PQM: A New Quantitative Tool for Evaluating Display Design Options

Software, Electronics, and Mechanical Systems Laboratory
3M Optical Systems Division

Jennifer F. Schumacher, John Van Derlofske,
Brian Stankiewicz, Dave Lamb, Art Lathrop

3M

Contents

Executive summary

Today, manufacturers can create displays with levels of resolution, luminance, contrast and color that approach the limits of human perception. However, due to cost, sustainability and other constraints, product designers usually don't maximize all performance characteristics in every display. Instead, they make choices about which configuration of characteristics will achieve the best mix of market appeal and profitability.

3M has developed a new tool to let product developers forecast how these design decisions affect perceptions of quality. The Perceptual Quality Metric (PQM) calculates expected viewer perceptions of quality based on display size, resolution, luminance and color gamut. Using this tool, product developers can quantify the perceptual quality improvements in products based on changes in display specifications, without investing in large and expensive consumer preference studies.

These assessments can guide the inevitable tradeoffs made in display design or, more intriguingly, drive toward a display that achieves unprecedented levels of perceived quality.

Initial applications of PQM reinforce the importance of high color gamuts and the attractiveness of 4K resolution, especially with larger displays. The improvement to 8K resolution shows little effect on perceptions of quality. These findings suggest that manufacturers will need to pay renewed attention to color management software and content providers will want to consider the use of high-gamut colors in their imagery.

3M's new quantitative tool lets product developers quantify perceptual quality improvements based on changes in display specifications.

Challenges in display design

Product developers now have extraordinary opportunities for advancing display performance.

New tools create displays approaching the limits of human perception in resolution and practical value in luminance. With new architectures and technologies, such as quantum dots and OLED displays can also offer a much larger color gamut—a performance characteristic that has lagged behind gains in resolution and luminance.

Most manufacturers do not produce displays that provide the highest levels of performance on all characteristics, even though they have the technical capability to do so. Instead, they make design choices based on market considerations: consumer cost constraints, consumer and regulatory concerns about energy consumption, or preferences for color and resolution.

Manufacturers can calculate the impact of these tradeoffs with relative ease. For example, if a product developer was considering a reduction in light sources, the impact on costs and luminance is measurable using readily available tools. The improvement in energy consumption is also simple to determine.

Measuring the impact of display characteristics against consumer perceptions of quality is much more difficult. It is even more challenging to gauge the change in consumer attitudes if a reduction in one performance attribute—luminance, for example—is accompanied by an improvement in another, such as color or resolution.

In theory, large-scale surveys could provide developers with detailed information on consumer preferences In reality, this research is expensive and short-lived. Key performance characteristics improve rapidly and consumer expectations change with them. For example, desirable resolution in 2007 was considered subpar in 2010. Not surprisingly, there have been few large-scale customer preference studies.

Computational modeling of preferences

As an alternative to studies, some researchers pursue computational models of consumer preferences, such as P.G.J. Barten's Square Root Integral (SQRI) metric. First published in 1987, the SQRI calculates expected viewer preferences for size, resolution, and luminance. SQRI's value is limited, however, in that it does not consider color gamut and contrast.

Recently, authors and other researchers at 3M developed the Perceptual Quality Metric (PQM), a new computational model based in part on the SQRI. PQM calculates expected viewer perceptions of quality based on display size, resolution, luminance and color gamut. Using this tool, product developers can predict how changes in display performance affect consumer perceptions of quality.

These assessments can guide the inevitable tradeoffs made in display design, or drive toward a display that achieves unprecedented levels of perceived quality.

A pilot eye-tracking study by 3M suggests that content with a higher color gamut receives more attention, as measured by aggregated fixation or dwell time, than content with a lower color gamut.

Developing the PQM

3M conducted a series of experiments to determine consumer preference for images displayed with varying luminance and color gamut characteristics.

