brightness Levels for HDR and SDR 8
Figure 2. Visual Illustration of Difference Between SDR and HDR Video 8
Figure 3. Illustration of Full Area Array Local Dimming 12
Figure 4. Power Impact of Aspect Ratio Scaling 14
Figure 5. 10-bit vs. 8-bit HDR CLASP Test Clip 16
Figure 6. Red Oaks in HDR vs. HDR CLASP Test Clip 17
Figure 7. HDR CLASP Test Clip vs. The Martian in HDR 18
Figure 8. HDR CLASP Test Clip vs. Ten Minutes of The Martian in HDR 18
Figure 9. Average Picture Levels of Sample Cell Phone Screen Images 19
Figure 10. LG 55EG9600 Power Measurements 28
Figure 11. Samsung UN55JS9000 Power Measurements 29
Figure 12. Samsung UN55JU7100 Power Measurements 29
Figure 13. Sony XBR-55X850C Power Measurements 30
Figure 14. RGB to Y'UV (= Y'CrCb) Conversion 32
Figure 15. Color Grading Example 33
Figure 16. Comparison of Typical HDTV and HDR Transfer Functions 35
Figure 17. Visual Illustrations of Color Volume and Color Space 36
Figure 18. Anatomy of a CIA Chromaticity Diagram 36
Figure 19. SDR to HDR APL' Comparison 39
Figure 20. Power Vs. AvgRGBLOP 42
Figure 21. Line of Best Fit and R2 Value 42
Figure 22. OLED Power and APL' over time 44
Figure 23. Edge Lit Power and APL' over time 45
Figure 24. R'G'B' 8-bit SDR Signal Levels Against Power for LG9600 (OLED) 46
Figure 25. AvgRGBLOP Histograms Assuming a 100 nit SDR Reference Display 47
Figure 26. AvgRGBLOP Histograms Assuming a 500 nit SDR Reference Display 48

Equation 1. R'G'B' to Y' Conversion Formula 31
Equation 2. HDR AvgRGBLOP 40
Equation 3. SDR AvgRGBLOP 41

1. Executive Summary

High definition televisions have improved remarkably in energy efficiency over the last decade, as labeling programs, incentives, and mandatory standards have steadily driven power levels downward. Large screen TVs that once consumed hundreds of kWh per year now routinely claim consumption of 75 to 150 kWh, while delivering better subjective picture quality and a host of new features and capabilities. However, it is important to understand how much of those apparent energy savings are the result of fundamental improvements to TV technology itself, vs. optimizing how TVs perform under the specific, necessarily artificial conditions of the test procedure. Part of that effort is to determine how much energy new television features consume that *do* routinely operate when televisions are viewed in the home, but may not be operating in the same way or measured at all during the test procedure.

NEEA recognized during the spring of 2016 that a substantial body of emerging research was pointing to serious flaws in the current US TV test procedure and energy use assumptions. As a result, NEEA commissioned this report, and the new research behind it, to understand how the test procedure could be improved, what qualities a new test clip should possess, and how much progress toward an optimal long-term test procedure could be made in the interim to inform current labeling and incentive programs.

ENERGY STAR®-labeled TVs typically enable an array of energy saving features in their default preset picture settings, such as automatic brightness control and motion detection dimming, to reduce screen brightness (and therefore power consumption) under particular circumstances. This yields impressively low energy use during the DOE test procedure, but energy use can be significantly higher under more realistic or typical viewing conditions and with other content viewed than the IEC test clip.

Modest improvements to the realism and representativeness of the current TV test procedure are possible by measuring televisions in additional, more power consumptive preset picture settings than the default (such as Vivid or Sports). Those values could be weighted with measured energy consumption in the default preset picture setting, depending on how persistent a TV's energy-saving features are, to come up with an improved estimate of its actual energy consumption in the field. Likewise, the energy savings from automatic brightness control could be measured or weighted in a slightly different manner to reduce the dominance of measurements under very dark viewing conditions, or the extent to which TVs dim their screens under those conditions.

Low-power mode energy use also deserves an update from the current procedure, primarily because it is now so common for new televisions to be connected to the Internet continuously once installed in the home. Also, because many televisions are downloading or uploading data and otherwise performing power-consumptive tasks when they might otherwise appear to their users to be switched off. That energy use is not being captured under the current test procedure.

Many of the shortcomings of the current test procedure are inherent to the test clip itself. Its scene cut frequency is unusually high, yielding higher energy savings from motion detection dimming than is typical when viewing other content. Likewise, it is not recorded in native 4K resolution, or encoded in 10-bit color or in any of the currently available high dynamic range formats. Each of those issues has the potential to boost power consumption as well, in some cases significantly.

This study included considerable effort to develop software tools that can analyze candidate video clips across a variety of attributes that are relevant to TV energy consumption. Software analysis revealed significant limitations in the ability of the average picture level (APL) metric to characterize video clips—APL predicted a surprisingly small degree of the second-to-second variation in power consumption in 2015 UHD televisions when displaying various test clips. If a new test clip is to be representative of real-world content, work is necessary to develop an accurate means to quantify brightness.

Detailed investigation of alternate metrics and the underlying processes used to encode standard dynamic range (SDR) and high dynamic range (HDR) video content revealed the difficulties of fairly comparing the relative brightness or APL of SDR and HDR content. This study lays the groundwork for the development of new and more comprehensive metrics for predicting the relative power consumption of TVs when playing a particular video clip (see technical appendices). The results will be useful long term with regards to HDR power consumption: as encoding becomes more common, performance specifications become standardized, more diverse types of content become available for analysis, and capability steadily makes its way into televisions of widely varying sizes and price ranges.

CLASP and its Europe-based researchers have developed a reasonably representative native 4K, 10-bit color, HDR-10 test clip that is appropriate for near-term testing efforts. Energy use measurements of four different 2015 UHD HDR TV models showed that the CLASP clip does a significantly better job of predicting typical TV energy use than the IEC clip does. The CLASP clip could therefore be useful for supplementing current TV testing efforts in the US and internationally, particularly to develop more realistic estimates of television energy use and savings when labeling and rebating 2016 and 2017 TV models.

Development of a new IEC test clip and associated DOE test procedure will require coordination among the international TV research community, and may take two to four years. That effort should include a comprehensive process to collect a representative sample of UHD video content from around the world, as was done more than a decade ago in the development of the current IEC test clip.

Updating the DOE test procedure will also benefit from further investment in analytic software to assess video clips. Standard test clips could be compiled more frequently than once every decade, to ensure they remain representative of typically viewed content. The energy efficiency community should also obtain or develop a handful of simple 4K, HDR-encoded test patterns that would allow quick comparison of UHD television performance and power consumption when screening candidate TV models for further testing.

Lastly, the energy efficiency community should continue to test and carefully assess noteworthy new 2016 UHD models as they are introduced. The UHD Premium industry specification and associated marketing efforts are likely to accelerate the race toward ever-brighter screens, with associated energy consumption impacts. Manufacturers are also continually and automatically updating the firmware on recently sold TV models, giving them new capabilities and changing their typical energy consumption patterns in ways that cannot be predicted from initial compliance testing alone.

2. Introduction

Television energy consumption has been measured in a largely consistent way within the United States for the last decade, based on the following process:

- Set up TVs for testing only in their default, out-of-the-box preset picture setting, with whatever energy saving features are enabled in that condition
- Play a standardized 10 minute IEC test clip
- Measure power consumption, second-by-second during that time, and average it
- Allow TVs with automatic brightness control (ABC) to reduce screen brightness as desired as the room gets darker and average the resulting power values as well
- Assume that average power consumption while the IEC clip is playing represents typical power use during the five hours each day that a typical TV operates
- Assume that typical power consumption during the other 19 hours per day is the same as a TV in the lab consumes when switched off and not connected to the Internet.

By this process and under these assumptions, high definition televisions have improved remarkably in energy efficiency over the last decade, as labeling programs, incentives, and mandatory standards have steadily driven power levels downward. Large screen TVs that once consumed hundreds of kWh per year now routinely claim consumption of 75 to 150 kWh, while delivering better subjective picture quality and a host of new features and capabilities.

But it is important to understand how much of those apparent energy savings are the result of fundamental improvements to TV technology itself, vs. optimizing how TVs perform under the specific, necessarily artificial conditions of the test procedure. As part of that effort, we also need to understand how much energy certain features and capabilities of new televisions consume that *do* routinely operate when televisions are viewed in the home, but may not be operating in the same way or measured at all during the test procedure.

NEEA recognized during the spring of 2016 that a substantial body of emerging research was pointing to serious flaws in the current US TV test procedure and energy use assumptions. As a result, NEEA commissioned this report, and the new research behind it, to answer the following questions:

- What are the key limitations in the current DOE/IEC TV test procedure that need to be addressed in the near term to support ongoing labeling and incentive programs?
- How effectively does the CLASP HDR test clip address the limitations of the current IEC test clip?
- How might the CLASP HDR test clip be utilized in the near term to supplement our understanding of TV energy use and performance?
- Which key technical issues are too complex or uncertain to be addressed in the near term, and may need to be considered through a longer term, multi-stakeholder IEC process?

Some of this report's policy recommendations involve detailed technical analysis that is impractical to document without the introduction of complex concepts and technical jargon. HDR video involves a more complex video production, distribution, and display ecosystem than prior standards like high definition TV (HDTV). In order to improve the readability of this report, the body of the report summarizes the key concepts, while Appendix E: Technical Background provides discussion targeted at technical members of policy development teams. NEEA submits this technical content for consideration by the expert teams assembled by the IEC and government policymakers, whose combined industry experience and technical knowledge outweighs that of the research team that assembled this report.

3. Methodology

This research approach used in this study involved the following steps.

1. Provide an overview of HDR technology based on secondary research and summarize empirical test results that demonstrate how much more energy TVs use when measured using the CLASP UHD HDR test clip.

   This research step involves comparing the measured power levels of four TVs (Appendix A: Models Tested) in the Ecos Research lab across three different video clips:

   a) The current HD IEC test clip

   b) An HD clip developed by Ecos Research for the Natural Resources Defense Council (NRDC) that represented current real world content with typical scene cut frequency, and

   c) An ultra-high definition (UHD) high dynamic range (HDR) test clip developed by the CLASP for use as an interim test clip while IEC stakeholders work to develop an international standard test clip[1].

   Specific attributes of test clips used in this study are listed in Appendix C: Test Clips.

2. Identify the broader challenges associated with developing a new test clip that is representative of real world content. Ecos Research performed this task through secondary research, lab experimentation and phone calls and emails with industry experts.

---

[1] Most of the power measurements conducted for this study are based on the 8-bit color depth version of the CLASP HDR test clip, as opposed to the newly released 10-bit version. Limited testing suggests that 8-bit verses 10-bit color depth has little impact on resulting power measurements as discussed in section 4.5.2 of this report, so we did not repeat tests conducted with the 8-bit clip.

3. Engage with CLASP and European consultants and policymakers to evaluate the use of the CLASP test clip on an interim basis. This involved the following steps:

    a) Measure the scene cut frequency of the CLASP UHD HDR test clip and compare it to the scene cut frequency of U.S. real world content.

    b) Assess the power impact of 8 vs. 10-bit color depth.

    c) Compare CLASP test clip power profiles to those of real world HDR content.

    d) Develop software capable of analyzing average picture level and other metrics for UHD HDR video clips.

    e) Assess several brightness-related metrics for use to ensure that future HDR test clips are representative of real world content in terms of brightness.

    Ecos Research measured scene cut frequency by documenting the video time stamps associated with scene changes and then organized the results in charts and tables.

    To develop software, Ecos contracted with a software vendor, VideoQ, to customize its video analysis software, VQPLA, and make it available for distribution to international standard and policy development teams.

4. Findings

4.1 Suitability of current DOE/IEC TV test procedure

Standardized energy test procedures for televisions and other devices serve several purposes. They enable repeatable, comparable energy consumption measurements that support policies such as regulation, mandatory and voluntary labeling, and incentive programs. Results of standard laboratory tests, combined with sales figures and field usage data, also allow estimation of regional and national energy use and savings.

The ideal test procedure would be easy to conduct, fair to all manufacturers and products, and representative of the energy consumed once consumers actually use the devices. The real world necessitates compromise, however. To minimize testing burden for manufacturers and regulators, standardized test procedures often involve a few hours' worth of data collection with default device settings and a single assumed usage pattern or duty cycle. This approach works well when test conditions are representative of real world use or when power consumption of a device tends not to vary much with user behavior. However, this is no longer the case for televisions shipping in 2016, which include far higher pixel density, brighter screens, new content formats, a dizzying array of potential picture settings, and entirely new methods of content delivery than those from a decade ago. The DOE/IEC test method—and in particular, the video clips that the tests are based on, which are the main focus of this study—need to be updated for the tests to serve their intended purpose. Table 1 summarizes the key points that support this conclusion.

Table 1. Limitations of the current DOE/IEC TV Test Method

Test Clip (Focus of this study)	The frequency of the current IEC test clip's scene cuts is much higher than in real world content, a fact that results in significantly lower power measurements during testing than real world use. A feature generically called motion detection dimming (MDD) reduces TV brightness when playing content with frequent scene cuts. (Horowitz 2016)
	The current IEC test clip is formatted as high definition (HD), but new ultra-high definition (UHD) content is available that offers higher resolution, high dynamic range (HDR), an expanded color gamut, and often greater color depth. These features tend to increase TV energy consumption significantly. (Horowitz 2015)
Test Configuration	The energy saving features in today's TVs are often set aggressively in the TVs' default, as-tested configuration to reduce reported energy consumption. However, even a small user change to brightness or contrast settings—in some cases a few clicks of the remote control—can disable these features with no warning or user notification. This lack of settings persistence means that test conditions are not representative of real world use. Field software updates also reduce settings persistence. For example, a field upgrade can add HDR capability for some models. (Horowitz 2016)
	TV users increasingly access content using smart TV apps, like those offered by Netflix or Vudu. Today's HDMI-enabled TV's are tested using only the HDMI port for input. (Alder 2014)
	The DOE test method requires testing with no internet connection, which is not representative of real world use of today's newly purchased smart TVs and may affect TV energy consumption. (NPD 2016; NRDC 2015)
Other	The DOE test method approach to automatic brightness control (ABC) power measurement may not yield representative values, given the equal weighting of power measurements taken at 3, 12, 35, and 100 lux ambient light levels and the lack of a 300 lux test condition. (Leritz 2016) CLASP and Intertek are exploring other, improved approaches to determining real world energy consumption with ABC enabled.
	The DOE standby test provides instructions to wait until the TV has stabilized before taking power measurements; however, some TVs use a significant amount of energy before stabilizing in standby. (Horowitz 2015)

Notes: This research focuses on gaps having to do with the International Electrotechnical Commission (IEC) test clip, referenced by the DOE TV test method.

Given these limitations, the energy efficiency community could elect to continue using the test procedure as is, with all usage assumptions unchanged, and focus its efforts solely on persuading national and international organizations to establish a new test procedure. Indeed, several international standards organizations, government bodies, and non-governmental organizations (NGOs) are engaged in efforts to update TV power measurement approaches. Table 2 summarizes these stakeholders and activities.

Table 2. TV Policy Initiatives in Progress

IEC	Initiated several workgroups to update the TV test method and associated test clip. Released technical report on HDR TV power measurement. CLASP, EPA, NRDC participating. (International Energy Agency 2016)
DOE	Initiated TV test method rulemaking with 6/24/2016 request for information (RFI). RFI comments are posted. (Department of Energy 2016)
ENERGY STAR®	Announced data collection effort and October version 8 kick-off webinar on 8/12/2016. (Environmental Protection Agency 2016)
California Energy Commission (CEC)	Has not commented on the timeline for updating California's TV energy efficiency regulation.
EU Commission	Updating Ecodesign TV Minimum Efficiency Performance Standard (MEPS) and energy label.
Collaborative Labeling and Appliance Standards Program (CLASP)	Engaged in Ecodesign MEPS and label updates. Developed new test clip in UHD SDR and UHD HDR formats. (CLASP 2016)
Consumer Technology Association (CTA) and Digital Europe	Engaged in the development of IEC standards and regulations.
Natural Resources Defense Council (NRDC)	Published studies of UHD TV energy consumption (Horowitz 2015) and MDD and settings persistence (Horowitz 2016). Provided public comment in response to DOE RFI.
The Northwest Energy Efficiency Alliance (NEEA)	Conducted research into HDR standards development and TV technology trends. Provided public comment in response to DOE RFI.

It is important to note in order to achieve comparable TV power measurements across international borders, energy efficiency policymakers generally rely on IEC standards. The relevant test standard for TVs is IEC 62087, "Methods of measurement for the power consumption of audio, video and related equipment" (International Electrotechnical Commission 2015) and the associated IEC test clip for TV power measurement. The DOE test method, used by ENERGY STAR and required by law, references IEC 62087 and the IEC test clip. This report shows that rapid innovation in video formats and TV technologies has outstripped the effectiveness of the current standard testing method for televisions: the U.S. Department of Energy (DOE) TV test method, based on the International Electrotechnical Commission (IEC) Method 62087, originally devised in 2006.

