Summary of Feedback on Three ACES 2.0 Output Transform Candidates

HDR / SDR Google Form SDR only Google Form

Responses as at 17th August 2022

HDR Responses: n = 9, 4 x Baselight, 5 x Resolve SDR Responses: n = 6, 5 x Resolve, 1 x Nuke

HDR users anonymized as H1, H2, etc. SDR users anonymized as S1, S2 etc

Setups

Of the HDR respondents, 8 had dedicated SDI / HDMI cards, and 1 used direct GPU output. Displays were 3 Sony X300s, 2 HX310s, 1 FSI XM311, 1 Dell QP3221Q, 2 Eizo ColorEdge and 1 LG C8 OLED (note, there are 10 monitors as one user had an X300 and an Eizo cg3146).

Of the SDR respondents, 3 had dedicated SDI / HDMI cards, and 3 used direct GPU output. Displays were 1 BenQ, 2 Dells, 2 Eizos and 1 Sony HX310. (Unsure why the HX310 user (S4) did not evaluate HDR as well)

All users graded in ACEScct, except the Nuke user who used graded ACEScg. The HDR users were evenly divided between using color space aware grading tools and simple lift /gamma / gain. Of the SDR users, only one used color space aware tools.

All the HDR respondents were in calibrated grading suites, except the LG OLED user (H5) and the FSI HX311 user (H6) who were in dark home environments.

Of the SDR respondents, S1, S2, S3 and S6 were in home / office environments, and S4 and S5 were in calibrated grading suites.

Tone Curve

All respondents felt the mid-tone contrast was "about right". Most felt the highlight contrast was "about right", with two HDR respondents and one SDR saying it was too high, and one SDR respondent saying it was too low.

There was less agreement about shadow contrast, with five HDR and two SDR respondents saying it was "about right", with all the rest saying it was too low, except one who felt it was too high.

All respondents said they could adjust the tone curve to where they wanted it, with a fairly even mix of "usually raising" and "usually lowering".



Do you feel the contrast was appropriately similar / different in the HDR and SDR versions?

Any other comments on contrast?

H1: It definitely feels a little crunchy in the low end by default. But this may just be a reaction having worked in the current aces for so long

H2: Question 3 - I didn't adjust the contrast, I just viewed the Candidates. Question 5 - the highlight contrast in HDR is good, but in SDR you've pushed the highlights down without affecting the midtones so much. The result is a "grey film" over the highlights compared with ACES 1.1. I'm mostly only interested in skintones and I appreciate the need for more highlight detail on cheekbones (IMHO this is where film shines over digital). I feel like you are going in the right direction with that, but you need to A) preserve the color B) drop the midtones somehow because right now faces are looking flatter compared with ACES 1.1 (again this is only in SDR - all candidates). Hope I'm making sense!

H3: All 3 candidates overall had better contrast than the current ACES 1.2

H4: contrasts are step forward compared to the original ACES Output transforms

H5: I don't recall any issues with contrast, but will have to go back and look closer. I recall all the gray charts looking alright. On grading I found some difficulty between the contrast controls and saturation control (mostly in the HDR scenarios.) Overall the tonality and contrast seemed to work well, but I plan to try on some B&W images to investigate further.

H7: I really like how version B reacts on contrast, but contrast only. I think it's oversaturated

H9: The current ACES's starting level contrast is much more smoother and there is more flexibility/wider range to adjust than the candidates. Compared to the current version, the blacks are darker and highlights are much brighter.

S1: I think it's the perfect starting point for going any direction depending on type of content.

S2: Starting point could be tiny bit lower contrast

S3: With the same contrast applied, candidates B and C generates more noise. Especially in the darker blues.

S4: ACES2 highlight rendition looks better compared to ACES1 but now slightly to low. Shadow contrast feels a bit too crushed. This is something I never liked about ACES1 and now it seems worse.

SDR vs HDR Matching

(from 1: "not at all" to 5: "very well")



How well did the SDR and HDR "match" for candidate A? 9 responses

How well did the SDR and HDR "match" for candidate B? 9 responses



How well did the SDR and HDR "match" for candidate C? 9 responses



Any general comments about SDR / HDR matching?

H1: Candidate b is very saturated in hdr. That saturation is not retained or represented in SDR. On some material that is quite saturated red I was viewing, the reds that felt more "orange" in hdr went pink using sdr

H2: I felt the chroma and hues matched well between the 3 candidates SDR vs HDR, but I feel like the previous SDR (version 1.3) matched the tonal feeling better. I feel a greyish, flattish thing going on in the highlights in the new SDR versions. I imagine you are trying to bring more detail in those highlights and I do see a marginal improvement - but it seems to come at a cost in terms of lost tonality (the speculars seem globally lowered).