In the first experiment, 14 adult subjects (equally divided by gender) were shown a set of 10 images over multiple trials. The images were randomly generated shapes (triangles, squares, and circles) of various sizes and random colors as a context-independent condition where humans would have no expectation of color.

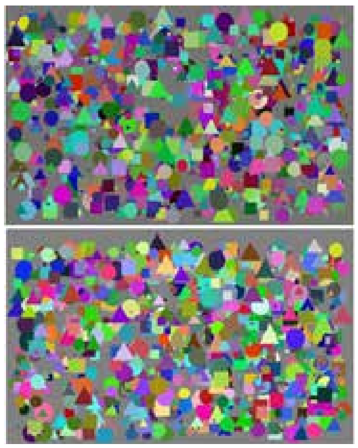


Figure 1. Random shapes and colors used as a context-independent condition such as a chart or graph, where humans would have no expectations of color ("memory colors").

In a second experiment, 24 adult subjects (equally divided by gender) viewed one image randomly selected from a set of five photographs. The subject viewed the same photo multiple times, with variances in color gamut and luminance. The photographs were used as a context-dependent condition to simulate display usage where subjects such as the sky have an expected color ("memory color"); photos included multi-color and single color (red, green, and blue) objects.

Figure 2. Photographs were used as a context-dependent condition to simulate display use where subjects have an expected color.

Each of 15 stimuli was processed to simulate every possible combination of four gamut sizes and three luminance levels, for a total of 12 simulations for each photo and image.

In both experiments, subjects were seated 36 inches from two high-performance monitors, arranged side by side. For each stimulus, two variations (with different luminance and color) were presented simultaneously on the two monitors (one image per monitor) for two seconds, after which white noise appeared until a preference was recorded. Every possible pairing was presented, including comparisons with identical characteristics. Additionally, to ensure one monitor was not preferred over the other, each pairing was repeated with the images on the opposite monitor. A total of 156 trials were given to each subject.

The data were used to obtain several computational models incorporating luminance and color gamut area. A validation experiment was conducted using 36 subjects (22 male). Five color gamut areas and four luminance levels were varied (creating 20 conditions) and subjects rated these display simulations for each of three images. The equation with the highest correlation to the validation data was selected as the final PQM.

Similar scores using different configurations

Initial results suggest that PQM is an accurate tool for forecasting how changes in color gamut, luminance and resolution will affect viewer perceptions of quality, as measured by Just Noticeable Differences (JNDs). One JND represents a difference between compared devices that is noticeable but does not have a large impact on preference; three JNDs represent a significant impact and 10 is considered substantial. The validation experiment showed a correlation of 0.97 between predicted quality values and actual values.

The implications of this tool are significant. By graphing luminance and color gamut, for example, developers can readily predict the relative impact of performance improvements or reductions without investments in large and expensive studies of consumer preferences.

Improvements in luminance and color are both nonlinear, starting with sharp increases that gradually become less steep. Gains in luminance affect perceptions of quality, even as the display approaches 400 cd/m^2, but between 200-300 cd/m^2 the return on improvement begins to diminish substantially. Improvements in color gamut, however, result in continuous strong improvements in perceptual quality, up to 120 percent of the Adobe RGB standard.

This example assumes ambient light of 310 lux. The results would obviously be affected by less favorable lighting conditions, such as a phone used outdoors or a television viewed in a bright room.