However, there are meaningful near-term changes to be made, on an interim basis, to improve the realism and representativeness of measured test results. Mathematical adjustments are also possible to change how measured values are weighted or considered in determining annual energy consumption. By individually assessing the various ways in which TV technology and content have changed over the last decade, and their impact on energy consumption, we were able to make a set of targeted recommendations to CLASP while it was developing an interim UHD HDR test clip. We were also able to conduct a series of measurements with their test clip that validated its usefulness in near term TV testing, as detailed below. Lastly, we were able to monitor ongoing policy developments and research findings in the US and internationally, and make strategic adjustments to the scope and focus of our research as it became clear that the EU, US EPA, and US DOE were all considering various ways to refine and update their measurements of television energy use.

Because this study was investigatory, its findings include descriptions and details of several technical aspects of television content and display—information that may be well known to TV designers, but not widely known to the energy policy community. Sections 4.2 4.5 below go into some depth to explain HDR, Average Picture Level, color depth, and other nuances of modern television. We then relate these factors to power consumption and implications for test clips and procedures. Appendix E provides extended detail on some of these topics.

4.2 Near term policy revisions considerations

NRDC found that energy saving features like automatic brightness control (ABC) and motion detection dimming (MDD) often get switched off in preset picture settings other than the default. They also noted that the specified backlight levels and other picture parameters are often set higher in preset picture settings other than the default setting. Manual changes by the user to any of those settings can often cause televisions to shut off optional energy saving features without warning to the user.

Even while relying on the IEC test clip, it is possible to conduct tests in other, more power-consumptive preset picture settings as well, and to weight those results with the measurements in default preset picture settings in some fashion to reflect the persistence and degree of applicability of energy saving features in each TV model.

Likewise, the energy savings from automatic brightness control could be measured or weighted in a different manner to reduce the dominance of measurements under very dark viewing conditions, or the extent to which TVs are allowed to dim their screens under those conditions. Many TVs operate at such low luminance levels when ABC is enabled and room conditions are quite dark, that resulting power levels are extremely low, but may not satisfy users accustomed to a much brighter picture.

Finally, there are good reasons to address low power mode consumption differently than the US test procedure currently does, primarily because it is now so common for new televisions to be connected to the Internet continuously once installed in the home. Many televisions are downloading or uploading data and otherwise performing power-consumptive tasks when they might otherwise appear to be off. That energy use is not being captured under the current test procedure.

Labeling and incentive programs would obtain more realistic values for TV energy use and savings from making these changes to current test procedures and calculations, but there are still deep and fundamental flaws in the IEC test clip itself that need to be addressed, as detailed below.

4.3 HDR Overview

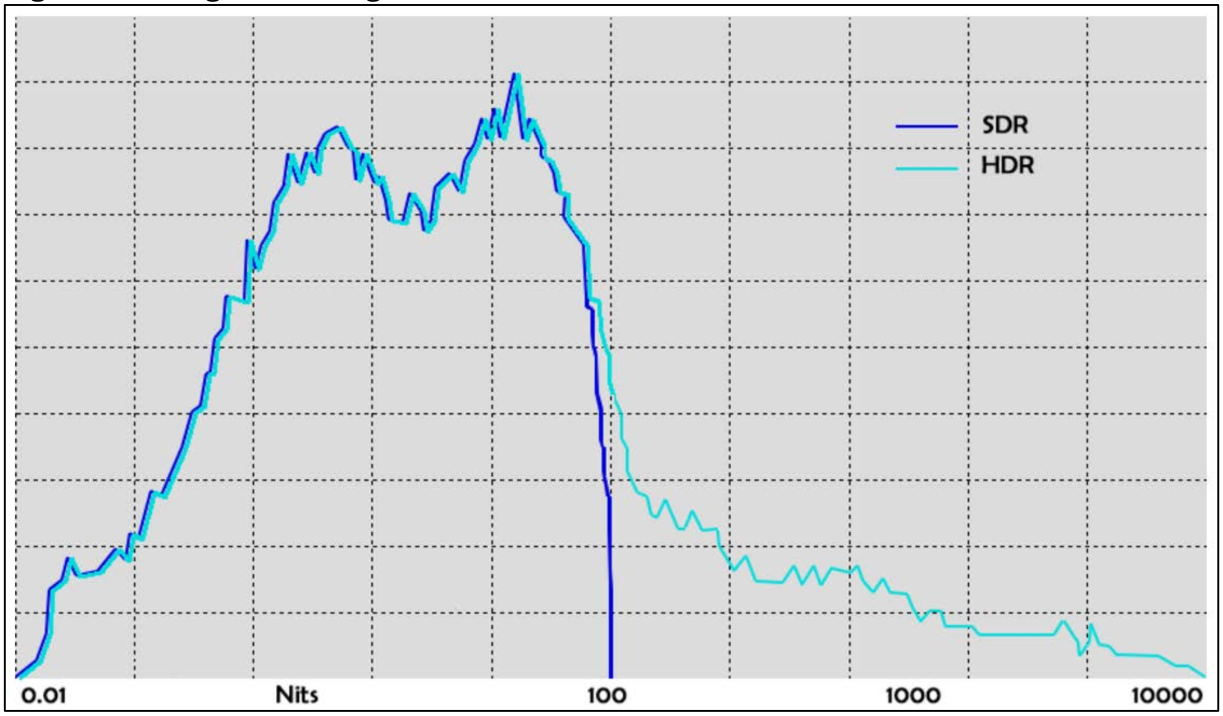
High-Dynamic Range technology enables TVs to deliver brighter highlights and greater discernible detail in both bright and dark areas of the screen, creating an almost three-dimensional feel. HDR requires both specially encoded content and HDR-capable televisions; displaying this format also uses significantly more energy than SDR content.

4.3.1 Key HDR Concepts

There are at least three distinct varieties of HDR content. This discussion focuses on HDR 10 since it is the most widely adopted by TV manufacturers and content providers; see section 11.1.5 for discussion of other HDR standards. **Figure 1** compares histograms of HDR and SDR brightness levels after the TV converts signal to screen luminance (i.e. cd/m^2, commonly referred to as nits). This graph shows that SDR content ranges from zero to 100 nits, but the HDR standard allows highlights (such reflections on shiny surfaces or emissive objects like the sun) up to 10,000 nits—brighter by two orders of magnitude.

Today's HDR LCD TVs cannot yet display 10,000 nit brightness—they reach 500 nits (e.g. Samsung JS9000) to 1,500 nits (e.g. Samsung KS8500) (Sony 2016; Dermers 2016). This raises the issue of how HDR TVs respond to brightness signals beyond their effective range. For example, the CLASP HDR clip has highlights that extend up to about 2,000 nits. Conceptually, a display with peak brightness of 1,000 nits would clip the right end of the curve in **Figure 1** so that all of the values above 1,000 would be displayed at 1,000 nits. However, the HDR 10 specification allows for "roll-off", which avoids a sharp cut off at 1,000 nits by enabling a gradual transition. This might seem an unimportant detail, but a sharply clipped video loses detail in the bright parts of the frames and is unacceptable to many viewers, so TV manufacturers typically use roll-off. Lab viewing of roll-off vs cutoff showed a dramatic difference in experience—the latter looked significantly degraded.

Figure 1. Histogram of Brightness Levels for HDR and SDR



Note: This histogram shows that HDR signals call for much brighter highlights than SDR technology. (Shaw 2016)

Figure 2 depicts the difference one might perceive between an SDR image and an HDR one. The difference between the two pictures is the addition of brighter picture levels (e.g. in the sky) and the added detail associated with the bright parts of the image. The power consumption associated with displaying the darker parts of the image is unchanged, but can be significantly higher in the brightest parts of the image.

Figure 2. Visual Illustration of Difference between SDR and HDR Video

Notes: Retrieved from (Shaw 2016)

For more in-depth technical background on HDR technology, refer to Sony 2016 and section 11.1 of this report.

4.3.2 Impact on TV Energy Consumption

TVs use more energy when playing HDR content. **Table 3** compares power levels for the four TVs tested in this study with the current IEC test clip (baseline), the HD test clip developed for a recent study by NRDC, and the UHD/HDR CLASP test clip. TVs tested with the CLASP clip used 30 to 80 percent more energy than when tested with the IEC HD SDR clip. Neither the CLASP nor the NRDC clips trigger MDD to any significant extent, so the difference between the baseline and real world content is largely due to the effects of MDD and UHD.

Table 3. Power Results Table

Clip and characteristics	LG 9600		SAM 9000		SAM 7100		Sony 850C	
	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase
IEC Test Clip (HD, SDR, 8-bit (Baseline)	136		125		100		108	
NRDC Test Clip (HD, SDR, 8-bit	142	4.4%	144	15.2%	125	25.0%	104	-3.7%
CLASP Test Clip (UHD, HDR, 8-bit)	175	28.7%	213	70.4%	178	78.0%	161	49.1%

Notes: Default Picture Setting with ABC Off. NRDC Real World HD SDR 8-bit Content developed for NRDC 2016. Detailed charts are included in Appendix D: Detailed Power Plots.

To better understand the power increase associated with HDR content, Ecos documented the preset picture setting (e.g. Vivid, Standard, and Cinema) and picture parameter (e.g. brightness, contrast, backlight) changes associated with HDR content in **Table 4**. For example, each television in the study makes automatic settings changes when HDR content is displayed. The two Samsung models both increase their backlight to 20, the maximum possible setting. The Sony 850C automatically switches to a preset picture setting called "HDR Video," where ABC is disabled and the brightness setting is increased to its maximum. The LG 9600 is the only TV in which a message appears explicitly alerting the user that HDR content has been recognized. When streaming video from an app such as Amazon, the LG preset picture setting reads "HDR (app)" an unlisted picture setting. However, when the LG displays HDR content from a USB source, no explicit picture setting change occurs, but the HDR alert remains and power levels rise. These setting changes are presumably responsible for much of the increases in power seen in **Table 3**.

Table 4. HDR Setting Changes Table

	HDR Setting Changes			
	Preset Picture Setting	**Picture Parameters**	**ABC**	**HDR Alert**
SAM 9000	Unchanged	Backlight automatically jumps from 18 to 20 with no notification	Unaffected	None
SAM 7100	Unchanged	Backlight automatically jumps from 12 to 20	Greyed out but remains	None

		with no notification	unchanged	
Sony 850C	Automatically changes to "HDR Video" and is greyed out	Brightness automatically changes from 30 (exact middle) to "Max" (60) with no notification	Disabled and greyed out with no notification	None
LG 9600	When streaming HDR content from an app, preset picture setting automatically changes to "HDR(app)" and is greyed out HDR content does not affect preset picture setting when played through USB	Cannot access picture parameters in "HDR(app)" mode Unaffected when content played through USB	Unchanged	"HDR is on" message is displayed (both USB and streaming)

Likewise, it is possible to compare television models playing the IEC test clip, the CLASP clip in SDR format, and the CLASP clip in HDR format, to understand the incremental power consumption jump associated with each. CLASP performed such an analysis on an assortment of models in British retail stores in the early summer of 2016, and shared those results with us as we were conducting parallel measurements on TV models in the US. As summarized in **Table 5** below, it is evident that HDR-capable UHD TVs tend to consume an average of 9% more power when playing a more realistic test clip than IEC, even when both clips have standard dynamic range, primarily because they do not accrue major power reductions from motion detection dimming. Moreover, power consumption increases by another 42% when moving from the SDR to HDR versions of the same UHD CLASP test clip, giving us some additional confirmation of the power consumption premium associated with HDR when nothing else is changed about the underlying content. ABC was not enabled for any of this British testing.

Table 5. CLASP Test Results from Retail TV Testing

TV Model Tested	Screen Diameter (inches)	IEC (watts)	CLASP UHD (watts)	CLASP UHD Increase over IEC	CLASP UHD-HDR (watts)	CLASP HDR Increase
Model A	50	77	79	2.6%	103	30.4%
Model B	55	131	132	0.8%	163	23.5%
Model C	55	70	102	45.7%	163	59.8%
Model D	55	132	130	-1.5%	138	6.2%
Model E	65	194	206	6.2%	448	117.5%
Model F	65	102	114	11.8%	175	53.5%
Model G	65	160	157	-1.9%	168	7.0%
Average	**59**	**124**	**131**	**9.1%**	**194**	**42.5%**

4.4 Factors to Consider When Developing a New TV Test Clip

Through this and related research, Ecos has identified several key factors that policymakers should take into consideration when developing a new TV test clip.

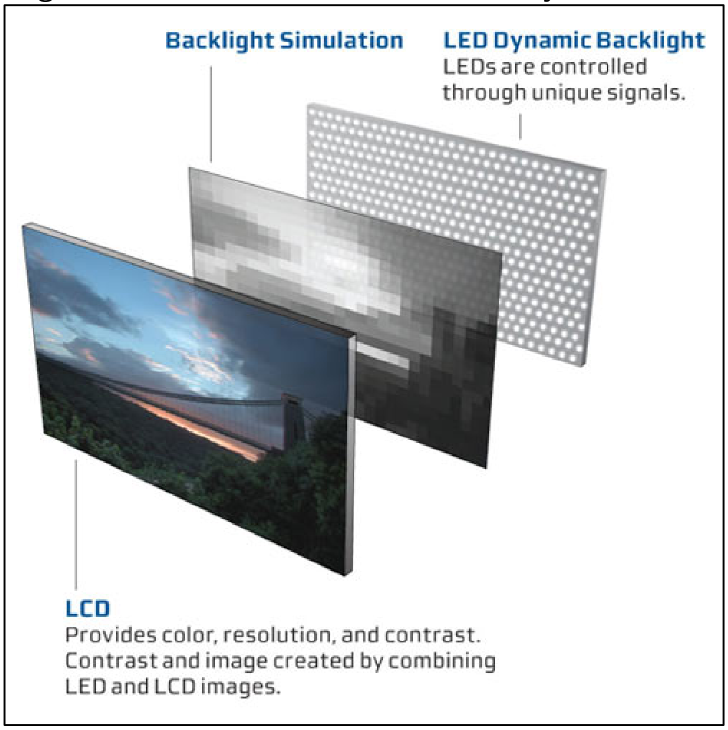
Luminance as a function of time: Test results suggest that frequent luminance changes—associated with scene changes and certain types of motion—trigger MDD. Therefore, luminance as a function of time should be representative of real world content. The analysis performed for this report focuses on scene changes; however, a more sophisticated, software-based approach could assess luminance as a function of time more comprehensively.

Picture Level / Light Output Power:
It is important to develop a video test clip that is representative of real world content in terms of its impact on backlight power consumption, since this is the most energy consumptive component of most TVs. This topic is explored further in sections 4.3.4 and 11.2.

Color: Color width, depth, and distribution impact TV power consumption. A new test clip should be representative across these variables.

Spatial distribution: Many of today's light-emitting diode (LED) liquid crystal display (LCD) TVs offer local dimming capability to increase contrast ratio and reduce power consumption. Local dimming is most effective for TVs with LED full area arrays as shown in **Figure 3** below. However, the more affordable and popular edge lit LED LCDs can offer a similar capability with fewer dimmable zones, defined by the number of edge light elements. In either case, one bright pixel in any given dimmable zone results in the associated LED element operating at high power. So, if every dimmable zone has just one or a few bright pixels among a sea of dark pixels, then the backlight array would operate at high power even though the picture level and light output would be low. The international community should evaluate the need for software tools that enable it to ensure that the next IEC test clip is representative in terms of spatial distribution.

Figure 3. Illustration of Full Area Array Local Dimming



Notes: Obtained from (Mochileiro Digital 2011)

Audio: International stakeholders should assess whether or not new audio technologies, such as object-based audio (Weary 2014), warrant inclusion in future test clips, because they result in the use of additional processing power.

