An Analysis of the Impact of Device Size on the Quality of Business Video Conferencing

Introduction
With the advent of video devices moving from large room systems to PCs and on to hand-held devices, there is a significant discussion about how video will be used across these devices. This paper is intended to discuss a critical issue that emerges in this area; that small devices and their cameras are ill-suited to business video. This paper will first discuss how business video is used and the requirements for video to be of value and then discuss the physical constraints of devices.

The conclusion is that small device video may be of limited business value, but great social value. While this may not be universal, it does represent a significant set of questions about the value of small device video and whether it will be adopted by business users.

Business Video
As discussed in the white paper "Video Value Series #1," the value of video to an interaction can be directly related to both the relationship of the individuals and whether the interaction is sales or task focused. Figure 1 is the model that that white paper defined that discusses how relationship, sales/task, and communications modalities relate.

However, in this case the value of video is in the ability of the video system to deliver the visual clues of understanding and acceptance that are necessary for the "selling" interaction. As this paper documented, in these situations, lack of proper

Figure 1 The impact of relationship and communications goals on media use
reading of these visual clues (or not having them at all in a voice only communications) can result in misinterpretation. In the second paper of the video Value Series, the value of video to middle-management discussions and negotiations was documented.

The question then arises, what is sufficient video for a business conversation? While there are business situations where using video to document something (a video of a broken part or a repair) and sending that in real time may be of value, the focus here is on human interaction. If we assume the value of video is clearly understanding those visual clues, then it can be argued that video MUST have two characteristics to be valuable: scale and resolution.

**Video Scale**
The term Video Scale here refers to the exposed size of the image of the person being transmitted. This can be anything from a simple head shot to a full-body view. What is clear is that a torso view (above the waist to the top of the head) is really required for reasonable analysis of video clues. The terms “poker face” and “body language” refer to a unique facet of human conditioning; while many of us are able to control our facial emotions, often our body language and position gives real clues as to our actual reactions and/or interest. Figure 2 shows a head shot and torso view of the same individual. It is clear that the interpretation of reaction is different between the two images. In the headshot, the look can be interpreted as potentially questioning or disagreeing or even whimsical, the torso shot clearly shows the crossed arms of disinterested or boredom.

In the History Channel video "Secrets of Body Language" (available on YouTube), a discussion of how the elements of body language ??? (not sure what is missing here)...consists of body posture, gestures, facial expressions, and eye movements. Humans send and interpret such signals almost entirely subconsciously." are discussed. The point is that it is much more than the face and the eye movements. In this video, the value of body, arm and hand positions and gestures are also discussed. The value of seeing the entire upper (upper?) body in a static video (versus moving people) is covered in great detail.
Resolution

The second critical aspect of business video is the resolution. For good visual feedback high resolution to see details are important. This generally resolves to HD video, where 1080P is generally the highest resolution available today. However, if you are too close to the display you can see pixels, the dreaded "screen door" effect of video. The challenge then is how to have that great video for most or all of the participants, without having the pixels visible to some.

First, it is important to understand video projection and how it works with the human eye. The human eye has about 2,000 cones (pixels) of resolution in the diameter of the core non-peripheral vision. This is where you see detail. Though the human eye is capable of receiving light through almost a 200-degree arc, the acute vision is limited to about 15-20 degrees. Of the about 7 million cones in the core visual field of that 15-20 degree arc (cones are for resolution, rods are for movement - there are about 100M rods in the peripheral field), a HDTV image captures the circular field of about 3.9 million as shown in Figure 4. As the rods are denser in that 15 degree core, this is probably closer to 4.5 million cones in the average eye. So, when viewed at exactly the right distance, a 1080x1920 (or called 1080 for short) image exactly matches the eye's inherent resolution.
Resolution is critical to detecting "micro-expressions," small changes in facial expression that may last as short as less than 100 msecs. To detect these expressions, the visual image must be of sufficient clarity to see these changes. Figure X shows two high resolution pictures on the left and two low resolution

![Figure 4 The impact of resolution on detail visibility](image)

Figure X shows two high resolution pictures on the left and two low resolution.
versions of the same image on the right, **on at 40% of the pixel width, other are 20%**. This is equivalent to going from 1920x1080 to either 480x320 or 240x160 (note the aspect is different here; it is the pixel density that is important). As can be clearly seen, the left images are clear, while the right images decrease the capability to perceive facial expressions.