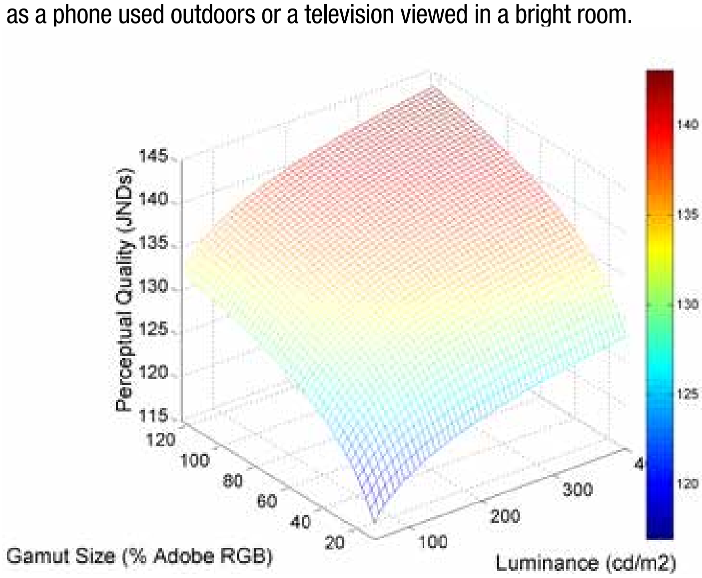


Figure 3: Model of interaction of luminance and gamut area with perceptual quality. A larger perceptual quality value indicates higher preference. Display modeled as 46-inch LCD TV with 1080 resolution and viewing distance of 1.5x the display diagonal (69 inches). With increases in color gamut or luminance, improvement in perceived quality is nonlinear.

In general, PQM suggests the highest luminance and a low color gamut will generate an acceptable quality value, but superior quality values are not achievable without a higher color gamut.

Furthermore, color saturation can be used to maintain high values if the developer opts to lower another performance characteristic. For example, if a developer sought to improve a display's energy efficiency by lowering luminance, expanding the color gamut can maintain or increase the display's quality value. Increasing color gamut can achieve excellent quality values, even at mid-range (250-300 cd/m^2) luminance levels.

Note that the same JND score can be achieved with higher color/lower luminance or lower color/higher luminance. As luminance drops from 350 cd/m^2 to 280 cd/m^2, perceived visual quality can be maintained by increasing the color gamut size from approximately 50 percent to 60 percent of the Adobe RGB standard.

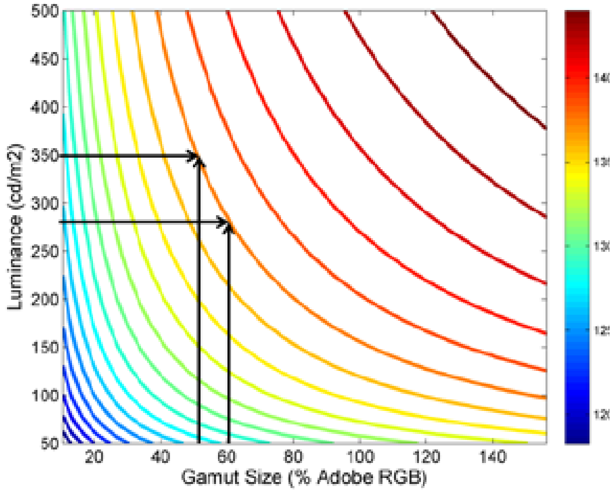


Figure 4: Isoquality curves for display quality show interaction between gamut size and luminance. Display modeled as 46-inch LCD TV with 1080p resolution and a viewing distance of 1.5x the display diagonal (69 inches). Note that the same JND score can be achieved with higher color/lower luminance or lower color/higher luminance.

Device comparisons

Figure 5 illustrates how these tradeoffs could apply in hypothetical devices with considerable variations in performance attributes. In this example, four models are analyzed using PQM: the first two models represent first generation devices, while the other models represent second generation models with better resolution. The model assumed a fixed viewing distance of 10 inches for all devices.

Models within each generation have significant differences in luminance and color gamut, but they achieve similar PQM values. In fact, first generation devices “A” and “C” attain identical PQM values, despite significant differences in color gamut and luminance. The same is true of second generation devices “D” and “E”.