H4: Candidate c from SDR to HDR lacks a lot saturation

H5: *A*, *B* & *C* all seemed to have more grain with rec2100 then with the 709 or 709sim (puzzled as to why???). The rec2100 had very good match using Dolby Vision to create the

HDR and SDR. A, B & C behaved differently in each scenario, but the 709 and 709sim were closer in their differences. A, B & C were behaving different comparing 709 SDR with 2100 HDR, And 709sim SDR with 2100 HDR differed differently. For me the main differences seemed to involve saturation compressions and hue shifts and maybe some luminosity shift (but lum might have been from the color issues.) Again issues between A, B or C were consistent in 709 and 709sim but varied with 2100.

H7: I much more like the colour matching on candidate A and contrast matching on B

H8: differences only really get visible in extrem lighting. matching for natural live images is generally very good. candidate B does not handle extrem colorful lights very well especially in SDR.

Hue Consistency

In general, all respondents felt that hue was fairly well to very well maintained in all zones by all candidates, with a little more disagreement about hue consistency in highlights.

Most respondents did not report hue changes with any particular grading operations, although a few SDR respondents reported some, although with no clear differentiation between candidates.

If you answered yes to any of the above, please give more detail about which grading operations in particular

H1: Pushing contrast around on b. I have a saturated blue wall I was grading. It seemed to shift between green and magenta subtly when exposing up and down

H7: brightness and contrast moves skintones too much for me, are too desaturated and I need to adjust them

H8: Candidate A: Blue shifts into cyan when raising exposure. Also green shifts slightly into cyan when raising exposure. Candidate B: Blue shifts into violett when raising exposure. so does Red to a smaller degree. C seems to hold hue better while changing exposure

S3: S-curve contrast. Frame 0018 for example or the red of the hot chile sauce on frame 002.

S5: All three candidates shift the hues more or less depending on the color. Some colors are more difficult to put back into place.

Only a couple of HDR respondents reported constantly making the same types of adjustments, but half the SDR respondents did. Particularly for candidate B.

If you answered yes to any of the above, please give more detail on the adjustments you are making.

H5: As there are many ways a image can look, I will try to give a couple general tendencies and one hue problem. In general for 709 and 709sim saturation seemed to go A weakest to B to C strongest... with B sometimes stronger for orange. In 2100 C gives the best skin, both dark and light, of any of the scenarios. However in both 709 A and 709sim A look to have the best skin for those groups although not as good as 2100 C. A major issue with C is that the blue is pushed toward cyan. At fist I thought it was just sat compression, but there is definite hue change (same sat) seen on vectorscope. Grading under this means there will not be any blue. Also with this cyan push there seems be some lum increase which in some of the images is quite obvious.

H6: Basic lift, gamma, gain, contrast, & saturation adjustments seem necessary from shot to shot to match luminance & contrast,

S1: Can. B -> Skintones feels slightly 'bendy' towards yellow/green from shadowed neutral to more yellow/green midtones to neutral highlights. Can. C blue lights render too cyan even when their exposure is lowered.

S4: Close to gamut colors, highly saturated colors seem to suffer with Candidate B. Tungsten practicals or fire turns quite brownish.

S6: I find all benefit from increased contrast, saturation and gain as a starting point.

Hue behavior was almost universally reported to be consistent between HDR and SDR for all candidates.

If you answered no to any of the above, please give more detail on the differences you are seeing.

H4: Candidate c lacks in color volume in HDR

H7: A seems to have different colour temperature between SDR and HDR versions

Please add any other comments regarding hue here

H2: Congratulations on the new handling of extreme Blues and Fuschias. Very nice! My preference by far is for Candidate A. Candidate B presents a big color boost that doesn't work and C is too desaturated/orange. Candidate A seems more consistent with what version 1.3 was doing anyways in terms of hues. That said, there is a shift towards pink overall in Candidate A that I find less attractive than HDR v1.3. If I had to choose today, Candidate A would be a big improvement over v1.3 for the details/tonality in the the HDR extremes. It crushes a bit in the toe, but that can be corrected easily. So can the pink shift (and I do hope you can remove that).