Another way to see this impact is as an image gets smaller. In this case the image resolution is the same, but as the image gets smaller it loses resolution as it is seen by a smaller number of your eye cones. Figure X shows the visible impact of reducing the **visible area on the ability** to see the image at a fixed distance. The left number shows the relative scale, the right the relative visible area. Scan around the four pictures looking at a specific visual element that can be seen in the 100% and see when it becomes non-visible. This clearly shows the impact of screen size and viewing distance. Generally the 50% scale/25% area is where visibility completely ends in seeing details.

![Image of facial expressions](image)

**Figure 5 The impact of image size on detail visibility**

It is important to note that resolution is a combination of two factors: the image resolution and the size of the screen as an intended angle of the eyes. In other words, a 1920x1080 image that is seen from a long distance will not be visible at its resolution, but rather based on the distance. Similarly, if you are too close to the screen, you will not see the entire image and the pixels you see will be larger, so ideal resolution is when the resolution is high and the screen is the right distance.
Generally, the distance from the screen is defined based on an angle. Both the Society of Motion Picture and Television Engineers and the THX standards use a reference of 50 degrees as the ideal angle to be from the screen. This optimizes the viewing of a motion picture, which is essentially the same as a videoconference. In this analysis, 50 degrees will be used as the ideal distance for viewing. Figure X shows this distance for screens of 21, 27, 50, and 80 inches. Obviously the larger screens are for room use, but this shows the range.

![Figure 6 Relative viewing distances for a variety of display sizes](image-url)
Telepresence - Ideal Video

The telepresence solutions that were deployed by many large enterprises is an excellent example of the principles discussed above. The telepresence system fixes the distance from the participants to the screen, assuring that each participant is sitting at the ideal distance to maximize pixels, while not being too close to see the pixels. In addition, the telepresence system generally assures that each participant is visible from above a tabletop, essentially a torso view. In Figure X, a classic view of Cisco Telepresence shows how this appears. Notice the people in the video displays. We can see the three people in the middle with their hands clasped are engaged and intent, while the person on the far right is sitting back and observing. In all cases, this positioning would be lost without torso video. The ideal distance and pixel density enables a clear view of the other factors of eye movement and expression.

This ideal combination of full torso display and resolution has meant that telepresence is an ideal remote system for conveying the critical video cues required for business video, or more precisely, interactive "selling." This is why executives, boards, and others that require that visual feedback have adopted telepresence.
Devices Define the Experience

Turning to video on devices, not room systems, we can begin to see how these concepts can be applied to the devices and the potential video experience they can deliver. Note that all of the examples included here were documented using devices readily available.

What the Camera Sees

In order to evaluate the experience and value of different devices, it is important to understand the angles for both video cameras and displays. For displays, all of the examples use 40 degrees, the midpoint of the 30-60 degree range (see the Video Conferencing Display Size white paper for more info). For cameras, the range is different. Generally the cameras have angles tied to the dimensions of the screen, so they range from 40-60. Figure X shows actual photos taken from devices ranging from a desktop PC camera to an iPhone 4. Each picture was taken exactly 12 inches from the surface. By taking

![Microsoft Life Studio Camera](image1)
![Apple iPad 2 Camera](image2)
![HP TrueVision HD Laptop Camera](image3)
![Apple iPhone 4 Camera](image4)

*Figure 8 Relative camera picture widths for four devices at 12 inches*

a picture of a ruler, the camera angles for both the wide and narrow sides can be calculated. Figure X is a table based on these observed measurements that calculates the wide angle, the aspect ratio of the camera and the narrow angle. Using the simple trigonometry tangent function, using the observed
camera width at a 1-foot distance, the angle of the visible image can be calculated. By using the aspect ratio of the camera, the angle in the smaller side can be calculated. As can be seen, the maximum angle of this set of cameras is about 64 degrees. Generally, 60 degrees is quoted as the maximum in most systems, so that it is used for an analysis of the ideal camera/viewing distance. This table can be seen clearly in figure X. In this figure, the red area is a representative camera angle with a 60-degree image angle in the wide aspect ratio. The top view of a body is located 40 inches from the camera to an image location just behind the ears. In a later section this positioning will be explained. The three black lines show the relative angles for the devices, with the laptop and iPad shown together at about 55 degrees.