Device Generation	Model	Display Diagonal	Resolution	Luminance (cd/m2)	% Adobe RGB	PQM Value (JNDs)
1	A B C	3.5” 4.0” 3.7”	960x640 800x480 854x640	541 365 449	53.13 109.38 69.79	131 129 131
2	D E	4.0” 4.8”	1136x640 1280x720	556 283	72.92 111.46	134 134

Figure 5. Different display specifications can attain similar perceptual quality values.

PQM can also reveal when additional improvements in a performance attribute will produce little or no change in consumer perceptions of quality—that is, when the attribute has been “maxed out.” For example, PQM analysis indicates that significant gains in the perception of quality can be achieved by upgrading resolution from 1080p to 4K. This was especially true when the display size increased and the viewing distance was held constant. The benefit from upgrading to 8K was negligible, however, even for the 65-inch display.

Impact on color management and content

PQM does not necessarily encourage the use of larger color gamuts. However, it does demonstrate that—if resolution and luminance are held constant—larger color gamuts will improve the perceived quality of the majority of consumer displays. It also suggests that higher gamuts can compensate for decreases in other performance attributes such as luminance and resolution.

The authors believe this demonstration of the power of color, combined with new enabling technologies such as quantum dots, will lead to more displays able to express a larger color gamut.

This, in turn, will have repercussions for the display industry:

First, a renewed emphasis on color management is likely. For years, many operating systems and programs have had insufficient or poorly implemented color management. Instead of interpreting the display's color capabilities, these systems and programs assume the display is capable of expressing colors corresponding to the sRGB gamut. Often, this creates undersaturated images and colors—a problem in any case, but especially on retail websites where images don't match the actual products. The weakness of this approach is exaggerated with high gamut displays or with gamuts that are not approximately the same shape as sRGB. As higher gamut displays become more common due to continued growth of OLED and the emergence of QDEF-enabled LCDs, well-executed color management at the system level becomes much more critical.

Second, a heightened color gamut could influence content. Once content providers have the ability to use expansive color, they will be inclined to use it. Intuitively, one recognizes that heightened color is preferable to a lower gamut. Initial research confirms this. **A pilot eye-tracking study by 3M suggests that content with a higher color gamut receives more attention, as measured by aggregated fixation or dwell time, than content with a lower color gamut.**

This is not to say that content providers will feel compelled to use a high color gamut—or color at all—in every circumstance. However, the tool of a higher color gamut will be used once it is available.

Additional impacts should become obvious as the PQM model is refined and extended. In the coming year, for example, the authors plan to expand the model by adding a measure of contrast. Further improvements could include a validation of the model for outdoor use and for video and animation.

As of this writing, 3M is evaluating how to make PQM available to its partners and customers, as well as to the broader display industry.

PQM in action: Is 4K resolution worth the cost?

The value of PQM is greater than assessing the impact of improving one performance characteristic while constraining another. The values generated by PQM can also reveal when an improvement will produce little or no change in consumer perception of quality—that is, when the benefit is "maxed out."

The latest models of ultra-high-resolution LCD televisions—4K sets—provide a good case in point. Objectively, the displays are a significant improvement; current HD sets with 1080p have one-quarter the resolution of the 4K sets. However, some reviewers have questioned whether that much higher resolution would translate to a meaningful increase in consumer perception of quality.

Based on PQM, the short answer is "yes" – especially among the largest displays – but the effect diminishes quickly after 4K, making 8K resolution a less attractive improvement.

At a viewing distance of nine feet, PQM suggests that on any display of 32 inches or more, the improvement from 720p to 1080p results in a meaningful improvement in perceived quality. (Figure 6.) On 42-inch and larger displays, the improvement from 1080p to 4K resolution creates a meaningful difference in perceived quality; the difference is dramatic for sets that are 55 inches and larger.