H4: All candidates compress reds little bit more towards magenta in SDR and HDR which is not favorable. A and B produce in SDR and HDR denser and kinda more neon blues which

is not what I expected to see. C creates lighter blue which almost get noisy. But none of them create out of gamut colors or lookalikes, which is very positive. Although hues were consistent between HDR and SDR there seemed to be some inconsistencies of those consistencies from A to B to C. This might be more related to saturation. Certain images would behave differently whether in A, B or C even though consistent from HDR to SDR. (for example: frames 0040 or 0005)

H5: Candidate C seems to be very good at holding hue while applying exposure changes.

S1: Can. C: I like the fact that colored lights stay very linear when they go very bright. However Can. A's default 'bendiness' to me looks better for fire and is easily adjustable for bright colored lights if desired.

A curious phenomenon I encountered was on the vectorscope all candidates 'jiggle' when the exposure exceeds it's boundary and start bending/folding back/desat which is easy to spot with exr 0062 (rec.2020 sweeps) of the provided samples. Optically I don't experience this as an issue but current ACES doesn't really have this. It's mostly visible in the blue/cyan region and Can. B even has a bit of this in normal exposure ranges. Of course I'm also not sure how accurate Resolves built in vectorscope is.

S3: Candidate C is breaking the signal with some hues. Please check the red T-shirt on frame 0019 with Candidate C as the ODT (Resolve ACES project). Candidate B also but not so strong. Candidate A is performing better. No grading applied.

Zone Saturation

HDR Respondents

Was the saturation in shadows:



Was the saturation in mid-tones:



Was the saturation in highlights:





How easy / hard is it to adjust (add or remove) saturation where you want it?

Do you find yourself consistently adjusting saturation one way or another?



SDR Respondents

Was the saturation in shadows:



Was the saturation in mid-tones:



Was the saturation in highlights:



How easy / hard is it to adjust (add or remove) saturation where you want it?





Do you find yourself consistently adjusting saturation one way or another?

Any other comments on saturation?

H2: While I agree in principle with brining adding richer saturation in Candidate B, the implementation just feels like a wash (too much "juice" on skintones), so I prefer candidate *A*. That said, the latter could have less saturation in the shadows, and less pink overall (I know, that's a Hue comment but a big deal for me ;)).

H3: Candidate C dealt with Saturation generally pretty well. Candidate B was far too saturated

H5: 2100 C seems to have more connection of sat and lum making it a bit more difficult to adjust. Compression of yellow, orange red and blue were noticeable and worked best in 2100 C. In 709 and 709sim sat compression was not as nice. In general for 709 and 709sim A sat was weak and C strong and sometimes none right. In general for 2100 C was best. However, there is a problem with the blue in C being pushed to cyan.

H8: Saturation in shadows seems a bit high for candidate A and B - a bit like the current HDR ODTs in ACES 1.3. I almost always desaturate them a bit. I did not do so with Candidate C - I liked the starting point for shadow saturation in C. Candidate C preserves saturation in the highlights very well when going from HDR to SDR. I liked that in Candidate C there is a bit more highlight saturation in both HDR and SDR.

S2: With candidate B and C boosting saturation boosts skin tones more than expected, more than candidate A, potentially causing more work with skin tones. Candidate C has color cast in neutrals like whites and greys, absent in the other two candidates. With color space aware controls it's easy to get the saturation to desired level with all candidates, in most cases.

S3: "Oranges" are too saturated on Candidate B.

S6: *I* find Candidate B midtone saturation is too high, plus it feels like there is a shift towards red/orange. It really stands out against A and C as well as vanilla ACES 1.2.

Conclusions



Overall how easy was each candidate to grade under?

It appears that there is less agreement about how easy Candidate C is to grade under, presumably due to people's unfamiliarity with the non-linear responses which occur when grading under a CAM. More people found Candidate A the easiest to grade under.

Preferences

HDR Respondents



Indicate your preferred candidate for SDR and HDR

If you could only choose one candidate, which would it be? 9 responses



If you were to eliminate one candidate, which would it be? 9 responses



SDR Respondents

If you could only choose one candidate, which would it be? 6 responses



If you were to eliminate one candidate, which would it be? 6 responses



While there is a clear majority in favor of eliminating Candidate B, the preference for A and C is split, with respondents looking at SDR only preferring Candidate A, and those who were also looking at HDR preferring C (presumably due to its better CAM based HDR / SDR matching).

Was there anything which drew you towards or put you off one particular candidate?

H1: *B* is too saturated and does strange things to fire. So does a tbh. C treats colours well but is slightly desaturated and what we're fleshy magenta skin tones seem to have a bit of green creepy in. I think I also may be missing some of the twists from aces 1.0 so a Imt would be great.

H2: C is too desaturated by far.

H3: Candidate C in most images didn't push the saturation too far and wasn't skew too much to one direction, except white was ALWAYS pushed to yellow tones, but overall was the best out of the 3.