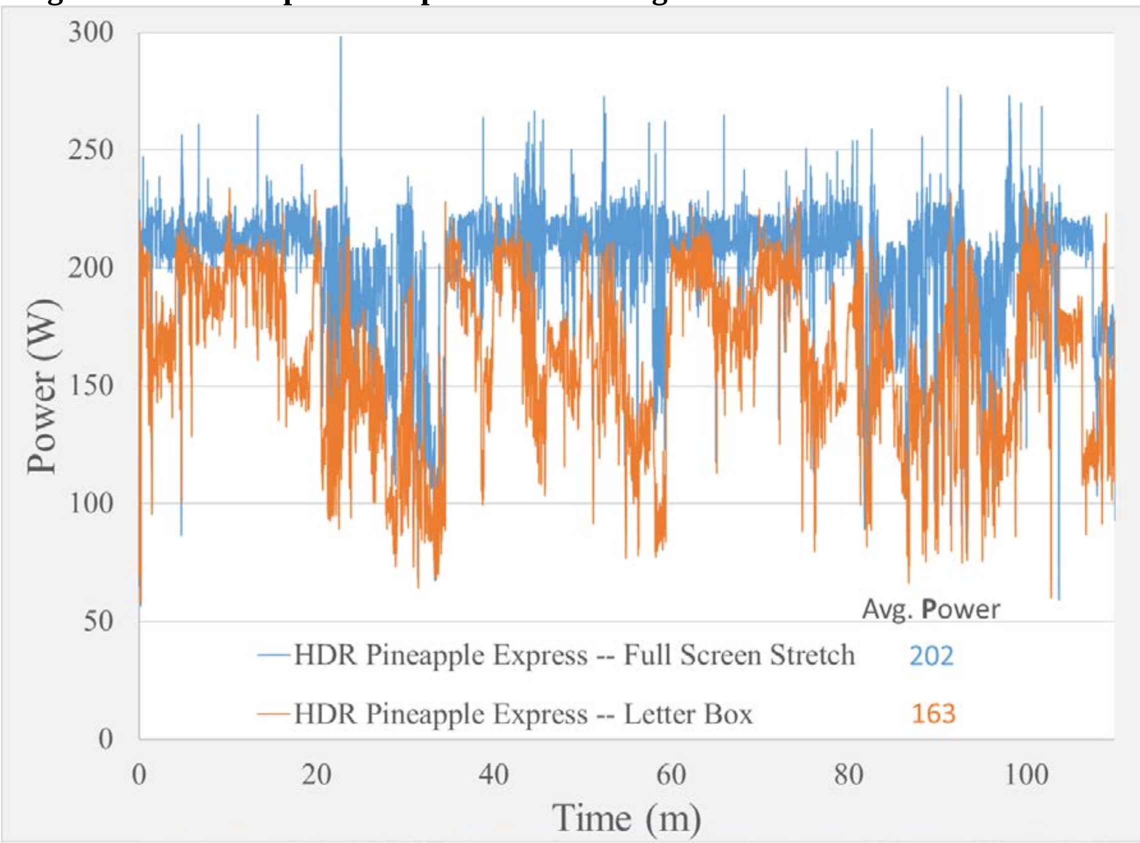
Content availability: Most of today's HDR content is available in Blu-ray format or as streamed video (e.g. Netflix and Vudu). The content is heavily skewed towards movies and TV series with little news or sports available. It is difficult to obtain legal copies of this content in a format that can be either a) distributed as part of a test clip or b) analyzed for the purpose of creating a representative test clip. Furthermore, best practices for color grading and image processing HDR video content have not yet developed, so today's real world content may not be representative of HDR content once de facto approaches evolve. For these and other reasons, it is important to engage with content creators. Content owners may be more likely to engage if video analysis tools provide insight into picture level and other important test clip development metrics while guarding the confidentiality of other data that content owners wish to guard.

At this time, it is not possible to analyze typical commercial HDR video content with VQPLA, the software tool used in this study. VQPLA is intended to work with digital video files. It is capable of calculating picture level and light output for Blu-ray discs with the use of a personal computer peripheral Blu-ray player; however, there are no PC Blu-ray players on the market than can play commercially sold HDR Blu-ray movies and shows. Likewise, it is not possible to analyze either streamed HDR movies or HDR content broadcast by major pay-tv service providers, due to copy protection on the underlying data.

Multiple versions and compatibility: There are two versions of today's TV test clip, standard definition (SD) and high definition (HD). There may need to be more versions of the next IEC test clip, depending on how interoperable the various HDR standards are with available TVs. Determining the appropriate number of versions of the next test clip is complicated because different HDR standards are used for different types of content (e.g. live video vs. off-line content). It appears that IEC stakeholders may be able to develop a test clip in Dolby Vision format that can be written onto a UHD Blu-ray disc that is compatible with both HDR 10 and SDR TVs—see section 11.1.5 of this report for more discussion on this topic. In this case, IEC may offer three or four test clip versions: SD, HD, UHD, and perhaps 8K. However, additional testing must be done to confirm the compatibility assumptions behind this approach. For example, this research project revealed compatibility issues when trying to play BT.2020 HDR 10 content on some 2015 HDR displays.

Other considerations: The international community should consider whether the next IEC test clips should include movie content, which is often darker and sometimes displayed in letterbox aspect ratio. Power levels are lower when playing movies in letterbox aspect ratio on TVs with local dimming, since the top and bottom backlights can be off (**Figure 4**). Closed captioning can also increase TV energy consumption. These factors may be (or may not be) important enough relative to warrant inclusion in future IEC test clips.

Figure 4. Power Impact of Aspect Ratio Scaling



Notes: TV Model - Samsung UN55JS9000; TV Setting Parameters - Default Picture Setting with ABC Off. Full-screen stretch aspect ratio was 16:9.

4.5 CLASP Collaboration

At the start of this study, the research team determined that the most impactful way to influence the development of an updated test clip was to collaborate with CLASP on this activity and to add value by:

1. Assessing whether or not the scene cut frequency of the CLASP test clip is representative of real world content,
2. Measuring the impact of color depth on power consumption,
3. Comparing CLASP test clip power profiles to those of real world HDR content,
4. Developing software capable of analyzing average picture level and other metrics for UHD HDR video clips, and
5. Assessing the value of these tools and the metrics they generate.

This section of the report summarizes key findings of these research tasks.

4.5.1 Scene Cut Frequency Analysis

The CLASP test clip is roughly representative of real world scene cut frequency. The CLASP test clip scene cut frequency, ten cuts per minute, is lower than what Ecos measured for U.S. real world content, fifteen cuts per minute (**Table 6**). CLASP chose to adopt a level lower that fifteen based on the results of European real world content scene cut measurements. A broader assessment of scene cut frequencies across geographies is required for the development of an updated IEC test clip.

Table 6. Scene Cut Frequency Analysis

	CLASP Test Clip	IEC Test Clip	NRDC U.S. Real World Content	NRDC U.S. Real World Content Without Commercials	NRDC U.S. Real World Content Commercials
Clip Length (min)	10	10	12	10	2
Total Scenes	97	262	185	108	77
Average Scene Length (sec)	6	2	4	6	2
Total Cuts	96	261	184	107	76
Cuts per min.	**10**	**26**	**15**	**11**	**38**

Notes: Table adapted from NRDC 2016. CLASP HDR test clip is added. Real World Content is twelve minutes long with commercials appearing for last two minutes.

4.5.2 Color Depth Evaluation

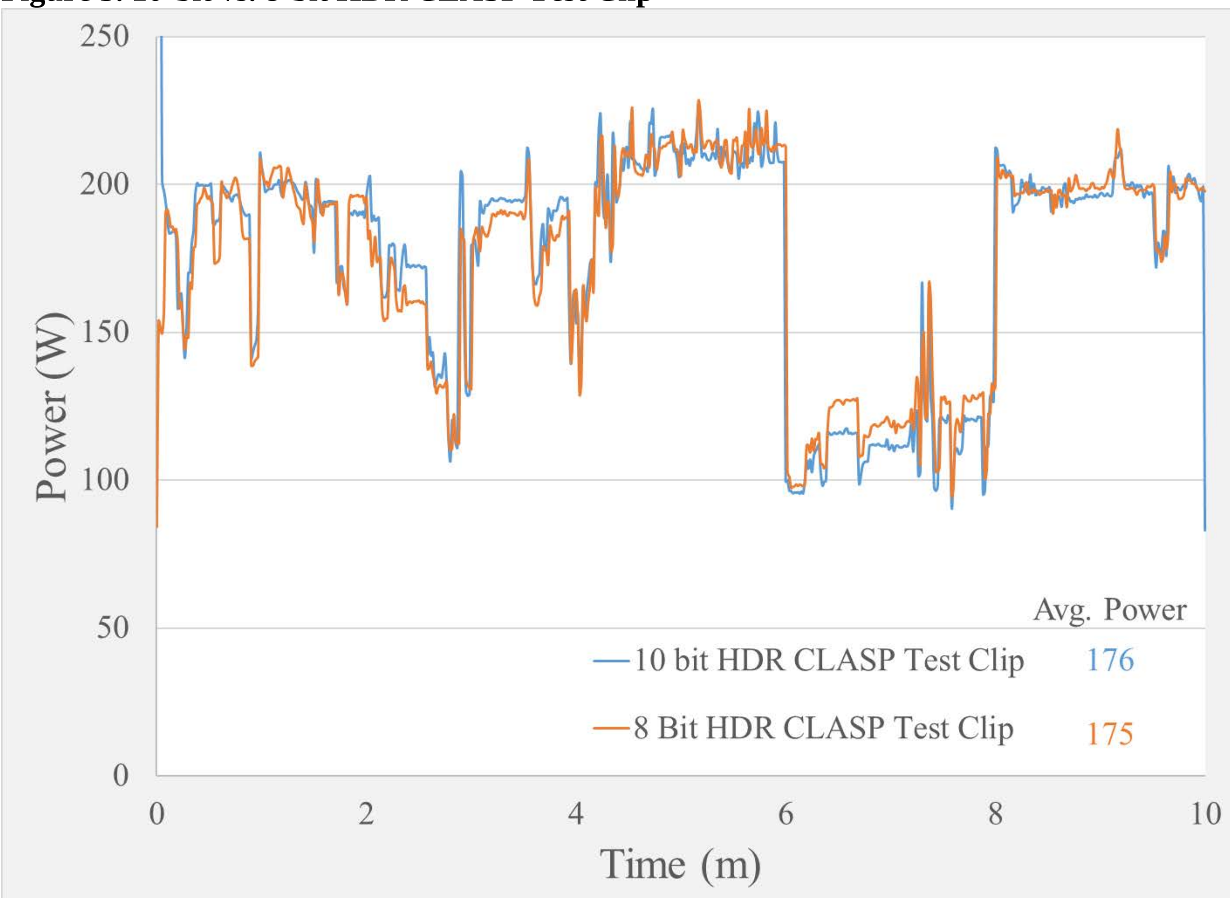
Color depth refers to the number of bits used to represent each primary color component of a pixel. Greater color depth allows for more precise rendering of color. Previously 8-bit had been the standard color depth in video files, but HDR technology now commonly uses 10 and 12-bit color depth. Ten-bit video can specify 1024 (2^{10}) distinct shades each of red, blue, and green, while 8-bit video can only specify 256 (2^{8}) distinct shades of red, green, and blue. Color depth alone should not affect the brightness of a video file, but color depth in HDR video enables viewers to see much more image detail, resulting in a more engaging viewing experience. (Sony 2015)

Early versions of the HDR CLASP Test Clip used in this study had a color depth of 8-bits. CLASP developed a 10-bit version of the clip later in this study. Through tests of both clips, Ecos determined that the 10-bit clip required a small amount of additional power to display on the TVs tested, but not enough to warrant repeating the 8-bit tests with the 10-bit clip.

Table 7. 10-bit vs. 8-bit HDR CLASP Test Clip Power Results

	LG 9600		SAM 9000		SAM 7100		Sony 850C	
	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase	Avg. Power (W)	Power Increase
8-bit CLASP Test Clip	175		213		178		161	
10-bit CLASP Test Clip	176	0.6%	215	0.9%	182	2.2%	161	0.0%

Figure 5. 10-bit vs. 8-bit HDR CLASP Test Clip

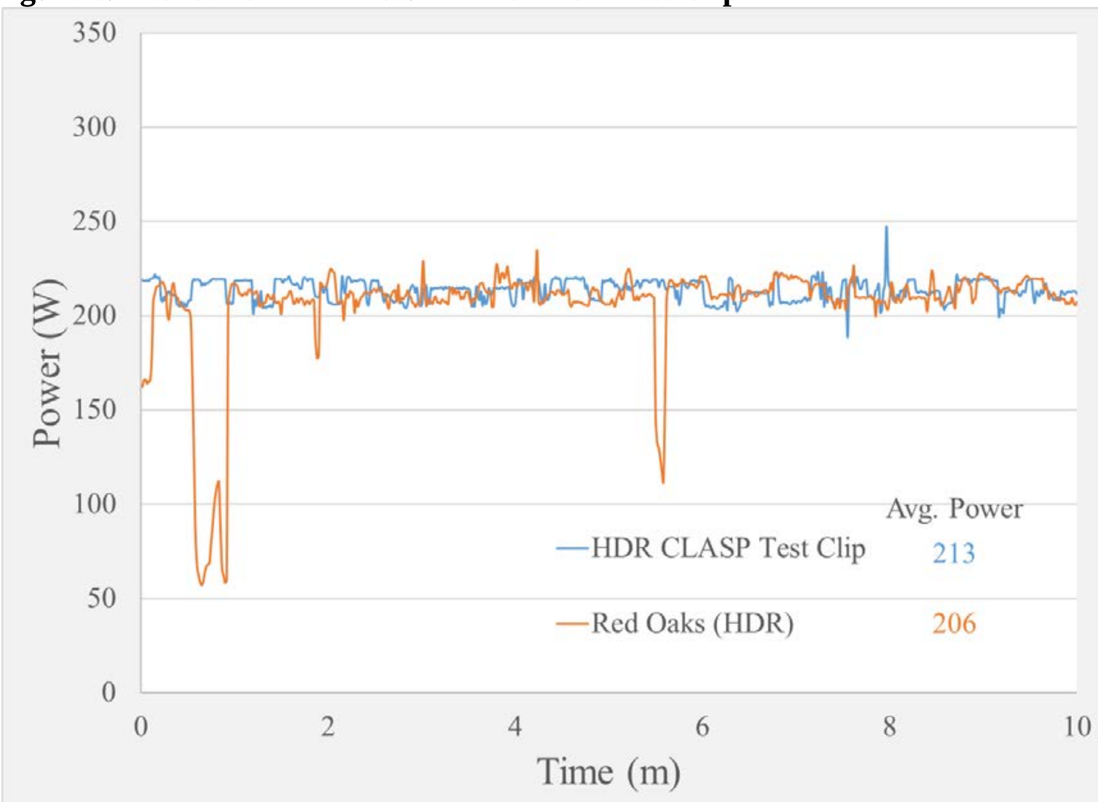


Notes: TV Model – LG 55EG9600 (OLED). TV Setting Parameters – Default Picture Setting with ABC Off. Organic Light Emitting Diode (OLED) displays are emissive displays, meaning that their subpixels each generate light independently.

4.5.3 Real World Content Comparison

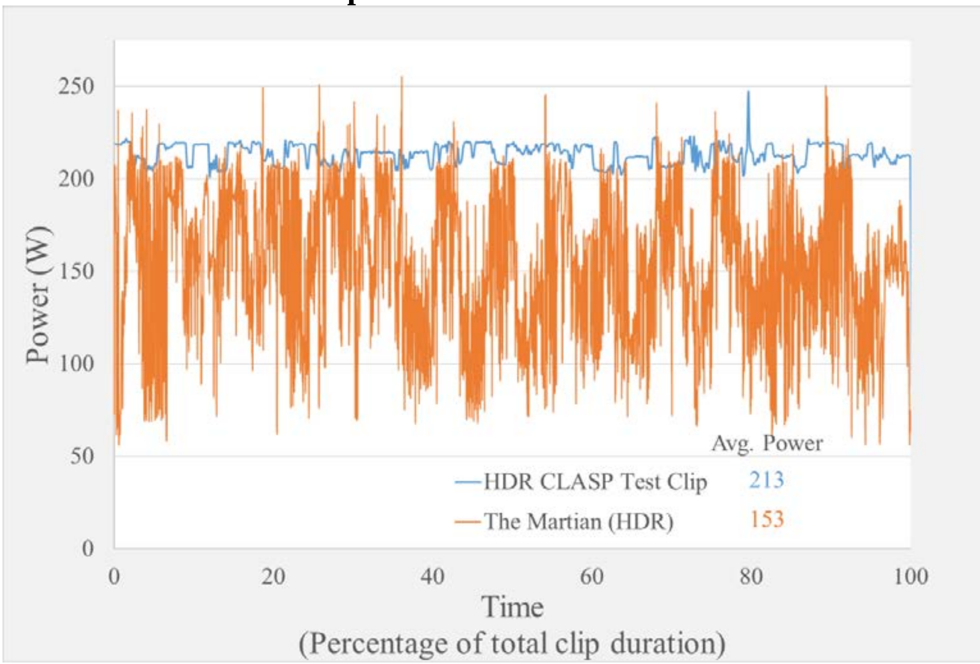
To better understand to what extent the CLASP test clip is representative of real world content, Ecos compared the power profile of the CLASP HDR test clip, as tested on the Samsung JS900, to two samples of professionally produced HDR content, an Amazon produced drama called *Red Oaks* and a Hollywood movie called *The Martian.* The power profile of the CLASP HDR test clip is similar to the profile for *Red Oaks* (**Figure 6**), but not to *The Martian* (**Figure 7**), which has darker backgrounds and dramatic rises and drops in brightness over time. These differences demonstrate the high degree of variability that exists in the color grading of HDR content and the need for further study of a wide variety of HDR content as HDR color-grading practices evolve over time.