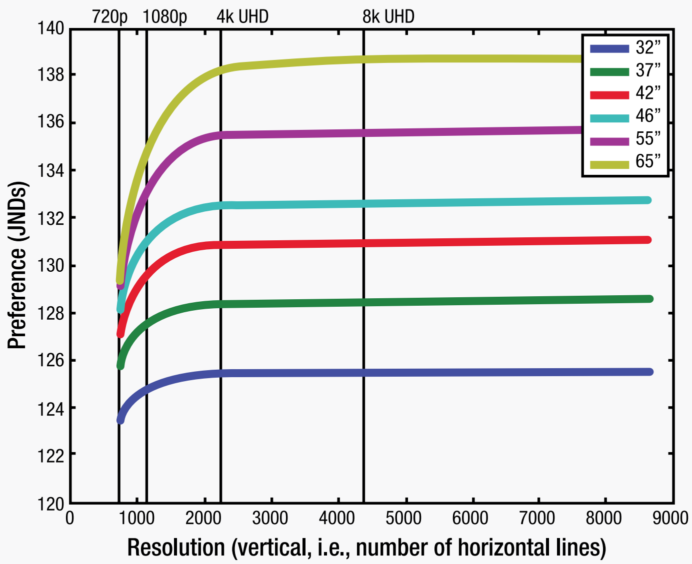
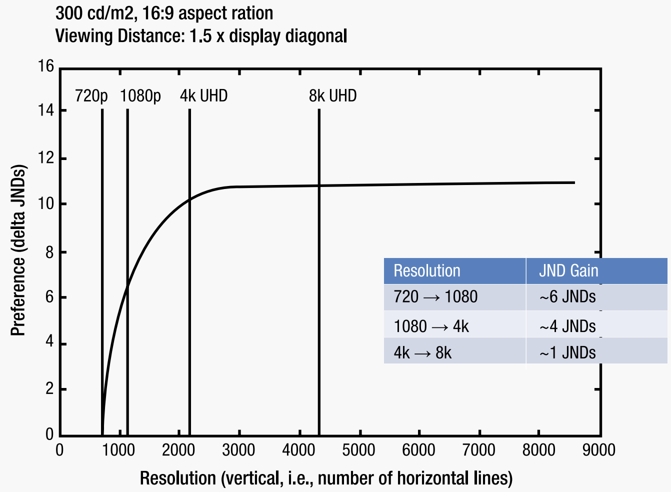


Figure 6. At 300 cd/m2 and a viewing distance of nine feet, improvement from 1080p to 4K creates a strong increase in perceptual quality of larger TVs. Improving resolution from 4K to 8K creates a measureable but much less powerful increase in perceptual quality.

At either range (nine feet or 1.5 times the diagonal), there is a measureable, but considerably less powerful increase in perceived quality when the resolution improves from 4K to 8K.



Resolution	JND Gain
720 → 1080	~6 JNDs
1080 → 4k	~4 JNDs
4k → 8k	~1 JNDs

Figure 7. At the recommended viewing distance – 1.5 times the display's diagonal – the improvement from 1080p to 4K creates a strong increase in perceived quality. When the resolution improves from 4K to 8K, there is a measureable, but considerably less powerful, increase in perceived quality.

Obviously, the introduction of displays with ultra-high resolution raises issues about content and infrastructure, specifically, the bandwidth necessary to allow video streaming. PQM does not address these important considerations. That said, product developers might find some value in the metric's analysis, which concludes that ultra-high resolution does appear to improve consumers' perceptions of quality, up to about 4K, especially on larger (42-inch and above) displays.

Supplementing PQM

PQM demonstrates how product developers can manage perceptions of quality as they alter performance characteristics. Final decisions on how to configure those performance characteristics will usually be determined by cost or the manufacturer's technical capabilities. In some instances, the application—how and where the display will be used—provides additional information to supplement the PQM analysis and guide the configuration of luminance, color, size and resolution.

This is particularly the case when the display is presented in a retail or other environment where viewer attention is crucial. Here, initial research shows, a display with a higher color gamut will receive more attention than a lower color gamut display, even when the two displays have similar PQM scores.

