



February 3, 2025

SUBMITTED ELECTRONICALLY

Chairman Brendan Carr
Federal Communications Commission
45 L Street, NE
Washington, DC 20554

RE: Joint Comments on FCC's Access to Video Conferencing Proposed Rule (CG Docket Nos. 23-161, 10-213, and 3-123)

Dear Chairman Carr:

The American Council of the Blind (ACB), the American Foundation for the Blind (AFB), and Perkins School for the Blind (Perkins) appreciate the opportunity to provide joint comments on the Federal Communications Commission's (FCC) proposed rule on access to interoperable video conferencing services (IVCS). Our organizations collectively represent communities and individuals who are blind, deafblind, or have low vision, and we are united in our missions to ensure that people have equal opportunity to IVCS. Founded in 1961 and comprised of thousands of members and 66 state and special-interest affiliate organizations, the ACB is a national member-driven organization of and for individuals who are blind, low vision, and deafblind. AFB is a national nonprofit that creates equal opportunities and expands possibilities for people who are blind, have low vision, and are deafblind through advocacy, thought leadership, and strategic partnerships. Perkins is the worldwide leader in education services for children and young adults who are blind, deafblind, or visually impaired. Perkins also partners with organizations to help them create digital products, services, and experiences that engage and include all people, regardless of their abilities.

We commend the FCC for its continued work to enhance accessibility in video conferencing platforms for individuals with disabilities. Video conferencing is an essential communication technology that enables individuals to participate in remote and hybrid professional, educational, and personal arrangements. Accessibility of video conferencing platforms is not consistent, and while many platforms have accessibility features, others are difficult to navigate for individuals who are blind, deafblind, or have low vision.

The FCC seeks feedback on the proposed performance objectives and if those objectives sufficiently address the accessibility challenges experienced by persons with disabilities. Our perspective is that the current measures are not adequate to meet the needs of individuals who utilize assistive devices and technologies to fully participate in video conferencing. Persistent issues include limited compatibility of video conferencing platforms and the full suite of video conferencing features with screen readers, Braille displays, and other assistive technologies as well as user interfaces that remain difficult to

navigate for individuals with visual impairments. Amending the proposed performance objectives is an opportunity to advance the goal of accessibility by ensuring assistive technology integration is guaranteed across all video conferencing platforms. To that end, we respectfully offer the following comments on the proposed objectives.

1. Performance Objectives as Applied to the Blind, Low-Vision, and Deafblind Community

a. Broad Application and Compatibility with Assistive Technology

We urge the FCC to apply performance objectives broadly across all subcategories rather than a fragmented application of standards that risks perpetuating accessibility challenges for individuals with disabilities encountering varying levels of accessibility depending on the platform used. Furthermore, we recommend the FCC mandate compatibility with the full array of assistive technologies to address the current challenges of incompatibility with assistive technology, such as Braille displays for the deafblind community, customization options, and screen readers. As discussed in more detail below, we also urge the FCC to expand this coverage to include key functionalities including, but not limited to, using braille for reading captioning text for deafblind individuals. It is essential to ensure integration with existing assistive technology and pave the way for a more accessible future. Accessible technology does not demand excessive costs or compromise the functionality or design of the platform.

b. Visual Information in Both Auditory and Tactile Forms

Furthermore, the FCC should split its performance objectives at 47 C.F.R. § 14.21(b)(1)(i). One performance objective should read: “Provide auditory modes that do not require user vision.” A separate performance objective should read: “Provide tactile modes that do not require user vision or hearing.” In addition, we support the modification of the performance objectives at 47 C.F.R. § 14.21(b)(2)(i) to read: “Provide visual information in *both auditory and tactile forms*.” This would enhance accessibility for people who are deafblind or who otherwise require that controls and information be accessed tactilely. For example, tactile functionality options covering braille displays would allow deafblind persons to use braille to read captioning text. Braille enables people to access information without relying solely on auditory means, which is crucial for learning and independent living. Devices like refreshable braille displays—electro-mechanic devices that connect to computers and tablets to deliver braille outputs—have enhanced accessibility for braille readers. In addition, we believe that the screenshare feature is an area where tactile options could be impactful. We applaud the FCC for acknowledging the need to provide visual information in tactile forms.

2. Terminology in the Context of IVCS

The FCC requests feedback on the distinction between the terms “audio description” and “visual image description” in the context of IVCS. We support these definitions, as set forth in the proposed rule, and believe that a clear distinction between these terms is needed to avoid conflating accessibility concepts.

The term “audio description” refers to a feature that is required for television and other video programming pursuant to the FCC’s part 79 Rules. Under those rules, audio description of a program’s key visual elements must be inserted into the natural pauses in the program dialogue. This term is widely recognized in the blind and low vision community.

According to the FCC, the term “visual image description” refers to a related feature, described by a commenter as functionality that generates real-time descriptions of visual information for people who are blind or low vision. As a preliminary matter, we wish to take this opportunity to clarify that this term includes “visual interpreters.” The term “visual interpreters” is more commonly used by the blind or low vision community when referring to the individual providing the visual image description service from a third-party provider on a video conferencing platform. This individual is arranged for by the end user or their representative with the third-party provider and is not the responsibility of the video conferencing service. Examples of visual interpreters includes Be My Eyes or AIRA.