Figure 6. Red Oaks in HDR vs. HDR CLASP Test Clip



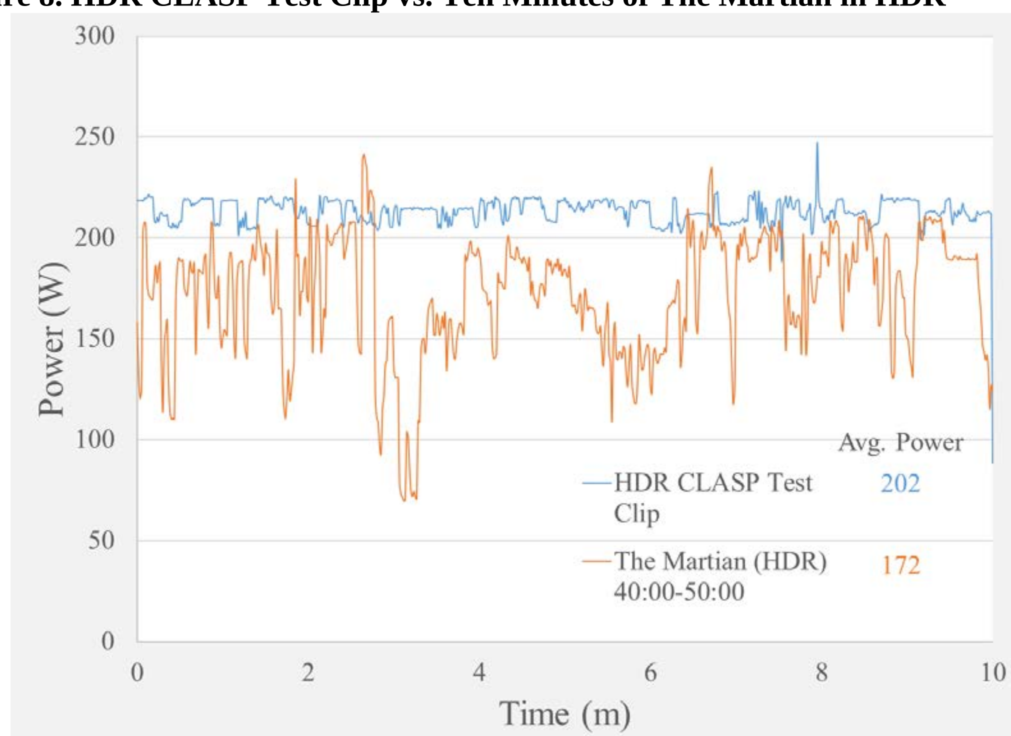
Notes: TV Model – Samsung UN55JS9000. TV Setting Parameters - Default Picture Setting with ABC Off. Red Oaks streamed from Amazon Prime Instant Video.

Figure 7. HDR CLASP Test Clip vs. *The Martian* in HDR



Notes: TV Model – Samsung UN55JS9000. TV Setting Parameters - Default Picture Setting with ABC Off. The Martian was played via UHD Blu-Ray. TV is edge lit with local dimming. The entire duration of The Martian is represented compared to the 10 min CLASP test clip.

Figure 8. HDR CLASP Test Clip vs. Ten Minutes of The Martian in HDR

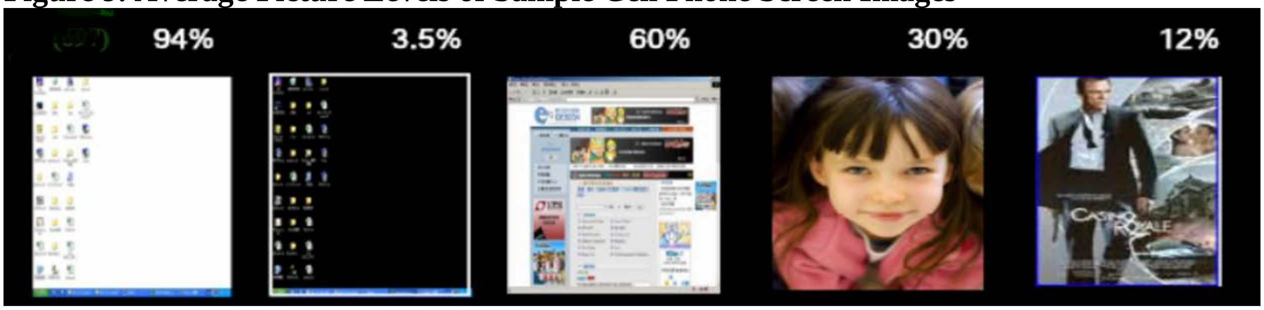


Notes: TV Model – Samsung UN55JS9000. TV Settings Parameters – Default Picture Setting with ABC Off. The Martian was played via UHD Blu-Ray.

4.5.4 Metrics Assessment

It is important to develop a video test clip that is representative of real world content in terms of its impact on backlight power consumption since a typical TVs backlight is its most energy consumptive component. The development of the current IEC TV test clip involved using a metric known as gamma-corrected average picture level (APL') to ensure that the test clip was representative of real world content in terms of APL' and APL' distribution, a key determinant of backlight power. An SDR video frame or clip in which all of the pixels are pure white has an APL' of one-hundred percent while an SDR frame or clip in which the pixels are pure black would have an APL' of zero percent. **Figure 9** illustrates the concept by showing several images of cell phone screen images and their associated APLs.

Figure 9. Average Picture Levels of Sample Cell Phone Screen Images



Notes: Image from (4D systems 2008).

Not all TV technologies scale power to APL' as proportionally as others, but APL' has been an important metric because it tends to correlate to power; a test clip designed without efforts to ensure that the test clip's impact on backlight power is representative would lead to power measurement results that do not reflect typical power levels. For example, because the red, green and blue sub-pixels in an OLED display generate their own light, OLED TV power tends to scale to APL' to a larger degree than in an LCD TV the backlight of which remains lit at a fairly consistent level regardless of average pixel brightness. This is true of emissive displays in general, which are displays the subpixels of which emit their own light as opposed to displays in which the subpixels operate primarily by blocking light generated by a white backlight.

LED LCD displays are the top selling variety in the non-emissive display category. They operate by filtering white light generated by LED backlights and then using liquid crystals to block light at the subpixel level to achieve the right mix of colors and shades. Many LED displays scale the brightness of individual LEDs to picture level to achieve a better contrast ratio resulting in an improved correlation between picture and power levels. The fact that power does not scale perfectly with APL' and that different TV technologies scale power to APL' to different extents is not important. Power *tends* to scale to APL' and similar brightness or light output metrics, so it has been important to develop a video test clip that has an APL' value that is representative of real world viewing so that TV test results are representative of real world use. And it remains important to have a metric that will enable IEC stakeholders to ensure that the next IEC test clip has a representative impact on TV backlight power via the use of APL' or a similar metric.

However, APL' cannot be compared between SDR and HDR content, so histogram comparisons between the CLASP HDR test clip either the SDR IEC clip or real world SDR content are invalid for the reasons stated in section 11.2 of this document. Other metrics discussed in section 11.2 provide theoretical advantages and some level of comparability. However, empirical studies documented in section 11.2.2 of this report reflect relatively poor correlation between all of the metrics evaluated and TV power consumption, likely due to the existence of automatic brightness limiting (ABL) and other non-linear factors discussed in in section 11.2.

Therefore, this report recommends that technical committees review all three metrics for use in developing the next IEC test clip. In the short term, the most practical approach may be to work with content providers to collect hours of real world HDR content and compare the APL' of this content to proposed interim test clips. Studios may need to provide special access to content files that are capable of being analyzed by the picture level and light output analysis tools.

5. Conclusions

Recent advances in TV technology have brought higher resolution, greater color depth and realism, higher contrast, and other improvements that make TV more compelling to watch. However, these improvements come at the cost of incremental energy consumption. International policymakers and standards bodies have begun the task of updating TV test methods to measure the energy TVs consume when playing HDR content and to ensure that TV power measurements better reflect energy consumption during typical use. The findings from this study point the way toward a comprehensive test method revision, as well as more modest, interim revisions that can be applied quickly to voluntary labeling and incentive programs.

Table 8 below summarizes the key findings of this study. The leftmost column lists the array of broader factors to consider when developing the next IEC test clip. The middle column assesses the suitability of the CLASP HDR test clip for interim use. The rightmost column lists actions needed to further the development of the next generation IEC test clip and versions thereof.

While the CLASP HDR clip does not fully address all issues that will be important to consider in the long-term IEC clip development process, it dramatically improves on the current IEC test clip. The CLASP test clip is much more representative than the current IEC test clip in terms of scene cut frequency, and it includes early UHD HDR content that produced TV power profiles comparable to those measured from a popular US drama (Red Oaks). As such, it is a welcome addition to the tools available to policymakers and program administrators now as they seek to more accurately characterize television energy use and savings.

Table 8. Summary of Key Study Findings

Broad list of factors	Suitability of CLASP HDR clip for interim testing	Actions needed for future IEC test clip
Luminance over time (e.g. scene cut freq.)	Scene cut frequency approximately representative of Europe and U.S. content.	Benchmark scene cut frequency across major international TV markets.
Picture level	Similar power profile to HDR drama *Red Oaks* but different from the film *The* Martian. Light output power histogram analysis suggests that it is generally representative.	Acquire and analyze video content internationally and further develop software tools.
Color	Needs further study.	Further study effects of color distribution, width and depth.
Spatial distribution of bright pixels	Not able to analyze with existing software tools.	Acquire and analyze video content internationally. Develop tools and methods.
Audio	Not likely to be a major determinant of energy consumption.	Further study.
Content availability	Need to engage content studios.	Work with studios to acquire content that can be analyzed via software.
Content evolution	Color grading methods and tools are evolving, meaning “representative” is a moving target. CLASP clip was developed with today’s tools, so is appropriate for use on an interim basis.	Better understand when methods will stabilize.
Multiple versions and compatibility	The CLASP HDR clip is encoded in HDR 10 format, an open standard and the most common format in the marketplace.	Further study compatibility claims and verify with tests before determining what array of IEC test clips to develop.
Movie content, letterbox aspect ratio, closed captioning	Not prevalent enough to warrant inclusion in an interim test clip.	Study whether these are important enough factors to warrant inclusion in test clip development effort.

Much work remains to coordinate the methods, findings, and conclusions presented in this study with technical working teams organized by standards bodies or government agencies. The broad range of factors listed above need further study, and the metrics and tools used to assess proposed test clip content relative to real world content need to be assessed by technical experts familiar with all aspects of HDR content development, distribution and display. Energy efficiency advocates should engage with IEC workgroups and in government rulemaking, and continue to hone tools and methods in preparation for those engagements. Specific tasks include:

1. Update the VQPLA software or an equivalent software tool to assess the average picture level and light output of video clips or frames at a pixel level to enable policymakers to understand with greater precision the differences between SDR and HDR video streams

and the likely impact of backlight power. This deeper understanding will also enable policymakers to develop more appropriate tools and methods for ensuring that future test clips are representative of real world content.

2. Develop or obtain a set of standardized SDR, HDR 10, and Dolby Vision test patterns for purposes of quickly measuring and comparing luminance and power consumption across various TV models and preset picture settings.
3. Conduct measurements similar to those conducted here on a wider variety of 2016 UHD TV models, including those employing newer display technologies, multiple HDR modes, and updated software controlling the persistence of energy saving features, to better understand the representativeness of these findings with the latest TV models.
4. Collect additional samples of popular UHD HDR content from around the world that can be analyzed by video software to ensure wide variation in picture levels and representativeness of sample content from which a final IEC test clip can ultimately be assembled.
5. Participate actively in the IEC and DOE test procedure development processes.

6. References

Alder, Emily. 2014. TVs Are Becoming the Next App Battleground. Business Insider. Retrieved from: http://www.businessinsider.com/the-smart-tv-app-industry-4-2014-2

Cineramax, Peter. 2015. Fox's first 4 UHD titles to include HDR mastering Non Dolby good for Beamer HDR. Digital Cinema Initiatives. Accessed from http://dci-forum.com/index.php?topic=362.0.

Collaborative Labeling and Standards Program (CLASP). 2016. New Video Test Sequence for Televisions. Retrieved from http://clasp.ngo/Resources/Resources/Headlines/2016/New-Video-Test-Sequence-for-Televisions

Department of Energy (DOE). 2016. Energy Efficiency Program: Test Procedure for Televisions; Request for Information. Washington, D.C.: Department of Energy. Accessed from https://www.regulations.gov/docket?D=EERE-2016-BT-TP-0023

Dermers, Cedric. 2016. Peak Brightness of TVs: Max luminosity and HDR highlights. Rtngs TV Reviews. Retrieved from http://www.rtings.com/tv/tests/picture-quality/peak-brightness

Digitalmediasuite. 2012. Color Correction 101: Examples of Color Profiles in Hollywood Blockbusters. Digitalmediasuite. Accessed from https://www.youtube.com/watch?v=OS6Tw5UT-8s&feature=youtu.be.

Discovery Scientific. 2015. YUV, YCbCr, YPbPr colour spaces | DiscoveryBiz.Net http://discoverybiz.net/enu0/faq/faq_YUV_YCbCr_YPbPr.html.

Environmental Protection Agency (EPA). 2016. ENERGY STAR Version 8.0 Specification Development Launch and Data Assembly. Retrieved from https://www.energystar.gov/products/spec/televisions_specification_version_8_0_pd

Gates, Chris. 2013. Color Grading in Post Prodction. Videomaker. Accessed from https://www.videomaker.com/article/c3/15710-color-grading-in-post-production.

Horowitz, Noah. 2015. The Big Picture: Ultra High-Definition Televisions Could Add $1 Billion to Viewers' Annual Electric Bills. San Francisco. Natural Resources Defense Council (NRDC). Retrieved from: https://www.nrdc.org/sites/default/files/uhd-tv-energy-use-report.pdf

Horowitz, Noah. 2016. TBD report on MDD and persistence. San Francisco. Natural Resources Defense Council (NRDC). Retrieved from: TBD.

International Electrotechnical Commission (IEC). 2016. TA 12 Subcommittee(s) and/or Working Group(s). Retrieved from http://www.iec.ch/dyn/www/f?p=103:29:0::::FSP_ORG_ID,FSP_LANG_ID:5681,25#1

International Electrotechnical Commission (IEC). 2015. Audio, video, and related equipment – Determination of power consumption - Part 1: General, Part 2: Signals and Media, Part 3: Television Sets. Frankfurt, Hesse: International Electrotechnical Commission. Accessed from https://webstore.iec.ch/publication/22581.

International Telecommunications Union (ITU). 2015. Parameter Values for Ultra-High Definition Television Systems for Production and International Program Exchange. Accessed from: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.2020-2-201510-I!!PDF-E.pdf.

Leritz, Nick. 2016. Comments from the Northwest Energy Efficiency Alliance on the DOE Request for Information on the Television Test Procedure. Portland, OR. Northwest Energy Efficiency Alliance (NEEA). Retrieved from https://www.regulations.gov/document?D=EERE-2016-BT-TP-0023-0003

Limoor, Josh, and Vincent, Tim. 2016. How do I Tell my Story in HDR. Technicolor. Accessed from https://www.youtube.com/watch?v=x1gPyMsPoYs&feature=youtu.be.

LionDoc. 2012. Accessed from https://commons.wikimedia.org/wiki/File:CCD.png.

Mochileiro Digital. 2011. LED Dynamic Backlight. Retrieved from http://www.mochileirodigital.com.br/wp-content/uploads/2011/07/LED_DynamicBacklight1.jpg.

NPD. 2016. 49 Million U.S. Internet Homes Now Own a Connected TV or Attached Content Device. NPD. Retrieved from: https://www.npd.com/wps/portal/npd/us/news/press-releases/2016/49-million-us-internet-homes-now-own-a-connected-tv-or-attached-content-device-according-to-the-npd-group/

Shaw, Steve. 2016. What you need to know about UHDTV - HDR, HLG & WCG. Sioux Falls, SD: Murideo News. 7/8/2016 entry. Retrieved from http://www.murideo.com/news.

Schulte, Tom, and Barsotti, Joel. 2016. HDR Demystified Emerging UHDTV Systems. Shoreline, WA: SprectraCal. Retreived from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0ahUKEwjBwvntrcnOAhUN-GMKHSWwDQMQFgg8MAU&url=http%3A%2F%2Ffilecache.drivetheweb.com%2Fmr5vpo_showstoppers_nab%2F178426%2Fdownload%2FHDR_Demystified.pdf&usg=AFQjCNHC8cK2XMK4YqhRmJY7vq7wc5unxA.

