Proposal to make the validation process go more quickly

Velocity has always been an opportunity area for VQEG in how we execute our work. One area that the ILG wishes to accelerate is the VQEG objective algorithm validation process. The ILG proposes that the AVHD project try a different division of labor, to try to foster speed and efficiency.

This proposal is intended to be discussed and implemented within the AVHD project. This approach, if successful, could be adopted by other VQEG projects.

# Opportunity Area:

* Speed-up writing validation test plan
* Speed-up implementing validation test (i.e., model scope is obsolete technology)
* ILG has limited internal funding to supplement proponent fees thus other funds are required
* Proponent has limited funds to pay substantially larger fees so new structure needs to be identified
* Reduce the total amount of time from execution to report to be less than our traditional 1 year or more
* Reduce the political thrash around unnecessary positioning around design and purpose leading to a smoother process

# Goal:

* Increase speed of entire validation test cycle
* Minimize time after model submission
* Keep fees low
* Rebalance cost / benefit ratio for each organization (ILG & proponent)
* Retain value of VQEG approach (e.g., independent validation, robust variety of HRCs)
* Ensure adequate pool of secret source content for testing
* Drive standardization and visibility of work that VQEG is producing
* Accommodate low numbers of proponents (2 to 4)

# Proposal:

* VQEG specifies “Best Practice” documents. AVHD has already begun drafts:
  + Definitions
  + Subjective Test Method
  + Model Requirements
  + Data Analysis
  + Report Structure
* Proponent should write & propose a test plan
  + Although anyone could write & propose a test plan, proponents are best situated to understand the appropriate scope and application of their models.
  + This should be a concise and to the point document
  + Focus is new issues to be discussed (e.g., number of tests, type of HRCs)
  + Where possible point to a “Best Practice” document previously approved by VQEG
* VQEG discusses the proposal during a face-to-face meeting, audio call, or interim meeting
* VQEG schedules enough time (during the meeting) to decide and commit

# Proposed “Best Practices” for ILG & Proponent Roles

VQEG could write and approve “Best Practices” for ILG & proponent roles based on any recent validation test (e.g., RRNR-TV, HDTV, Hybrid Perceptual Bit-Stream). However, those Best Practice documents would not address the problems of how to speed up the validation process. The following is a proposal for a new division of tasks between ILG and proponents that is intended to address all of the above opportunities and goals. The following plan is intended for two or more proponents.

The ILG welcomes alternative ideas that address the opportunities and goals identified above.

## Idea:

* Time critical ILG tasks occur before model submission
* Time critical tasks after model submission performed by motivated proponent
* Offload from ILG to proponents some expenses and time consuming tasks

## Project Co-Chair Role:

* Lead a discussion on the proposed validation test plan proposal
* Reach agreement with VQEG (e.g., approve test plan, approve modified test plan, reject test plan)
* Decide upon a schedule
* Edit and publish the validation test plan (if approved)

## ILG:

* Role: oversight, supplies some new / secret source, select scenes, design subjective tests, redesign subjective tests if unbalanced, perform data analysis, verify model values, and review final report
* Omit: ILG does not run subjects, create PVSs, write most of the final report, or subcontract running subjects for a proponent (see below for alternatives)
* Add: ILG to write and publish a “Subjective Test Design Guideline” document by model submission deadline.
* Timeline: ILG does scene selection before model submission
* Timeline: ILG designs valdation subjective tests before model submission
* Goal: eliminate ILG deadlines after model submission that have the potential to slow the time line down due to competing priorities

## Proponents:

* Role: organize coordination audio calls, create HRCs, implement tests, report quality imbalance to ILG (if test needs to be redesigned), run subjects (after model submission), write final report
* (Add) Each proponent
  + Supplies new source / secret source under CDVL terms, DRM free, or public domain (well before model submission) Proponents are required to submit a minimum amount of footage for use in the validation testing (amount to be determined). [Issue to be discussed; no agreement on terms of source scene distribution (this test only, NDA, all tests, etc.); Concerns were expressed that it may be too expensive for proponents to supply source footage under CDVL terms, DRM free, or public domain that is appropriate for validation testing.]
  + Supplies several HRC options (e.g., variety of software and/or hardware, hire company to coding video)
  + Informs the ILG about capabilities to create HRCs, give example PVSs (well before model submission)
* Constraint: model dropped from consideration if subjective test not submitted on time
* Fees pay for time consuming ILG tasks (e.g., scene selection, secret source scenes, test design/redesign, verify model)

## Notes:

**ILG running subjects for a proponent:** An ILG who contracts to run subjects for a proponent must withdraw from the role of being an ILG for the that validation test.

**Other labs running subjects for a proponent:** Proponents can buy the service of running subjects from another lab (e.g., if a proponent cannot do it themselves due to facilities or expertise). This must be reported to the ILG. The ILG will oversee such agreements. The fee terms are negotiated directly between the test lab and proponent. Such agreements must be open so that everyone knows about this agreement (but not the costs).

**Source Video Sequences:** A complication with proponents providing source video for the Multimedia Validation Test was licensing terms. This proposal requests usage terms that avoid the involvement of lawyers. An intended by-product of this proposal is that VQEG should build an increasingly large pool of source video sequences that can be used for future validation testing. See also notes above.

**Proponent Deadlines:** Naturally, deadlines after model submission can be delayed if the delay has wide support (e.g., unforeseen events, unrealistic deadlines).

**Final Report:** It should be possible to write most of the final report by identifying the appropriate “Best Practice”, “validation test plan” and “subjective test design” documents. This will shorten the length of the report considerably. The report should mainly consist of the following sections:

* Executive summary
* ILG official data analysis
* Scatter plots
* Proponent comments