In a pilot study conducted by 3M, five subjects were shown nine colored images. Each image was manipulated to produce four different color gamuts: standard RGB (sRGB), saturated green, saturated red, and saturated red and green.

The four color gamuts for each image were displayed simultaneously for one and three seconds to avoid scanning heuristics; placements were also varied to eliminate location artifacts.

Eye movements were tracked and the time of fixation on each image was aggregated. During the three-second trial, subjects fixated on the saturated green and saturated red and green images longer than the sRGB images. Results of the one-second presentation showed a similar pattern in mean fixation duration.

This demonstration of the power of color, combined with new enabling technologies such as quantum dots, will lead to more displays able to express a larger color gamut.

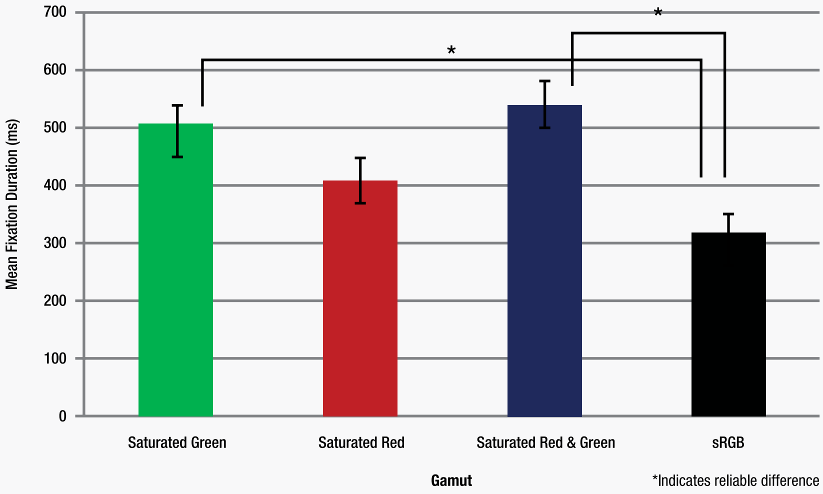


Figure 8. During a three-second trial, subjects fixated on the saturated green and saturated red and green images longer than the sRGB images.

Given the sample size, it is difficult to draw detailed conclusions on the interaction between specific images and color gamuts. Also, the device used in this pilot study had a red primary that was only slightly more saturated than the sRGB standard, which likely had a negative impact on the "Saturated Red" test case.

Despite these limitations, the study has implications for product developers. It suggests, for example, that the perception of overall quality should not be the sole consideration when designing a display for use in digital signage. A display with a high color gamut will attract more attention (as reflected in longer fixation times) than a display with a lower color gamut. Likewise, content developers should consider the relationship between attention and color saturation as they choose the icons and images used in retail displays.

The authors are currently considering additional research, using a larger sample and more capable devices, which will be able to draw more detailed conclusions on the relationship between fixation time and color gamut.

—

Barten, P.G.J. (1987). "The SQRI method: a new method for the evaluation of visible resolution on a display." *Proceedings of the Society for Information Display,* 28: 253-262.

"The truth is, as nice as these TVs are, you probably won't see much difference. A 60-inch 4K won't look dramatically better than the 1080p TV you have in your home right now unless you shove your nose into the screen. The average person's eyes can't see the difference when sitting 10 feet away from a 60-inch TV." http://www.wired.com/gadgetlab/2013/01/the-4k-push-ces-2013/

See also "Why Ultra HD 4K TVs are still stupid," http://reviews.cnet.com/8301-33199_7-57566079-221/why-ultra-hd-4k-tvs-are-still-stupid/

3M Optical Systems Division
3M Center, Building 235-1E-54
St. Paul, MN 55144-1000
U.S.A.
1-800-3M HELPS
3m.com/displayfilms

Printed in U.S.A. dz15245