Sony. 2016. HDR (High Dynamic Range). Tokyo: Sony. Accessed from https://pro.sony.com/bbsccms/assets/files/cat/hdr/latest/MK20109V1_1_HDR.pdf.

Volpatto, Walter. 2016. NAB16 Walter Volpatto (P1) from Fotokem on grading HDR and ACES. CineTechGeek. Accessed from https://www.youtube.com/watch?v=Ouja_kfAM5g&feature=youtu.be.

Weary, Jay. 2014. Audio Objects for 4KTV. TV Technology. Accessed from http://www.tvtechnology.com/audio/0014/audio-objects-for-4ktv/222964.

Wesson, Philip. 2014. Color Grading, LUTS and FilmConvert. Phil Wesson. Retrieved from http://www.philwesson.com/blog/2014/7/27/luts-and-color-grading.

Yurek, Jeff. 2012. Color space confusion. Milpitas, Ca: Nanosys. Retrieved from https://dot-color.com/tag/cie-1931/.

Zink, Michael. 2015. MANAGING HDR CONTENT PRODUCTION AND DISPLAY DEVICE CAPABILITIES. Burbank, CA: Warner Bros, USA; Wavelet Consulting LLC, USA. Accessed from http://www.ibc.org/__media/PDF/michael_zink_managing_hdr_content_production_and_display_device_capabilities.pdf.

4D Systems. 2008. Introduction to OLED Displays Design Guide for Active Matrix OLED Displays. Retrieved from http://www.ruf.rice.edu/~mobile/elec518/readings/display/4D_AMOLED_Presentation.pdf.

7. Appendix A: Models Tested

The four TV models tested are detailed in **Table 9** below:

Table 9: TV Models Tested

Data	LG 9600 2015 55"	Samsung 7100 2015 55"	Samsung 9000 2015 55"	Sony 850C 2015 55"
MFG	LG	Samsung	Samsung	Sony
Model	55EG9600	UN55JU7100	UN55JS9000	XBR-55X850C
Shorthand	LG 55EG9600	SAM UN55JU7100	SAM UN55JS9000	SNY XBR55X850C
Resolution	4K	4K	4K	4K
Backlight	N/A	LED Full Array	LED Edge Lit	LED Edge Lit
Panel Technology	OLED	---	Quantum Dot	Quantum Dot
Local Dimming	N/A	Yes	Yes	No
Date of MFG	Jan-15	Feb-15	Feb-15	Apr-15
Serial Number	501RMDZ71930	03K33CYG201591M	03UB3CCG200608K	5009168

8. Appendix B: Test Equipment

For all power measurements, a Chroma 61602 reference power source provided a stable AC waveform to the TV. A Yokogawa WT-310 analyzer interfaced to a custom designed LabVIEW control panel logged power measurements in real time. Ecos configured test equipment and tolerances per the Department of Energy (DOE)/IEC standards with a sample rate of 250 milliseconds (ms) for all measurements. During lab-based ABC testing, Ecos used a Konica T-10A meter to measure room illuminance. All test equipment (**Table 10**) held valid calibration certificates at the time of test and met or exceeded tolerances defined by the DOE Final Rule 78 FR 63823.

Table 10: Test Equipment List

Type	Make	Model	S/N	Calibration Date
AC Power source	Chroma	61620	616020002628	04/22/16
Power analyzer	Yokogawa	WT310-D-C1/G5	C2RA21009V	04/19/16
Blu-ray player	Oppo	BDP-103	B242U01435045057	N/A
HDR blu-ray player	Samsung	UBD-K8500	0ABN1RCH305908T	N/A

9. Appendix C: Test Clips

Ecos was able to play the CLASP HDR test clip only via USB memory stick connected directly to the TV in some cases. In these cases, some of the TVs had default settings that are different than the default settings for HDMI input. In order to maintain comparability of results for these cases, Ecos manually set the TV settings to the defaults for HDMI input when testing with the CLASP HDR clip on USB stick.

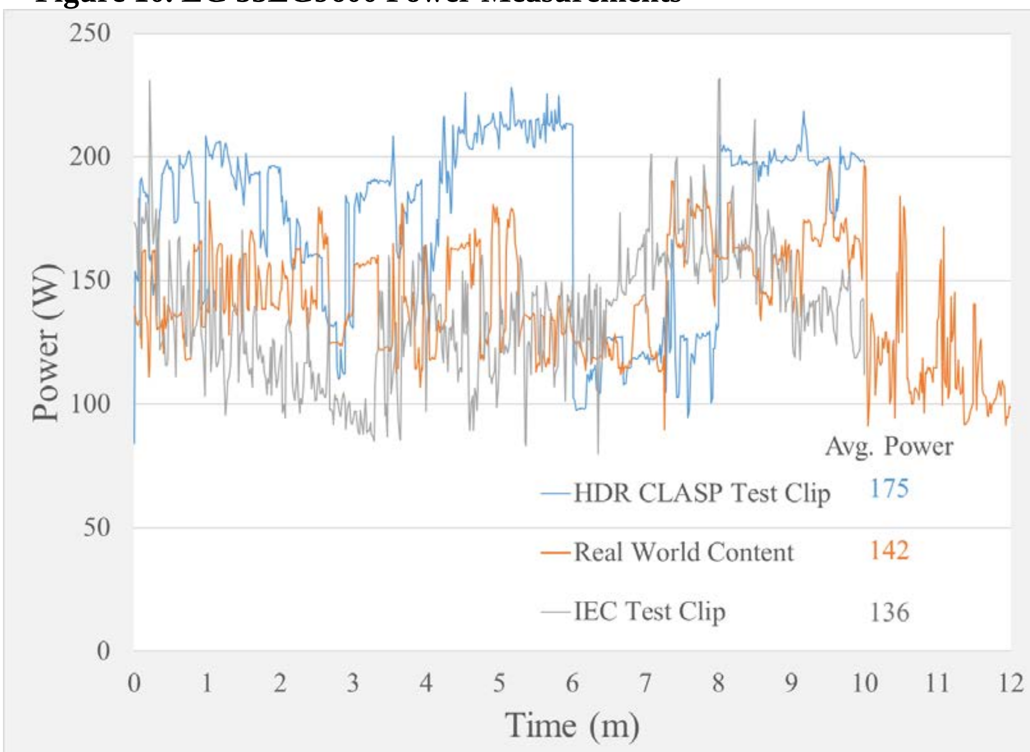
Table 11. Test Clip List

	CLASP Test Clips			IEC Test Clip	Real World Content	Ecos Scene Cut Video
Dynamic Range	HDR 10	HDR 10	SDR	SDR	SDR	SDR
Resolution	2160p	2160p	2160p	1080i	1080p	1080p
Color Depth	10-bit	8-bit	8-bit	8-bit	8-bit	8-bit
Transfer Function / Gamma Curve	ST2084		BT.709	BT.709	BT.709	BT.709
Clip Length (min)	10	10	10	10	12	1/6
Average Scene Length (sec)	6	6	6	2	4	1
Use	Color depth analysis	Primary HDR test results	Histogram analysis	Power & scene cut frequency measurements and histogram analysis	Power & scene cut frequency measurements and histogram analysis	Scene cut frequency analysis
Video Input Method	USB memory stick connected directly to TV			DVD in Blu-ray player with HDMI 2.0a cable	USB memory stick connected directly to TV	Laptop connected via HDMI cable
Source	(CLASP 2016)			IEC	Ecos for NRDC	Ecos

10. Appendix D: Detailed Power Plots

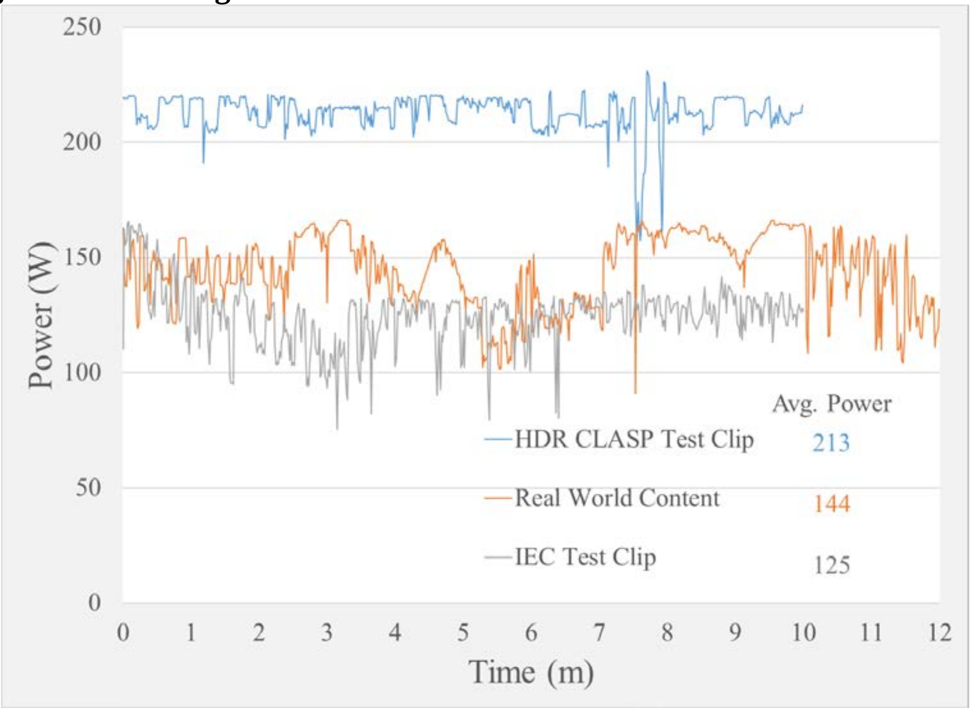
Figure 10 through **Figure 13** below show the power vs. time plots that are the basis for the average power values shown in **Table 3**. The average power value for each power measurement is shown next to the associated legend entry in the same color as the associated data series. While the IEC and CLASP test clips are ten minutes long, the real world clip is twelve minutes long: ten minutes of typical television show content and two minutes of commercials. An upcoming NRDC report will includes additional details about the type of content included in the real world test clip. The power increases reflected in these tests represent a combination of factors including the effects of MDD and 4K HDR content.

Figure 10. LG 55EG9600 Power Measurements



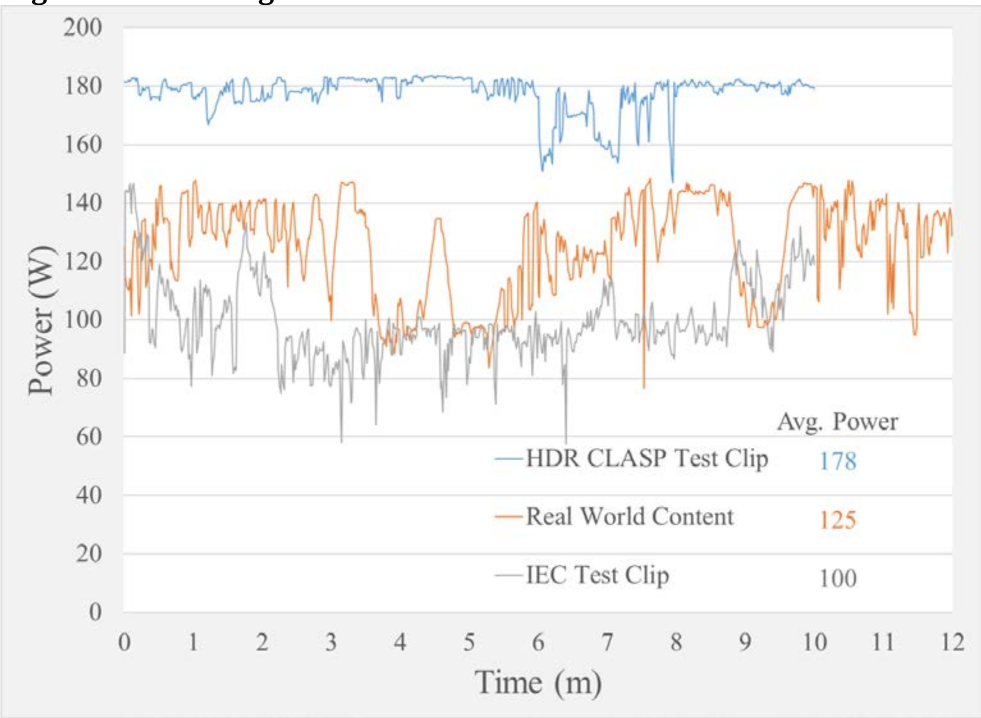
Notes: TV Settings Parameters - Default Picture Setting with ABC Off

Figure 11. Samsung UN55JS9000 Power Measurements



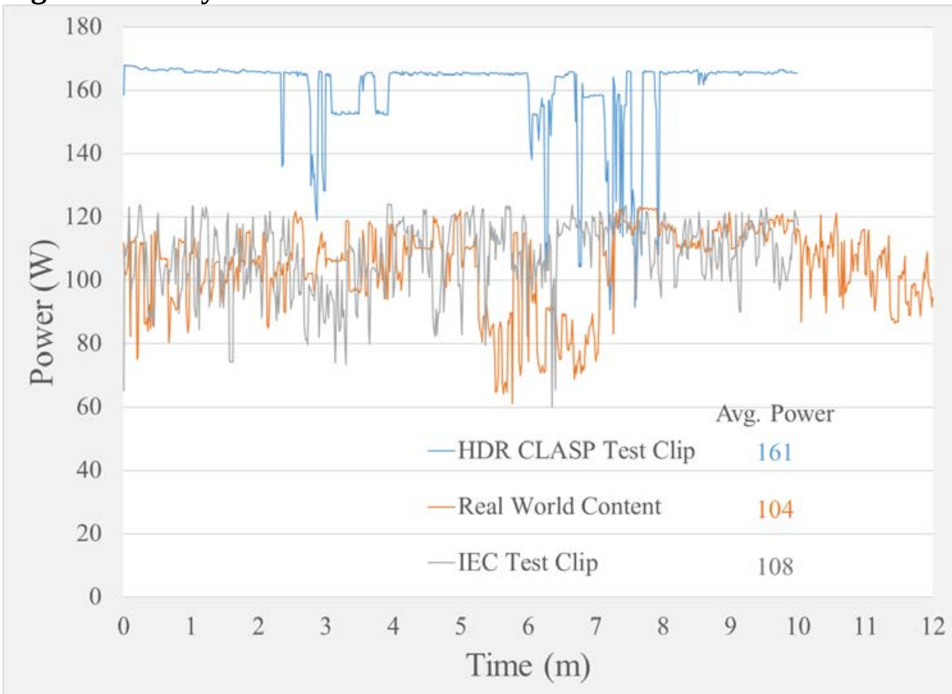
Notes: TV Settings Parameters - Default Picture Setting with ABC Off

Figure 12. Samsung UN55JU7100 Power Measurements



Notes: TV Settings Parameters - Default Picture Setting with ABC Off

Figure 13. Sony XBR-55X850C Power Measurements



Notes: TV Settings Parameters - Default Picture Setting with ABC Off.

11. Appendix E: Technical Background

This Appendix is targeted at technical participants and standards setting and regulatory discussions related to TV power measurement, not to a general audience. The contents of this Appendix have been developed with input from some industry experts; however, they have not been peer reviewed. This section of the report provides technical background on HDR beyond the overview provided in section 0. It also provides in depth analysis of alternative picture level metrics.

11.1 HDR Technical Background

Section 0 provides an overview of HDR technology. This section provides additional technical detail to help the technical audience understand the in depth analysis that forms the basis for this reports findings.

11.1.1 Video Pipeline

Some of the concepts described in this Appendix require an understanding of key video pipeline terms and process flows. A simplified list of video capture, delivery and display steps is included below to convey these concepts.

1. Video images are captured by a video camera that senses red, green and blue (RGB) light levels, and converts the analog sensor output to digital RGB signals, typically 16-bit for modern cameras (Sony 2016).

2. The video is processed through a gamma curve—or electro optical transfer function

(EOTF) in the case of HDR—to generate gamma-corrected signals R'G'B' where the apostrophe denotes "gamma-corrected." In today's digital world, these transfer functions are designed to enabled the optimal degree of signal compression and picture fidelity. Section 11.1.3 provides additional background and examples on gamma curves and EOTFs.

3. The R'G'B' is converted to Y'UV (= Y'CbCr) signals to enable more compression, where Y' represents luma or brightness and UV (sometimes expressed as CbCr) represent color space.

   APL' equals Y' averaged across all the pixels in a frame or video clip. It is not appropriate to compare the APL' of an SDR clip to that of an HDR clip because Y' is calculated using different constants per **Equation 1**.