This diagram makes it clear that almost all cameras have about 60 degrees of angle in the wide image. This 60-degree angle will be used as a standard to analyze the impact of the camera.

<table>
<thead>
<tr>
<th>Device</th>
<th>Observed Width (inches)</th>
<th>Tangent</th>
<th>Wide Angle (degrees)</th>
<th>Aspect Ratio (W=1)</th>
<th>Narrow Angle (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft LifeStudio HD Camera</td>
<td>15</td>
<td>0.63</td>
<td>64.0</td>
<td>16x9</td>
<td>38.7</td>
</tr>
<tr>
<td>HP TrueVision HD Laptop Camera</td>
<td>9</td>
<td>0.38</td>
<td>41.1</td>
<td>4x3</td>
<td>23.8</td>
</tr>
<tr>
<td>Apple iPad 2 Camera</td>
<td>12.2</td>
<td>0.51</td>
<td>53.9</td>
<td>4x3</td>
<td>31.9</td>
</tr>
<tr>
<td>Apple iPhone 4 Camera</td>
<td>10</td>
<td>0.42</td>
<td>45.2</td>
<td>4x3</td>
<td>26.4</td>
</tr>
</tbody>
</table>

Figure 9 Calculated camera angles for four popular devices

Figure 10 camera angels for devices compared to 60 degree angle
The Visible You

As was discussed earlier, the key to the person viewing is to be able to see enough of the person being seen to feel comfortable with the image. Figure X shows a representative 6-foot tall figure with representative scale and the required visible field to see the arms and slightly (about 2 inches) above the top of the head. This required visible field is 26 inches tall and 28 inches wide (to accommodate some arm motion away from the body). This required visual field for viewer comfort will be used as the basis for analyzing all camera positions.

Figure X shows how this visible field translates into both portrait and landscape orientations, the top is a 16x9 (typical HD wide-screen camera) with a "landscape" orientation. The darker grey bands show the area captured by the camera aspect to maintain the minimum size on the smallest side. As can be seen, to generate the 26-inch tall image, a width of just over 46 inches is required. The second image shows the same for the camera and device in the "portrait" orientation. As can be seen, with a true 16x9 camera and display, the landscape orientation is superior. For some small devices that use 4x3 aspect ratios, the portrait may be better, but most new devices have the 16x9 standard. A 16x9 screen and camera, along with a landscape orientation, will be used for this analysis.
The 60/50 Camera/Screen Experience

Using the combination of a 60-degree camera angle and 50-degree viewing angle (both on the widest side of any aspect ratio), it is possible to calculate the ideal viewing distance and then from that the required screen size.

Figure X shows the representations of our video "conferencer" in the side and top views. These views include the 26-inch image height and 28-inch image width. These images and dimensions are based on an average 5'-11" inch male. As will be noted later, all analysis is linear to relative size of the individual as that defines the required camera field.

By placing the conferencer in the camera field, it is possible to see how the distance from the camera to the user can be defined. Figure X shows the conferencer in the visual field in both a top and side view. As can be seen, the vertical dimension drives the length of the field of view from the camera. With a 26-inch height and a 36-degree vertical angle...
(the vertical angle of a 60-degree side angle lens), the length to the 26-inch height is 40 inches (using the tangent function in trigonometry). This visual point is at the widest point, at the back of the shoulders. The dashed line shows the eyes, which are five inches closer to the camera.

Using this model for the ideal camera position for a 16x9 camera and our model figure, it is now possible to overlay the required screen size to maintain a 50-degree viewing arc in the wide aspect. Again, using simple trigonometry, it is possible to calculate that a 50-degree angle and a 35-inch viewing distance (the 40-inch image distance less the difference the eyes are forward). This calculates to a screen width of 32.7 inches. Figure X shows both the image relationship and the required screen size in a top view that shows both as horizontal images.