Candidate A just went pink all the time and seemed to have some sort of clipping affect to most red values. It was only fantastic at Blue tones every other transform had issues with blue.

Candidate B was pushed too far in all directions, much too saturated and hue shifted with highlights blown out. I would throw this one out completely.

H4: B looked overall most complete

H5: Out of all the frames included and some of my own candidate 2100 C worked best. Also keep in mind that I used 2100 with Dolby for both HDR and SDR. If no 2100 then next would be A, and for only a couple frames B. (This is judging without any grading.) Most of the time B seemed to have something wrong with the saturation in comparison to the others. B also seemed to have some smooth sat/hue issues (seen in orange in frame 0001.) (Also note that I found all candidates to be good with neutrals and tonality.) All seem to have difficulties with the blue and green screen hues, although pulling quick keys seemed no trouble. The lack of blue in C is unacceptable, but I expect this can be fixed. Also note that although I select 2100 C as best, C in 709 and 709sim is not so good and at times less than A or B. There is also an issue that there might be what looks like "grain" in 2100 (A, B & C.)(check out frames 0009 and 0054.)

H7: I don't like colder tones in highlights SDR candidate A

H8: candidate B handled extrem lighting situation not as well when going from HDR to SDR. shadow saturation was a bit high for B and A. HDR/SDR match for candidate C was very good and C was very stable in Hue when applying exposure changes. Also in chart no. 61 the upper row of colors was more uniform in luminance for candidate C - blue was much darker in A and B. That makes Candidate C for me a better starting point for Look-LUT development and work with luminance curves. Shadow saturation looked very natural in C. Also when going from HDR to SDR in C Saturation in Highlights in SDR were well preserved without showing artifacts.

S1: Can. B was unpleasant to start off with because I always felt I needed to fix the skintones which Can. A gives a very pleasing result out of the box. As mentioned before I liked the qualities C gave with regards to colored lights but I could achieve a similar look easily with A whilst having more qualities that I preferred in most scenes, especially skintone.

Can. C felt had quite saturated brights which also exaggerated tones close to neutral white making it trickier to get a clean look from a white-ish scene unless you were spot on with balance or use means of desaturation. I felt this was especially evident in aces_1920x1080_graded.00320.exr from the old big pool. But less noticeable overall.

Cranking saturation from 50->100 gave the most pleasing result with Can. A in the extreme scene 0063 (lightsabers) where it felt that color gradation was best maintained blending across hue and saturation. In most scenes Can. A feels to have a slight bias towards blue

when increasing saturation. Can. B very much towards reds and makes blues appear more magenta ish. Can. C feels the most balanced across all. My preference here would still be A.

S3: Yes. Candidate A is less saturated. More clean when treating with "neon colors" (Frame 004) and the highlights. Better skintones (Candidates B and C are have more saturation).

S4: A in general feels quite balanced already, but I would say it's feels a bit too muted in terms of saturation. Blue tones (sky) seem to be quite dense which I actually like, but it shouldn't be that dense for a general viewing transform. Skin looks most pleasing out of the tree candidates.

B is probably my favorite, but the hue shifts in highly saturated colors are a real issue for me. Rendering of sky is better and skin tones are fine.

C feels too saturated for me and the colors tend to become quite exaggerated and bright. Skin tones become a little too yellow.

S6: I love C. It feels like the image is "ready" out of the box and just needs gentle tweaking to bring out its best. A is a "blank slate" that is very flexible, but needs a little extra work to get a final image. B feels oversaturated and has a color shift towards orange/red which I found not desirable.

Any Other Comments?

H2: Super well done on the out of gamut blues and fuschias. I think A is a winner if you remove some pink. Can you just keep version 1.3 colors/hues?

H3: Current ACES and all 3 candidates could not deal with yellow-orange-ish tones well at all. It either went towards primary yellow or straight to red-ish orange.

H4: I think the contrast ratios developed to a nice direction, but the compression of saturation is somehow unnatural (especially C). The compression of reds towards magenta and higher saturation is not favorable.(all candidates). Skin tones are not a step forward, I prefer in that regard the original ACES transform. Blues tend to get too dense and too neon (A+B). all candidates could have more saturation in HDR highlights. Overall no problems with out of gamut colors, which is great.