Equation 1. R'G'B' to Y' Conversion Formula

$$Y' = R' \times K_R + G' \times (1 - K_R - K_G) + B' \times K_B$$

Where:

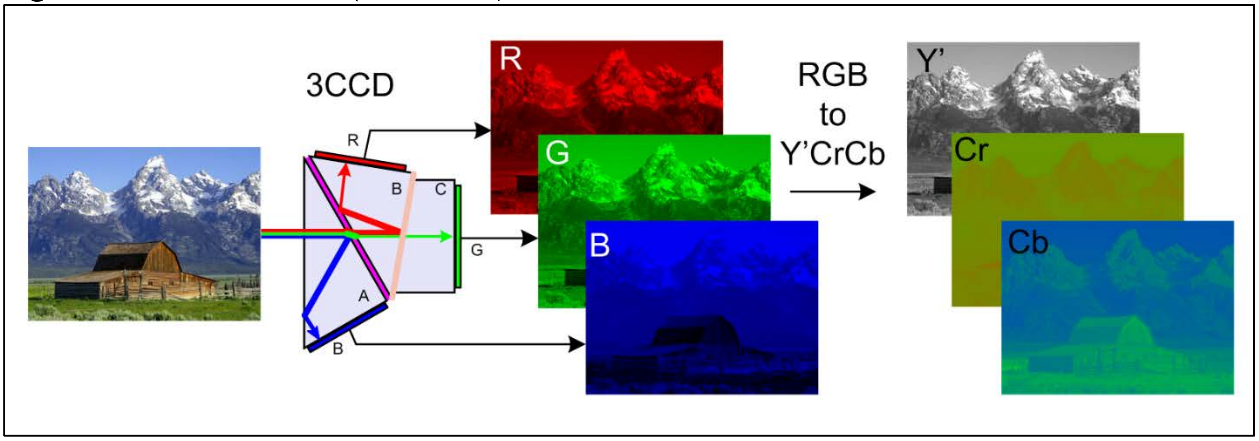
- *Y' and R', G', and B' are defined in section 11.1.1.*
- *For SDR (BT.709 color space)*
  - $K_R = 0.0722$
  - $K_B = 0.2126$
- *For HDR (BT.2020 color space)*
  - $K_R = 0.0593$
  - $K_B = 0.2627$

(Discovery Scientific 2015; ITU 2015)

4. Chroma subsampling can be used to compresses the video stream. The eye is more sensitive to Y' detail than color space detail, so Y' is allocated more bits that U or V signals in typical chroma subsampling approaches.

5. **Figure *14*** illustrates the decomposition of a color image into RGB components and to YUV components.

Figure 14. RGB to Y'UV (= Y'CrCb) Conversion



Notes: Public domain images retrieved from (LionDoc 2012). 3CCD refers to a camera with three image sensors or charged couple devices (CCD). Note that although it is not shown in this simplified picture, the transition from RGB to Y'UV involves the conversion to R'G'B' using a gamma or EOTF curve then to Y'UV.

6. The Y'UV or Y'CbCr signal is compressed for transmission.

7. The TV receives the compressed video stream and performs chroma resampling.

8. The TV converts the Y'UV or Y'CbCr signal back to R'G'B'.

9. The TV applies a gamma curve or EOTF to convert the signal back to RGB signals.

10. The RGB signals are converted to analog signals and sent to display panel components for rendering.

11.1.2 Color Grading

Before video compressions and transmission, the RGB video is color graded; a colorist adjusts the image captured by the camera to achieve a specific mood or other story-telling objective (Gates 2013). Colorists tend to color separately for SDR and HDR video. There are automatic algorithms that can create one format from the other, but the results are less than optimal. Some TV's, the Sony 850c for example, offer a preset picture setting that automatically up-scales SDR content to HDR. Based on views expressed by colorist Timothy Vincent (Limoor, Vincent 2016), up-scaling may prove to be little more than a marketing bullet point while studios and service providers work to make native HDR content more broadly available.

The role of the colorist is illuminated through examples. Given the limited color and contrast range of SDR video, a colorist may let a brightly lit part of an SDR clip wash out because they would have to make the rest of the clip too dark in order to capture detail in the bright part. The enhanced contrast supported by HDR reveals the detail of this bright part of the clip. However, the added detail reveals information that is inconsistent with the story (e.g. a jet airplane high in the sky for a movie set in 1850) or the added detail inadvertently draws attention away from the intended focus of the scene, the main character's face. These factors may cause the colorist to recolor the whole image when creating an HDR version to meet the director's intent for the scene resulting in different color grading between SDR and HDR versions.

According to Timothy Vincent, colorists prefer to develop the SDR version first since the SDR version tends to pale in comparison to the HDR version. If the colorist develops the HDR version first, then she or he tends to exaggerate the contrast in the SDR version in a futile attempt to achieve the striking effects achieved with HDR video. The bottom line is that the average brightness or light output of real world HDR video is different from that for SDR video, and average levels will continue to evolve as best practices for color grading HDR content evolve.

Figure 15 provides an example of how color grading can be used to convey meaning or set mood in a story. (Limoor, Vincent 2016; Digitalmediasuite 2012; Volpatto 2016)

Figure 15. Color Grading Example

Notes: Color grading can create set the mood in a scene or convey other meaning. HDR opens new storytelling possibilities for colorists. (Wesson 2014)

11.1.3 Transfer Functions

The ST2084 specification requires that all HDR 10 TVs map signal level to nits per the EOTF shown in **Figure 16**. Scaling—or allowing manufacturers to deviate from the prescribed mapping of signal level to nits—is not technically permitted for HDR TVs, but it is permitted and common practice for SDR TVs. An SDR TV supporting a peak luminance of 500 nits typically maps an 8-bit signal level of 235 to 500 nits in its brightest preset picture setting and scales proportionally down to its lowest light output level. An SDR TV supporting deeper blacks than those in the SDR specification scales the TVs response to the low end of the SDR signal range to utilize those deep black. So, it is not true that the HDR specification allows darker blacks than SDR TVs can offer, as sometimes claimed.

A word of caution about reading **Figure 16**: the ST2084 EOTF is applied to each of HDR 10's R', G', and B' color signals entering the TV. **Figure 16** shows the output of this transformation—R, G, and B light output signals—as "nits" (i.e. the y-axis label), but neither the eye nor a calibrated light meter would not see the resulting levels of nits for each color individually. This is confusing, but ***Table 12*** helps illustrate the point. A compliant HDR 10 TV with R'G'B' signal input of (510,510,510) would have a measured luminance of 100 nits, flat white. Each individual color would measure something less than 100 nits with a calibrated light meter, even though application of the EOTF results in a "false nits" value of 100 for each color. So caution must be used when interpreting **Figure 16**. (Zink 2015)

Table 12. Example Calculation of Average Light Output Power

	Red	**Green**	**Blue**	**Combined RGB**
Electrical input: 10-bit signal level, R'G'B'	510	510	510	--
Optical output: RGB ("false nits")	100	100	100	--
AvgRGBLOP (nits)	--	--	--	100

Notes: It may be helpful to refer back to **Figure 16** to understand, for example, why a signal level of 510 corresponds to light output of 100 nits.

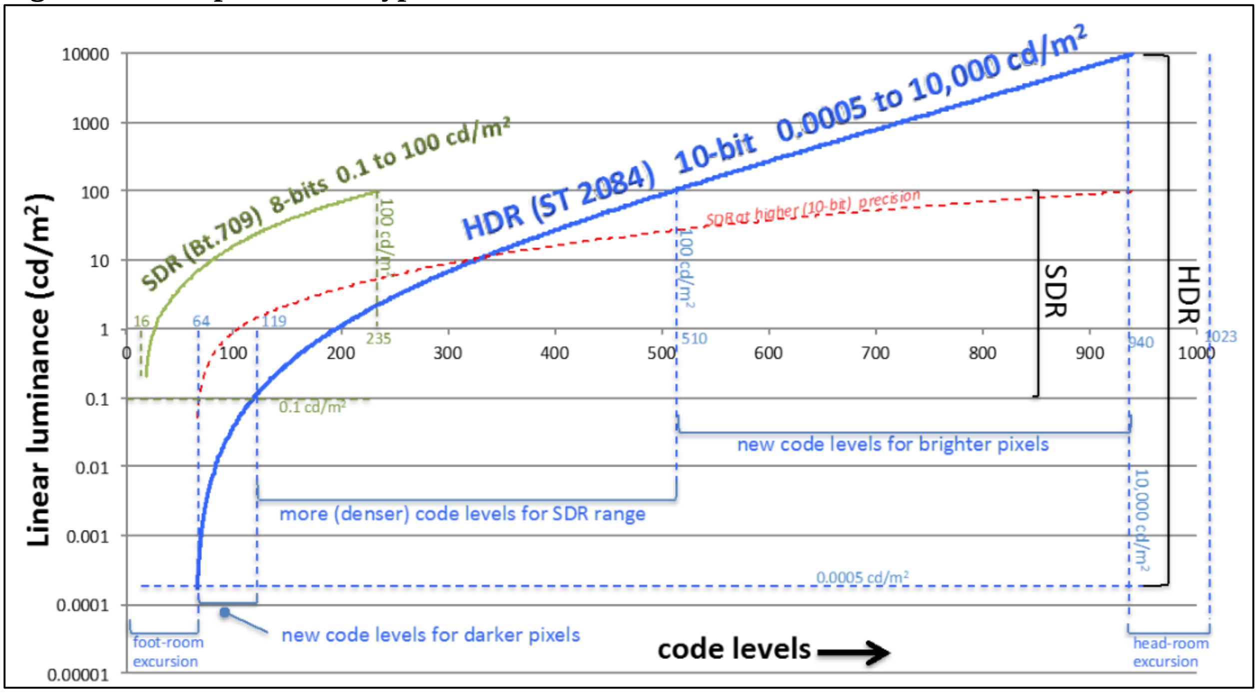
A few reminders about **Figure 16**:

- Both BT.709 (SDR) and ST2084 (HDR) reserve a few code levels at either end of the range that are not normally used. These are called head-room and foot-room.

- The reference range for BT.709 (SDR) is 0.1 – 100 cd/m^2 (nits). In other words, most SDR content is color graded on a 100 nit reference monitor in a dark room with the intent of ensuring that it looks great on a 100 nit display in a dark room setting. But BT.709 allows for scaling with the TV's capability to both brighter and darker levels than covered by the spec. An SDR TV capable of 400 nits peak brightness and 0.05 nits on the low end might modify the BT.709 transfer functions below—when set in a bright mode like Vivid or when located in a bright room with ABC enabled—so that the curve extends proportionally from 0.05 – 400 nits. In a dim room with Cinema mode selected, such a TV may operate in the reference range of 0.05 – 100 nits.

- The ST2084 spec does not allow for scaling with HDR video, but manufacturers appear to bend the rules to enable HDR TVs to be manually changed to preset picture settings that have different brightness level.

Figure 16. Comparison of Typical HDTV and HDR Transfer Functions

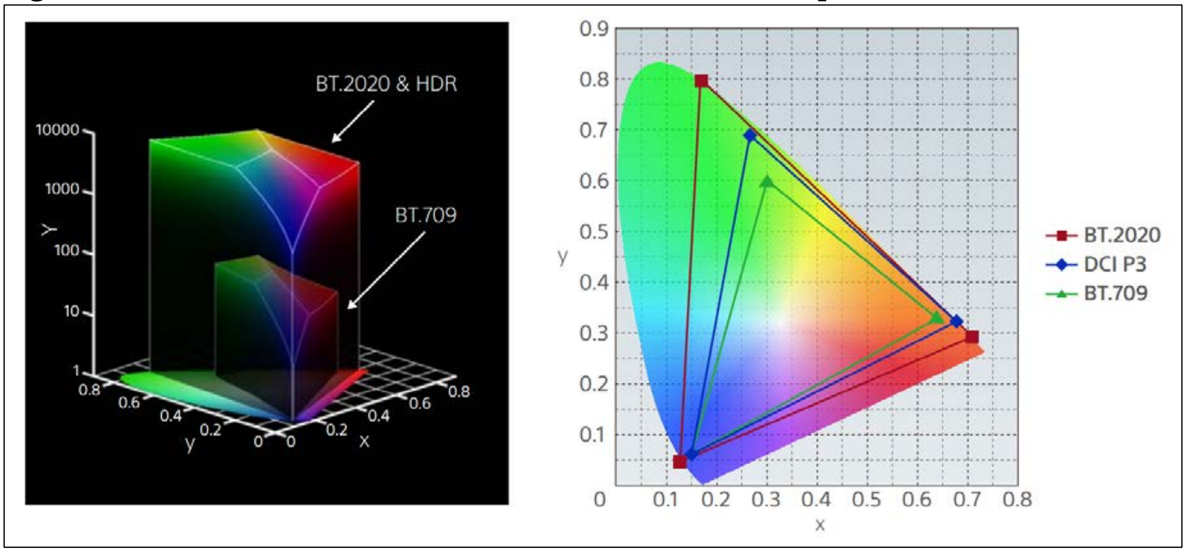


Notes: ST2084 defines an absolute mapping of signal level to optical output for HDR content. SDR video does not use an exact mapping; the SDR TV can scale the maximum input signal level to its peak brightness capability. The case above shows the reference case of an SDR display with 100 nit peak brightness. Retrieved from (Cineramax 2015).

11.1.4 Color Mapping

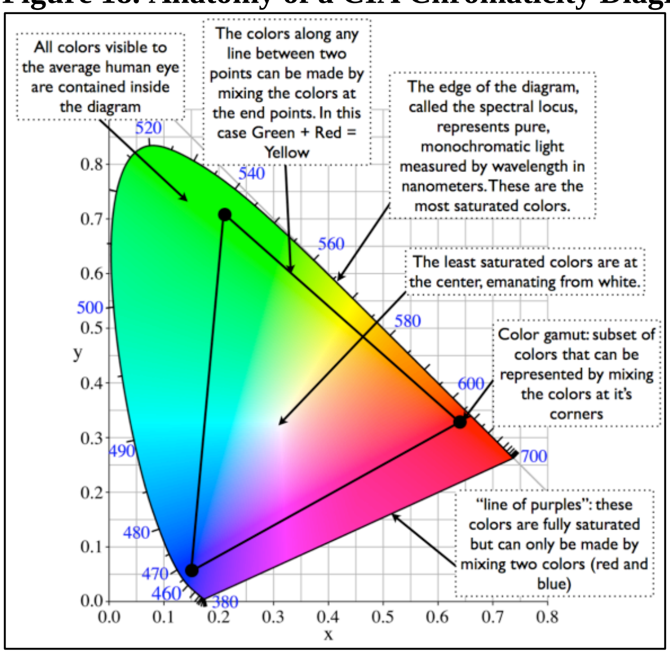
The HDR 10 standard requires that content is mastered in BT.2020 color representation. HDR 10 TVs must support BT.2020 signal input even though they may be physically limited to displaying 90% of DCI P3 colors. A comparison of HD TV (BT.709) versus HDR color space is shown in **Figure 17.** For guidance on how to read a chromaticity diagram, see **Figure 18**.

Figure 17. Visual Illustrations of Color Volume and Color Space



Note: The addition of luminance turns a color space map (right) into a color volume map (left). The vertical axis in the left diagram represents peak luminance in nits. Retrieved from (Sony 2016).

Figure 18. Anatomy of a CIA Chromaticity Diagram



Notes: Retrieved from (Yurek 2012).

11.1.5 HDR Compatibility

There are two versions of today's TV test clip, SD and HD. There may need to be additional versions of the next text clip depending on how interoperable the various HDR standards are with new and legacy TVs and whether or not compatibility promises are realized. Determining the optimal number of test clip versions is complicated because different standards are used for different types of content (e.g. live video vs. off-line content). It appears from the information in

Table 13 that IEC stakeholders may be able to develop an HDR test clip in Dolby Vision format that can be written onto a Blu-ray disk and is compatible with both HDR 10 and SDR TVs. **Table 13** lists the four major HDR standards in existence today and their respective compatibility with each other, SDR TVs, Blu-ray players and set-top boxes (STB).