A 32.7 inch screen width translates to a diagonal screen measurement of 37.5 inches. As most monitors and televisions are sold with diagonal measurement, this allows the comparison to the ideal screen. A typical 25-inch desk monitor has only 44% of the viewable area of the 37.5 inch monitor shown. Figure X shows a more accurate side view of the vertical dimensions and layout of our ideal scenario. In this view, the camera is actually mounted to the top of the monitor, and the relative distances are shown.

The table in Figure X is a calculation for both orientations. It assumes the screen matches the camera.

---

<table>
<thead>
<tr>
<th></th>
<th>Camera Distance Calculations 60 Degree Focal Angle for Wide Aspect</th>
<th>Required Screen Size Calculations 50 Degree Visible Field for Wide Aspect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wide/Horizontal angle viewing (degrees)</td>
<td>Narrow/Vertical angle (degrees)</td>
</tr>
<tr>
<td>16x9 HD Camera and Display - wide angle is horizontal</td>
<td>60</td>
<td>36.0</td>
</tr>
<tr>
<td>16x9 HD Camera and Display - wide angle is vertical</td>
<td>36.0</td>
<td>60</td>
</tr>
</tbody>
</table>

---

Based on this analysis, we now have a calculation for the "ideal" desktop experience. However, very few of us have a 37-inch desktop monitor. How does the typical experience with a smaller device compare to the ideal?
The Actual Desktop Experience – 27-Inch Monitor

Figure X shows both a side and top view of a typical 5'11" office worker with a video system using a 27-inch diagonal desktop monitor. The left views are top and side views showing the viewing angle of the screen at the ideal camera distance base. The red lines show the 50 degree ideal viewing angle in comparison to the actual viewing angle that positioning for good camera viewing delivers. The right shows the ideal screen viewing distance. On the right, the position from the left for ideal camera is shown as a greyed image. The distance between the ideal camera and the ideal viewing is the gap. The user is challenged to position themselves to either be able to see the other participants or have the other participants see their full body. As can be seen here, with a 27-inch monitor it is possible to position close in the middle and get a reasonable image and screen view. As the screen gets smaller, the gap will increase.
The Actual Desktop Experience - 21 Inch Monitor

Moving down to a 21-inch monitor reduces the visible area by 40%. Figure X shows how this impacts both the viewing angle when positioned for a good camera view and the distance for good viewing. By reducing to the 21-inch monitor, the gap is now almost half the distance to the screen from the good camera position. If the user moves to see well, all arm and body views are gone.

Figure 19 21 inch desktop monitor experiences
The Laptop Experience

Moving from the desktop to a 15-inch laptop causes a major change. With the smaller laptop screen, the difference between the ideal camera and viewing distance is now more than double. The ability to position for both reasonable viewing and reasonable projection is almost lost. In fact, generally the user will position close to the viewing distance, resulting in a camera view that is essentially a head shot. This significantly reduces the experience of the other parties and significantly exacerbates the eye angle. If the user is looking at the screen and "talking" to the images instead of actually looking at the camera, this will result in a very poor expertise (experience?) for the receiver.

Figure 20 15.6 inch laptop monitor experiences
The Tablet Experience
Going down in size from the 15.6-inch laptop to a 10" tablet (the 7-inch versions are not considered here), the challenges get even worse. Again, the left views in figure X show how the user and the tablet positioned for the best camera angle and view for the other video participants. As can be clearly seen, the actual viewable area is now less than half of the eye’s visible field. Remember, this is a 50% reduction in visibility as it is a two-dimensional image. Obviously, the viewer’s capability to identify the details and body language responses in the incoming video will be seriously compromised. The result is that the user moves closer as shown on the right. Here again, the user has moved close enough to really see the image (and take advantage of that retinal display). However, now the camera image is significantly compromised. The arms are only visible if they are moved up. Note that this is shown with the device in the landscape view; in the portrait view, the actual image is larger, but the screen is now positioned wrong for viewing groups.