H5: I basically set up color management and first compared the triplets (A, B, C) for each group. Then I compared the 709 and 709sim in SDR. Then I used Dolby for 2100 HDR and SDR and compared. Then I compared HDR from 709sim and 2100. I built some 3 node binary switches to easily and quickly a/b selections. I used all the frames supplied plus a few of my own. I then started to grade some. To answer the questions I did some quick overall fake grades watching the scopes (also had three vectorscopes for high, mid and low.) I answered grading as easy but those question's answers should really come from well experienced pro graders who know how long things should take. Overall I found 2100 C to provide the best start. However, there is a blue shift towards cyan that needs to be fixed. I can appreciate that some may wish to have film look out of the box... but that is not what this ACES ODT should be about. This should be a technical transform that provides the best fit into the output color space for all the data captured. So yes, there will need to be compression to fit things in. And I realize that how far into the space the compression begins

and what form it takes and even what if any stuff is clipped (like colors no human sees) are part of this design. But to alter color that is seen (in this case blue) to adjust some "look" is not acceptable. I realize that many worked hard to get a certain look with film. Film has its own certain limitations and character. But the world we can see has much more and displays being manufactured now have more (not all we can see but approaching 2020.) The ODT must give all what we can see that a display can produce. If one then desires to limit that they can add a specialized LMT, LUT or whatever. Putting that blue back under that ODT is not workable. Now that I've noted blue... I also found that the red, although a seemingly nice compression, seems to also darken (for example frame 0029.) To me this seems out of place and the red should be a bit brighter. I realize the deeper the red the darker, but it seems a bit too dark. Again I don't want to fight the ODT. And lastly (but not finally) I would like to suggest a change in the color space from AP1 to that of the new ARRI log-C4. The reason being that more of the visible spectrum is covered with a minimum of unseen colors. AP0 is still fine, but AP1 is maybe ready to retire.

H7: Thank you for making me participate in your questionnaire.

H8: In the next round - if there is another - I would like to see a P3D65 gamma 2.6 ODT for dark surround SDR cinema grading please.

H9: While there seem to be less green in candidates A & B, the starting level of contrast has become stronger. The saturation in candidate B looks a bit too high. For candidate C, the situation-where certain colors in the highlights break-has been improved but overall color balance seems out of place.

Also, the highlights of all A,B,C candidates are pretty strong.

While the current ACES itself is already great, I hope that this new ACES 2.0 would really improve the aspect of greenness and the distorted highlights of certain colors that were noticed in the current ACES.

S1: *I* still have a lot to learn as a colorist so perhaps some of my conclusion don't fully make sense but these are my observations. I hope they provide some use and best of luck with the development.

S3: It is nice to see an improvement with the "out of gamut" colors and the treatment of the highlights (less clipped with a nice roll-off soft). From candidate C I like the hue in the blues but I would not implement that in an ODT but in the grading. I think Candidate A is the best one. Less saturated and less hue shift. Frame 0023 the lighting inside the pumpkin for example. Candidate A is dealing better with the borders. But in frame 0030 is creating noise in the red t-shirt similar to the actual REC.709 ODT. With a better performing in certain colors to avoid artifacts it is a great option. Nice job. Congrats to the entire team for this work.

S4: I very much respect the effort towards finding a better RRT for a future ACES version. I also welcome the move to conduct open testing and gather feedback from the professional community about this.

However, none of the proposed candidates is perfect yet. I couldn't choose one over the other if I had to. They are not good enough yet to replace the old RRT. My number one concern - which was not addressed at all - is the fact that the new RRTs are still very sensible to highly saturated blue colors. If ACES2 still relies on the gamut compression patch then I would say it's not worth the hassle. Another goal should be the perfect reversibility of the RRT. I can't see from the LUT you supplied if this has already been achieved, but it should be a must have quality of a new RRT. The improvements on the look level go into the right direction, but are a small gain given the compatibility issues the industry is likely to suffer.

Are there any other questions we should have asked?

H2: Which candidate has better color separation? I personally find that one hard to tell because some images would yield B (perhaps because everything is more saturated?), and some would yield A for being more subtle.

H3: What looks good out of the box. Or which candidate deals best with highlights/clipping.

H5: Yes, I found significant differences to comparing A, B & C depending on whether from the groups 709, 709sim or 2100. Some questions were hard to answer because of this and I had to generalize giving a less accurate answer. Also it would have helped beforehand to provide more instruction on what information might be specifically needed. For example, to grade in a way to look for reactions. Questions could have been added to specifically evaluate the compression of various hue saturation compression. The high sat in some of the sample frames can be evaluated. I found some differences which showed C as preferable in the 2100 series.

S3: Maybe how are you doing the saturation adjustments?

S4: How big of a concern is reversibility for you? Should the new RRT be solid enough to work without any form of gamut patch?