Table 13. Compatibility of Existing HDR Standards

<table>
<tr><th rowspan="2">HDR Systems</th><th rowspan="2">Use</th><th rowspan="2">EOTF</th><th rowspan="2">Metadata</th><th rowspan="2">Min. HDMI</th><th colspan="6">Compatibility</th></tr>
<tr><th>SDR TV</th><th>HDR 10 TV</th><th>Dolby Vision TV</th><th>Blu-ray</th><th>HLG TV</th><th>STB</th></tr>
<tr><td>HDR 10</td><td>Specified by CTA and UHD Alliance for HDR-compatible TVs and mandated for Ultra HD Blu-ray</td><td>ST2084</td><td>A single content layer with ST2086 static metadata</td><td>2.0a</td><td>No</td><td>Yes</td><td>Optional</td><td>Yes</td><td>No</td><td>?</td></tr>
<tr><td>Technicolor/ Phillips</td><td>Used primarily by pay-TV service providers like Comcast and OTT providers like Netflix</td><td>ST2084</td><td>A single content layer with metadata</td><td>?</td><td>Yes</td><td colspan="3">Optional via STB</td><td>?</td><td>Yes</td></tr>
<tr><td rowspan="2">Dolby Vision</td><td>Off-line production as opposed to live broadcast</td><td rowspan="2">ST2084</td><td>A base content layer with static metadata</td><td rowspan="2">1.4</td><td colspan="2">Optional</td><td>Yes</td><td>Optional</td><td>No</td><td>?</td></tr>
<tr><td>Real-time HDR content such as live broadcast or OTT applications</td><td>An enhanced content layer with dynamic metadata</td><td>No</td><td>No</td><td>Yes</td><td>No</td><td>No</td><td>?</td></tr>
<tr><td>BBC/NHK</td><td>Real-time broadcast in UK and Japan</td><td>HLG*</td><td>No metadata</td><td>1.4</td><td>Yes</td><td>No</td><td>No</td><td>No</td><td>Yes</td><td>?</td></tr>
</table>

Notes: *HLG = hybrid log-gamma. Based on (Schulte, Barsotti 2016).

Ecos Research has framed the process by which the number of test clips could be determined in **Table 14**. More analysis is needed to develop a recommended approach. For example, this research project revealed compatibility issues when trying to play BT.2020 HDR 10 content on some 2015 HDR displays, raising the possibility that even more versions are needed than reflected below.

Table 14. Preliminary Framework for Determining Number of Needed TV Test Clips

	Format	Resolution	Color gamut	Color depth	Target Display Max Brightness
SD clip (if needed)	480i	640 × 480	SMPTE C	8-bit	Not specified, scales to TV capability
ST2084-based clips	1080p	1,920 x 1,080	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	10,000 nits per ST2084
	4K	3,840 x 2,160	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	10,000 nits per ST2084
	8K	7,680 x 4,320	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	10,000 nits per ST2084
HLG-based clips	1080p	1,920 x 1,080	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	Not specified, scales to TV capability
	4K	3,840 x 2,160	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	Not specified, scales to TV capability
	8K	7,680 x 4,320	BT.2020 with backward compatibility	Supports 12-bit Dolby Vision, 10-bit (HDR 10), and SDR backward compatibility	Not specified, scales to TV capability

Notes: Based on (Schulte, Barsotti 2016).

11.2 Evaluation of Picture Level Metrics

This section of the report presents both theoretical and empirical approach to evaluating picture level and light output metrics for use in determine to what extend a proposed test clip is representative of real world content.

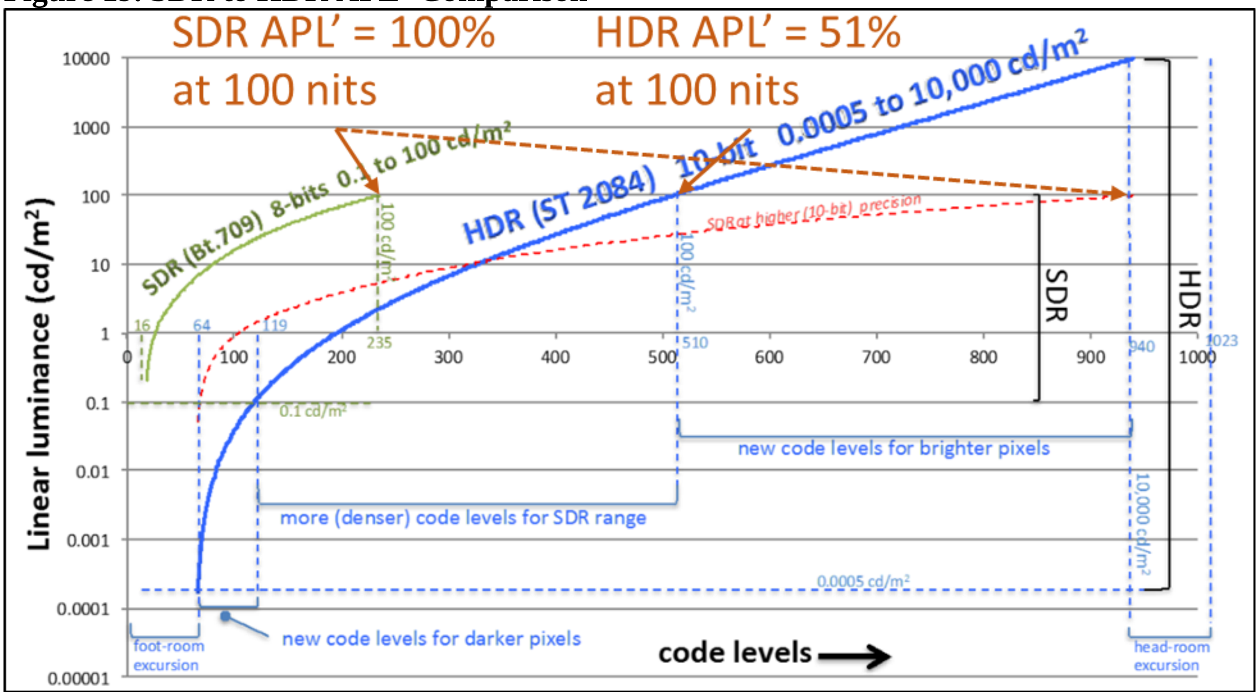
11.2.1 Theoretical Analysis of Approaches

From a purely theoretical standpoint, light output power appears to be the best approach compared to average picture level or average light level; although, the correlation study results presented in section 11.2.2 of this report cast doubt on this theory.

As described in IEC 62087 (International Electrotechnical Commission 2015), APL' is calculated by analyzing the Y'UV video signal before decoding the signal in the TV. IEC 62087 specifically refers to APL' as Gamma Corrected Average Picture Level (APL'), which is calculated by averaging each pixel's average luma (Y') value across a frame or clip. Because SDR and HDR video standards use different parameters to calculate Y' (**Equation 1**) and different transfer functions (**Figure 16**), APL' calculations for SDR and HDR are not comparable.

It is helpful to use an example to illustrate how APL' values differ for an HDR frame compared to an SDR one. This example assumes the use of SDR reference display peak luminance of 100 nits (see **Figure 19**). For reference white—100 nits in this case—an SDR video frame would have an APL' of 100 percent because reference white represents 100 percent of the maximum code level, factoring out code level headroom. For 8-bit SDR, reference white occurs at code level 235 out of a possible 235. However, in the HDR case, 100 nits occurs at code level 512 out of 940. Eliminating head and foot-room from the calculation, the resulting APL' is 100 x (512 – 64) / (940 -64) = 51.1 percent. So at the upper end of the SDR luminance range, SDR APL' differs from HDR APL' by almost a factor of 2.

Figure 19. SDR to HDR APL' Comparison



Notes: Retrieved from (Cineramax 2015).

The CLASP HDR test clip has an APL' value close to that of the IEC SDR test clip, 36 percent and 34 percent respectively. For a 10-bit HDR clip 36 percent correlates to a code level of 0.36 x (940 - 64) + 64 = 379, which is near the intersection of the 10-bit BT.709 and ST2084 curves (**Figure 16**). One could argue that as long as the APL' of real-world content remains in the range of this intersection, then APL' is a suitable metric. One could also argue that it is necessary to match test clip APL' to real world content APL' separately for SDR and HDR video, so it doesn't matter if APL' values are not comparable across SDR and HDR clips.

In addition, because APL' is calculated based on the external video signal, which does not scale linearly with the light output signals sent to the display components within the TV, APL' does not scale linearly to power, even in theory. It scales based on the BT.709 gamma curve or the ST2084 EOTF. It is not critical that the metric used to match a test clip to representative content scale linearly, but it would be more useful if it did.

From a theoretical perspective, it appears that a metric called light output power (LOP) would be better suited for the intended purpose. This metric represents the average of RGB (not R'G'B') signals across a frame or video clip (section 11.1.1). The RGB signals result from processing the external signal, used to calculate APL', through the appropriate transfer function. The average of the RGB signals represents the theoretical light output of the display, in nits. As calculated, light output power excludes the effects of the following variables, which represent features that are not standardized across brands or models:

- **Automatic Brightness Limiting (ABL)**: All displays have a maximum power limit for the display as a whole. No display can support peak brightness (1,000 nits for example) on every pixel across the display. When a display approaches the maximum power limit, it scales pixel brightness—typically in a proprietary way—to avoid power overload.

- **Luminance and color roll-off**: When HDR displays play content that includes peak brightness values or color gamut width in excess of the TV's capability, then rather than just remap all of the out-of-range values to the TV's peak value—a process called clipping—an image processor remaps the out-of-range values across the upper range of TV-supported values to provide a more gradual transition.

- **Automatic Brightness Control (ABC)**: ABC adjusts a TV's brightness according to ambient light levels.

- **Aspect ratio scaling**: TVs scale images to suit the preference of the viewing. For example, this may involve displaying a movie in letterbox format or stretching it so that the picture covers the entire screen, modes that require different display power levels.

The exclusion of the above variables provides a cleaner analysis of the video image itself without factoring what a TV may or may not do to that signal, but eliminating these variables does reduce the correlation between AvgRGBLOP and measured power consumption of lab TVs. This approach enables the international community to develop a test clip that is representative of real world content, not to develop a metric that will predict the light output or power consumption of individual display models.

The simple average of RGB values across an HDR frame or clip, expressed as AvgRGBLOP (**Equation 2**), is the expected output of an HDR 10 or Dolby Vision display for that frame or clip.

Equation 2. HDR AvgRGBLOP

HDR AvgRGBLOP = (AvgR + AvgB + AvgG)/3 in nits

Where:

- *VQPLA software uses the ST2084 transfer function to* *calculate R, G, and B from R', G', and B'.*

The equivalent SDR equation (**Equation 3**) uses R'G'B' values as input and contains exponents that approximate the inverse gamma function.

Equation 3. SDR AvgRGBLOP

$$SDR\ AvgRGBLOP = (R'^2 + G'^2 + B'^2)/3 \text{ in nits}$$

A third metric that Ecos and VideoQ have consider is an industry standard metric known as frame average light level (FALL), which is calculated as Max(RGB) averaged across a frame or clip. In this case, order of operations matters; the maximum of R, G, and B values is calculated for each pixel, and those maximum values are averaged across a frame or clip. MaxFALL represents the maximum frame AvgFALL for entire clip. FALL is commonly used to assess the relative brightness of various video clips—since the eye perceives monochromatic images to be generally as bright as white images. The International Broadcasting Corporation uses FALL for this purpose in a recent white paper (Zink 2015). Given that the focus of FALL is perceived brightness and not light output, AvgRGBLOP should in theory be the better metric for the purpose of test clip development.

11.2.2 Empirical Analysis of Approaches

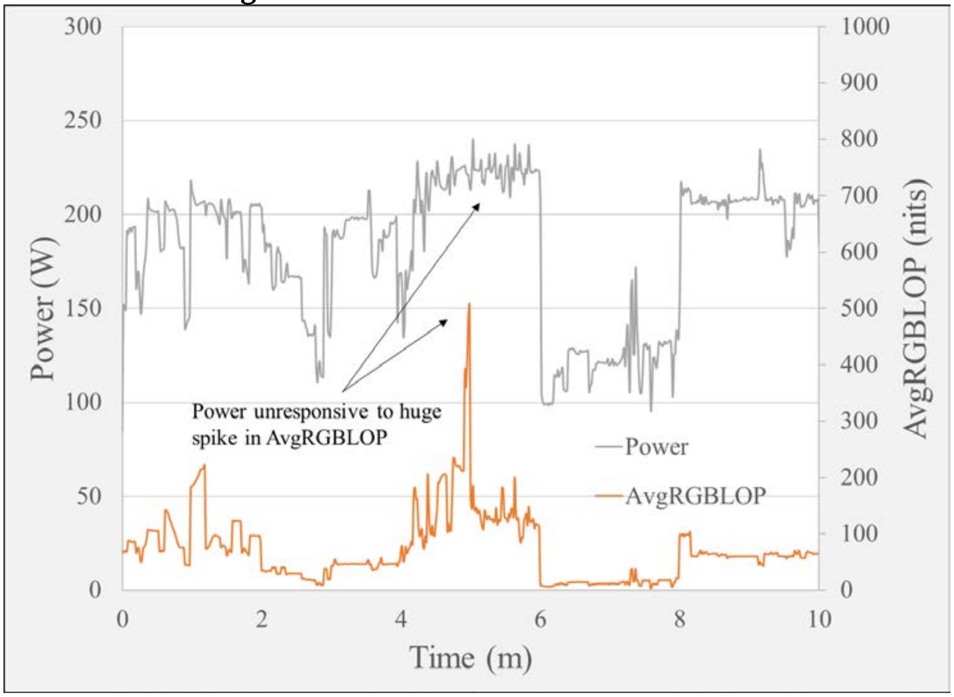
Through empirical tests, Ecos sought to verify that AvgRGBLOP was the best metric by verifying that it most closely correlated to TV power consumption. It did not; it performed worst of the three metrics evaluated reinforcing this reports recommendation to work closely with industry technical experts when developing metrics to update the test clip. This section summarizes correlation study results and discussion of limitations.

Correlation to TV Power Use

The biggest challenge associated with empirical correlation analysis is that the factors removed from our theoretical analysis—ABL, roll-off, ABC, aspect ratio scaling, and BT.709 luminance scaling—have the potential to reduce the level of correlation across the board. In order to reduce the number of variables, Ecos conducted power correlation tests with ABC and MDD disabled. Furthermore, the three clips that Ecos tested for this purpose were formatted in 16:9 aspect ratio, so they were displayed full-screen without scaling in all tests.

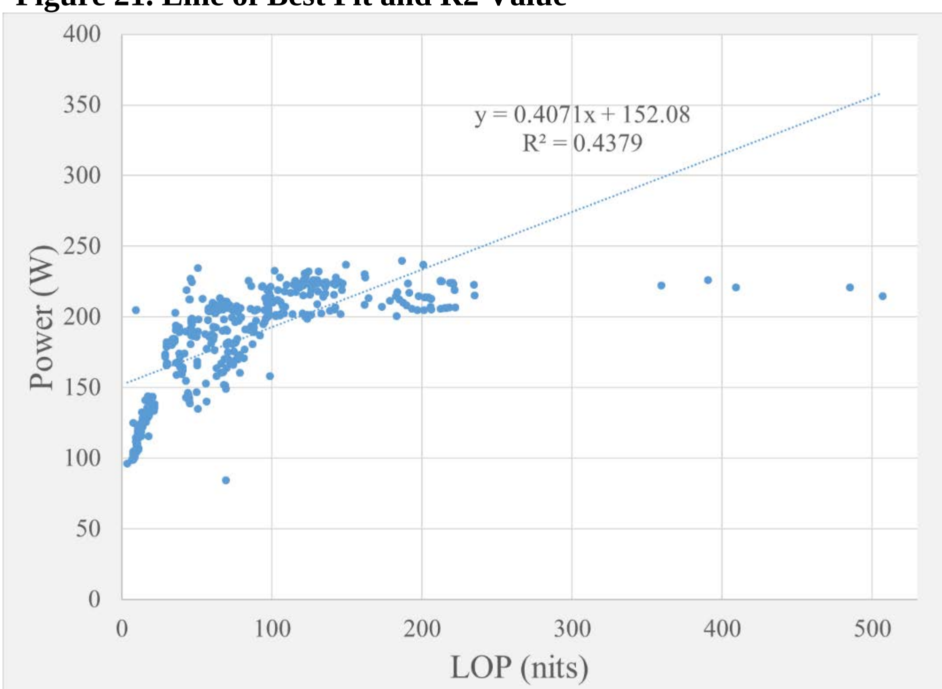
Figure 20 and **Figure 21** are provided below as an example of the process used to calculate correlation values. They represent data collected to calculate the correlation between LG9600 power and AvgRGBLOP by a) measuring power versus time for the CLASP HDR test clip, and b) plotting scatter diagram in order to calculate line of best fit and associated R^2 value.

Figure 20. Power Vs. AvgRGBLOP



Notes: TV Model – LG 55EG9600 (OLED). TV Settng Parameters –Default Picture Setting with ABC Off. AvgRGPLOP (Target Display Maximum Brightness (TDMB) set to 10,000 nits) vs. Power for CLASP HDR, 8-bit. ABL can be seen at about 5 minutes into the clip. LOP spikes but power is limited to below 250 watts.

Figure 21. Line of Best Fit and R2 Value



Notes: TV Model – LG 55EG9600 (OLED). TV Settng Parameters –Default Picture Setting with ABC Off. AvgRGPLOP (TDMB 10,000) scatter plot with R^2 value for

CLASP HDR, 8-bit. LG OLED TV. This plot also reflects a power limit of less than 250 watts.