Another factor that is somewhat visible in the laptop scenario becomes very visible in the tablet. As the device is typically on a desk and the user’s head is now 6-12 inches above the device, the user now appears to be "hulking" over the camera and peering down into the image. This becomes quite disconcerting to the person watching the video. Again, Figure X shows two actual shots at the two distances using an iPad 2.

Finally, the eye alignment issue is even worse, as the angle is greater between the center of the screen and the camera.
The Smartphone Experience

Reducing the display further to the 4-inch maximum for a typical smartphone exacerbates the situation further. Again, Figure Z shows the smartphone positioned for a reasonable camera image in the left views. In this case, two things become clear in this view. The first is that the phone must be on a stand as the users arms are not long enough to hold it (plus this eliminates many forms of body language), and second, the viewing angle is very small.

Holding the phone close enough to have the screen be large is shown on the right. Assuming the user can focus at this distance, the screen is now highly visible, but the camera image is virtually useless. In landscape, all that can be seen is part of the face, and in portrait the whole face is visible. Another issue with this is also visible, which is the device tends to drop to a low angle, creating the "nose hair" video view. This is an extension of the issue of "hulking" and comes as a hand-held device will not be held up at head height as the arms get tired.

Figure 22 4 inch smartphone monitor experiences
The Perception Factor

The analysis above can be reduced to a Perception curve based on the device and the screen size. Ignoring variances in camera angle and assuming they are all 60 degrees, then it is possible to map the relative screen area. As a reminder, Figure X shows the relative views and can be used to guide the analysis. Figure X is a map of the perception factor with suggested ranges on it for selection of a monitor size for individual business video.

At the left, the 37.5-inch monitor that yields a 50-degree viewing angle when the user is completely inside the 60-degree viewing field is the ideal. As the screen size shrinks in diagonal measurement, the visible area decreases at the square of the reduction in size. Based on the analysis, the devices from 27 to 37.5 inches will allow a reasonable compromise between viewing and camera. These devices are required for symmetrical video conferencing where it is important for both parties to have a complete experience. From the 15.6 of a laptop to the 27-inch desktop, the experience will be significantly compromised. This should be avoided unless economically impossible. It also can be used for an asymmetrical experience where one party does not require feedback, but merely needs to see that the other party is there. The caution is that such asymmetrical events are few. For example, in a video kiosk, the agent needs to be able to see the customer to sense their mood and the customer will feel much more comfortable with a torso view of the agent. Finally, as has been clearly demonstrated, devices under 15.6-inch diagonal screens are virtually useless for business video. While these devices may be great for video with family and friends, the video experience will make their use limited at best.
Wide Angle Options

One potential change is to make the front-facing camera in a device wider. While the desktop camera, the angle is already out to 64 degrees, wide, this is an option. However, distortion becomes an issue as lenses get wider, as does managing a focal field across this image. The view shown in Figure X (from Wikipedia) is a typical 28-mm camera lenses and has a 65-degree field of view. As the angle of the lenses is widened, the potential for distortion becomes higher. Figure X (also Wikipedia) shows three lenses and the distortion a wider lens creates. As can be seen, the lenses under 25 mm, the greater the 70 degrees of field begin to distort the image.

This becomes a significant challenge, especially to maintain price points for the cameras. Therefore, until there are major new developments, the 60- to 70-degree lenses are the largest we will see.

Conclusions

The point of this paper was to discuss whether small devices could really be used for a good video experience for both the person sending and receiving the video. As the analysis has shown, as devices shrink, the ability of the camera and device viewing angles to match up to enable both a good transmitted and visible image becomes difficult. In fact, many devices probably will not create a video experience that will be useable for advanced selling collaboration where the visual clues are essential.

As devices improve and cameras become more tuned to video, it is clear that enhanced lenses and field-of-view may reduce this issue. When buying a device or system, a user should see how the device performs using the analysis above. Before buying any device for video conferencing, experiment to find the ideal camera and viewing angles. Do an actual video and position yourself so the camera gives the "view of you" as discussed earlier, from the lower torso to two inches above the head. After doing this, put a document on the screen and try to read it from that distance. Then use the distance number to position yourself the right distance from the screen for optimal viewing, and see what the difference is. If you are more than 20-40 percent closer, then the device will not be good for video.