Table 15 shows the results of the R^2 analysis. Contrary to our theoretical analysis, APL' is most correlated to power use in the tests conducted followed by AvgFALL. Perhaps this has to do with the fact that APL' factors the human eye's sensitivity to different colors via the constants expressed in **Equation 1**.

Table 15. Line of Best Fit R^2 Values

	Content	APL'	AvgFALL_nit	AvgRGBLOP_nit
LG 9600 (OLED)	HDR Clasp Test Clip	0.75	0.47	0.44
	SDR Real World Content	0.45	0.54	0.42
	SDR IEC Test Clip	0.5	0.45	0.34
SAM 9000 (Edge Lit)	HDR Clasp Test Clip	0.06	0.08	0.08
	SDR Real World Content	0.27	0.29	0.25
	SDR IEC Test Clip	0.25	0.21	0.19
SAM 7100 (Full Array)	HDR Clasp Test Clip	0.34	0.17	0.15
	SDR Real World Content	0.2	0.25	0.21
	SDR IEC Test Clip	0.18	0.15	0.13
Sony 850C (Edge Lit)	HDR Clasp Test Clip	0.15	0.12	0.12
	SDR Real World Content	0.2	0.33	0.29
	SDR IEC Test Clip	0.29	0.21	0.25
Average		0.30	0.27	0.24

Notes: TV Setting Parameters - Default Picture Setting with ABC and MDD off. Gold shading denotes the highest level in a row in cases where the difference between highest and next highest is greater than 0.02.

Table 16 reflects the average R^2 values for the three video clips tested. These results are shared for information purposes. No conclusions were apparent from these results.

Table 16. Average R2 Values for Video Clips Tested

Test Clip	Average R^2 Values
HDR Clasp Test Clip	0.24
SDR Real World Content	0.31
SDR IEC Test Clip	0.26

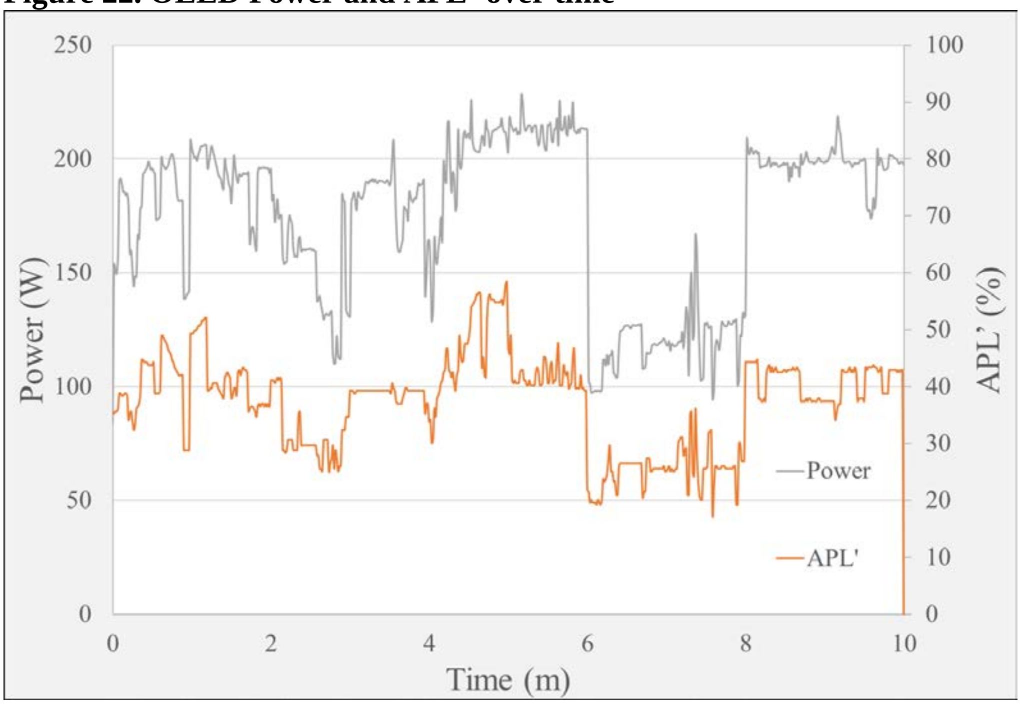
Table 17 shows that, as expected, OLED technology demonstrated the highest correlation with the three metrics evaluated in **Table 15**.

Table 17. Average R^2 Values by TV Technology Type

TV Technology	Average R^2 Values
OLED	0.48
Edge Lit LED	0.20
Full Array LED	0.20

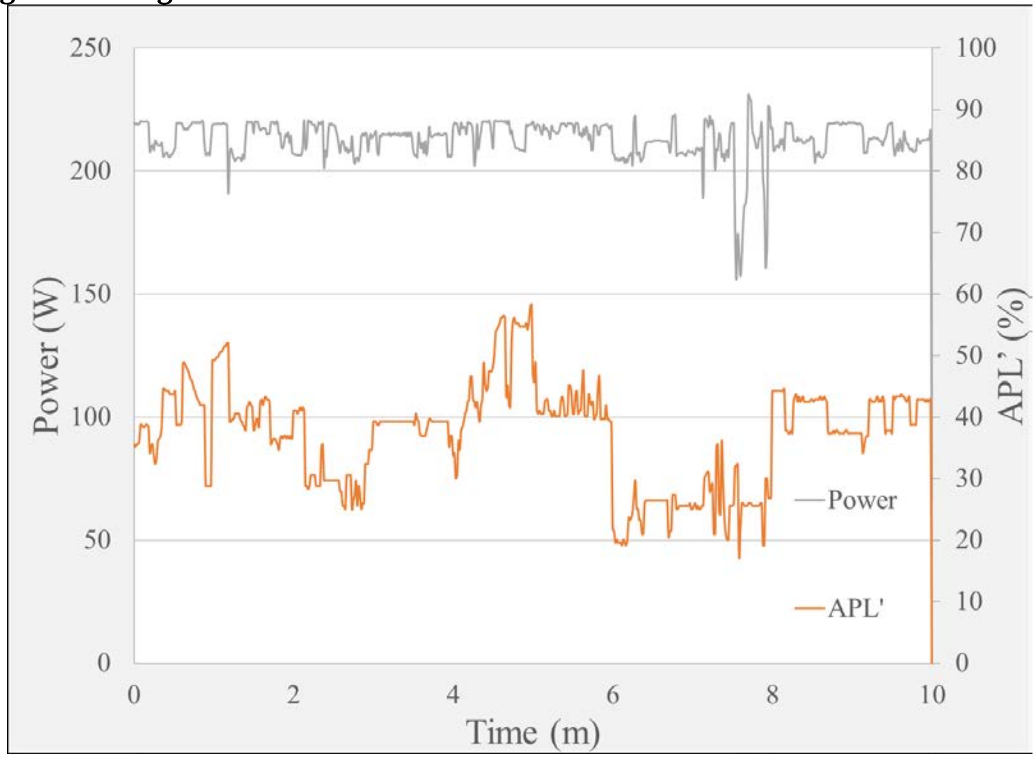
Figure 22 and **Figure 23** illustrate why OLED technology tends to have a higher correlation than Edge Lit LED. Because OLED technology is emissive, subpixel power scales to subpixel energy consumption, and power scales to APL'. Edge Lit LEDs scale power by dimmable zone, but not by subpixel resulting in a lower correlation. The difference is most clearly illustrated by minutes six to eight in the figures below.

Figure 22. OLED Power and APL' over time



Notes: TV Model – LG 55EG9600 (OLED). TV Setting Parameters –Default Picture Setting with ABC Off. Content – CLASP HDR Test Clip

Figure 23. Edge Lit Power and APL' over time



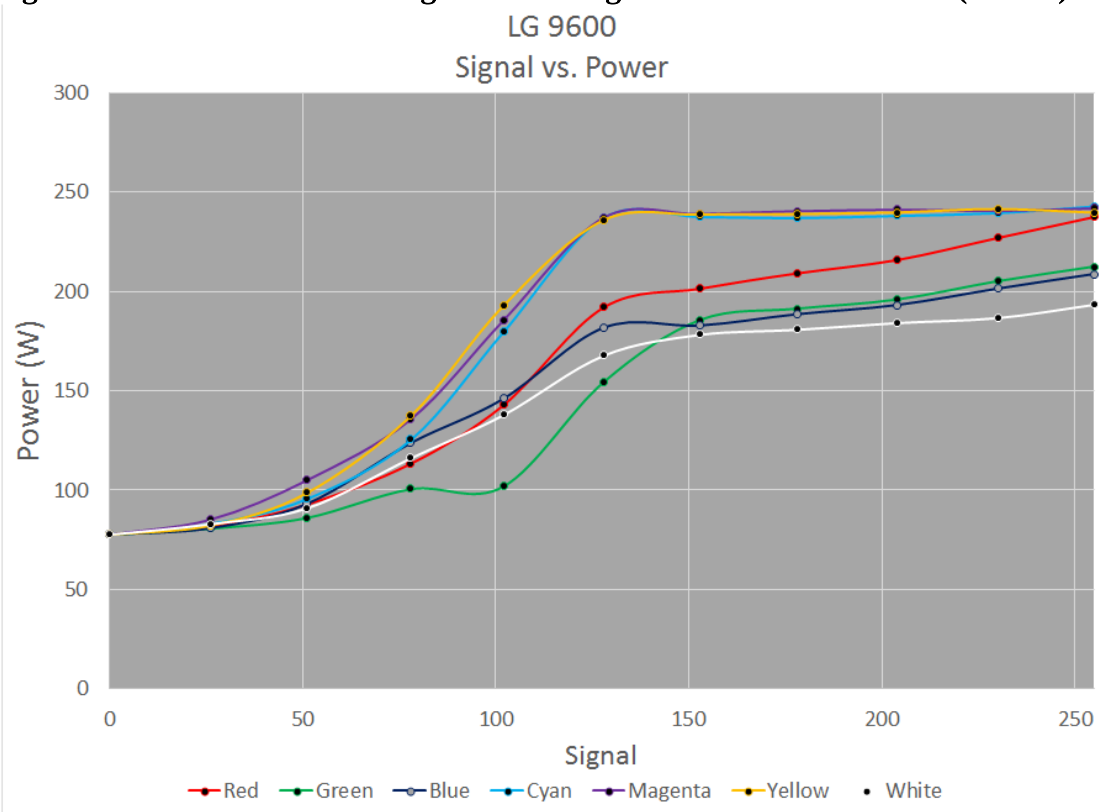
Notes: TV Model – Samsung UN55JS9000. TV Setting Parameters – Default Picture Setting with ABC Off. Content – CLASP HDR Test Clip

LG9600 (OLED) Observations

As discussed above, there are many variables that could affect the correlation between clip AvgRGBLOP and TV power use. For example, a deeper investigation of the LG9600 (OLED) display added another variable to the mix. This TV uses an OLED configuration called RGBW in which each subpixel has an independent white light source. The RGB subpixels use light blocking color filters to achieve their desired color. The fourth, white subpixel does not need a color filter, so it operates more efficiently. It uses less energy for this display to generate white light using the white subpixel than to generate through equal parts R, G, and B. What this means is that this TV operates more efficiently for relatively white frames than for monochromatic ones.

Ecos' Calman luminance meter software includes test patterns that enable one to deliver specific R'B'G' signals to a TV via HDMI cable. **Figure 24** shows the results of LG9600 tests with 10 signal levels each for Red, Green, Blue, Cyan, Magenta, Yellow, and White. Cyan Magenta and yellow are relevant because they all use only 2 subpixels, and therefore cannot use the more efficient white subpixel. Full Cyan for example has an R'G'B' signal level of (0,255,255); Magenta is (255,0,255); and Yellow is (255,255,0). Note that 255 represents the highest value for 8-bit video signals including head-room per the discussion of **Figure 16**.

Figure 24. R'G'B' 8-bit SDR Signal Levels Against Power for LG9600 (OLED)



Notes: TV Setting Parameters – Default Picture Setting with ABC Off; The green data point at signal level 100 may be anomalous.

The primary findings of and questions raised by this test are:

1. White frames generally use less power than frames saturated with one or two primary colors. This is likely due to the lack color filtering on the white subpixel.

2. The slope of all lines plotted in **Figure 24** decrease after signal level 128. This is likely the result of ABL for cyan, magenta and yellow given that the power level for these colors peaks at close to 250 watts, which appears to be the power limit in **Figure 20** and **Figure 21** as well. It's less obvious why elbows exist at this signal level for the other colors. There is no associated bend in the BT.709 gamma curve at this signal level per **Figure 16**.

3. It is not clear why the blue and green lines have negative slope at times or why the green and white lines cross. More technical expertise is needed to fully understand this data.

Once again, the fact that there are many variables that reduce the correlation between picture level or light output metrics does not mean that IEC stakeholders should not use these metrics as an aid in developing the next test clip. It does mean that engagement of industry technical experts is required to thoroughly vet any approach.

Histogram Analysis

IEC stakeholders used an APL’ histogram to ensure that the current IEC test clip had a representative distribution of APL’ levels compared to 40 hours of real world content. (International Electrotechnical Commission 2015) In this section of the report, we will examine the use of histograms to better understand the relative differences between similar SDR and HDR content.

Figure 25 shows several histograms of SDR and HDR video clips analyzed in this study with TDMB set to 100 for the SDR clips. It is clear from these histograms that, if viewed in a dark room with these assumptions, that the HDR clip would be much brighter than the SDR clips, which runs counter to the stated intent of the HDR specification. It is possible that colorists are actually color grading with a 300 – 500 nit peak luminance display as the target, which would be common for contemporary TVs used in mid-range to bright ambient light conditions. **Figure 26** shows that the SDR clips, when scaled to a 500 nit display, have AvgRGBLOP values that are similar to those for HDR content.

Figure 25. AvgRGBLOP Histograms Assuming a 100 nit SDR Reference Display

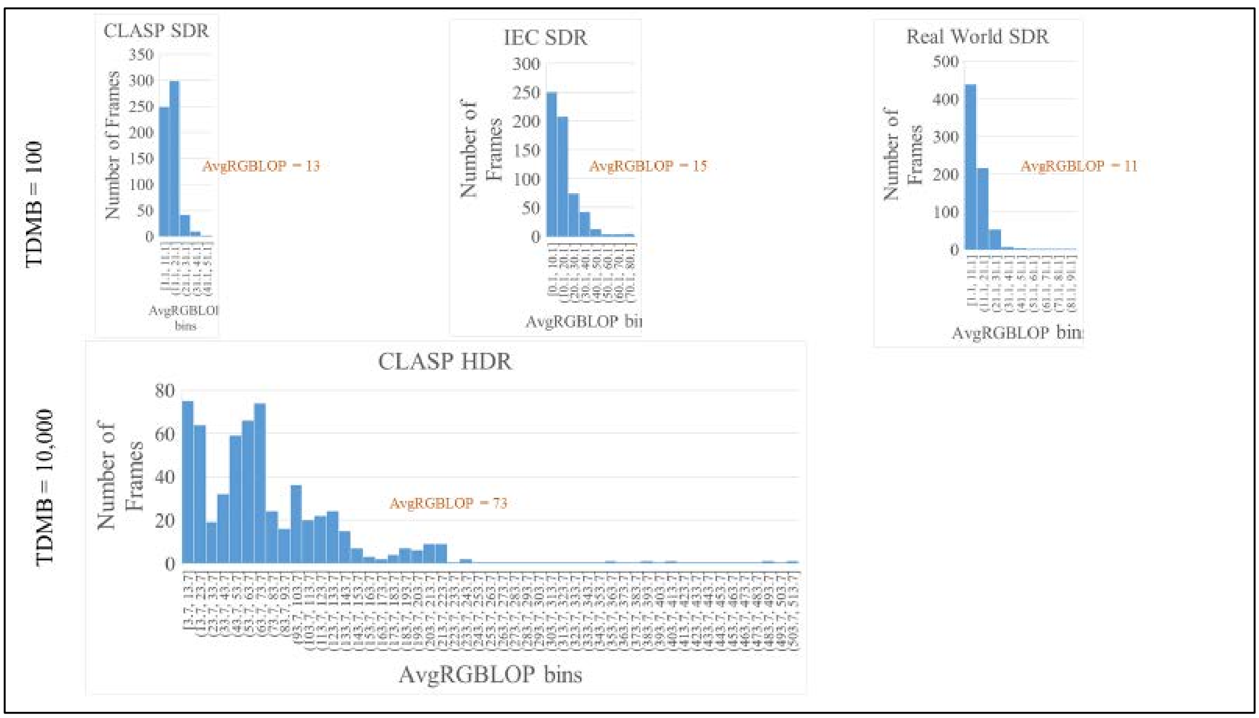


Figure 26. AvgRGBLOP Histograms Assuming a 500 nit SDR Reference Display