3. Provision of Audio Description and Visual Image Description Services

The FCC requested feedback on amending the performance objective relating to the availability of visual information for people who are blind or deafblind or have low vision to specify the provision of audio description and visual image descriptive functionality, as well as compatibility with third-party visual image descriptive services. We support amending the performance objective to include requiring audio and visual image description and compatibility with third-party services. Establishing clear guidelines will help ensure that industry stakeholders consider the full utilization of emerging technologies as potential accessibility solutions. These guidelines must also accommodate future advancements in assistive technology to maintain long-term compatibility with IVCS.

The FCC is also seeking input on the scope of visual information that should be provided through audio description in IVCS, including the potential requirement for platforms to support both built-in audio description and compatibility with third-party services. Many individuals who are blind or have low vision utilize third-party visual interpreting platforms such as AIRA and Be My Eyes for real-time visual image descriptions. These services connect users with volunteers and trained professionals who provide customized descriptions of on-screen content. Compatibility of videoconferencing services with third-party visual interpreting platforms is critical for end-user accessibility of content on platforms.

4. Separate Description Channel

To maintain the integrity of the main audio channel of a videoconferencing session, the FCC should require videoconferencing systems to provide a separate channel where audio descriptions or visual image descriptive services may be inserted, and where descriptions may be listened to discretely from the main audio. This approach would maintain consistency with how descriptions are currently delivered to television and movie audiences. To support such a feature, an application programming interface (API) would need to be developed allowing hosts to assign an identified audio describer or visual interpreter to a videoconferencing session. The user interface should also provide a mechanism for the host to designate any participant in the session as the audio describer or visual interpreter and grant them access to the discrete audio channel.

Once assigned, the describer/interpreter would provide descriptions into a discrete audio channel, and these descriptions would be transmitted to participants separately from the main audio channel. To hear descriptions, IVCS vendors would provide an option (such as an "AD" button or menu choice) in the user interface that would activate the audio-description channel, in the same way that pressing the CC button in current IVCS interfaces displays captions. Activating the channel must not deactivate the main audio channel; rather, both channels should play in parallel, allowing each participant to control their own accessibility settings without affecting others. The audio channel should also be customizable by users in different ways, such as the ability to choose to duck the main meeting audio by a set percentage level,

duck only when the participant is speaking, and choose what channels the main and audio description channels are heard. Additionally, if a host elects to record a session, the audio-description channel must be included in that recording to preserve accessibility. The discrete audio channel should be customizable, allowing for a user to adjust volume.

5. Accessibility Preferences

It is critically important that IVCS users' accessibility preferences are stored and retained within the IVCS platform, so users are not forced to change the settings each time they use the service. Specifically, the user's preferred setting should be tied to the IVCS. Users should have the flexibility to customize the IVCS platform for the duration of their choice. Users should not have to reselect the settings each and every time the user wishes to use the service. A person with a disability often chooses multiple settings in an IVCS for a variety of reasons to accommodate their disability. For instance, a person who uses screen reading software may choose to turn off the setting that reads chat comments aloud. This is often done because it is difficult to simultaneously listen to the verbal conversation and chat conversation at the same time. Resetting this preference every time a person uses the IVCS is burdensome and time consuming.

We recognize that preferences may be set by users through a general settings page (for instance, a general settings menu on a smartphone) or the individual IVCS platform's setting. If preference generalizations can easily be made across IVCS platforms, users should have this option. However, users should also be able to set their preferences on each individual platform. Different platforms may offer different options in communicating and sharing information. As a result, general setting options may not fully encompass all the different feature possibilities offered by one platform. The final rule should account for this specificity and future innovation. In other words, a general settings page on a smart phone is helpful, but it should not limit individualized settings needs for new platforms that go farther than other IVCS.

We also recognize that a user's preference may differ from device-to-device. Although a person may use the same IVCS across multiple devices (e.g., a computer or a smartphone), the person may tailor their preferences on the devices based on different circumstances. This may be due to the device's size or shape or the task that is being performed on the particular device. As a result, the accommodation needs may look different from one device to another. Further, the same settings for one account may not apply to all devices. We believe that such tailoring device-to-device should be recognized and accounted for in the final rule. Settings should also be tied to the particular device used.

Finally, the process to select certain accessibility options should be quick and easy to find (including the use of shortcuts or gestures)—with respect to both the IVCS settings, as well as the device's settings. A user may need to quickly change settings for a variety of reasons on a moment's notice. The accessibility settings should not be hidden or difficult to find, time consuming to scroll through, or complicated to understand.

6. Accessibility Features

a. Chat Function

People who are blind, who are deafblind, and who have low vision often use screen reading software. The screen reader reads aloud anything that is displayed on the screen when the text appears. As a

result, the screen reading software reads aloud any discussions that take place in the chat box. However, if a participant communicates in the chat box at the same time a participant is verbally talking, the screen reader voices the text at the same time the person is audibly talking. This makes it difficult for the user to follow both voices at once.

This problem can be mitigated by developing different settings to lessen the frequency of such overlap. For instance, a setting could be included that truncates the written messages that are read aloud. A screen reader user should have the ability to choose how often they receive the messages verbally or turn off the comments altogether. The user should also be able to turn the chat setting back on quickly if necessary. Additionally, if the person chooses not to hear the message every time, to alleviate overlap, the messages should be easy to find again when the person has the time to backtrack and read the chat conversation. The ability to move seamlessly between the chat box and the main screen should be developed to allow users to fully participate in the conversation and access both the verbal and written content of the discussion.

People who are deafblind also often use screen reading software that offers braille reading on their devices. The same situations described above are experienced by deafblind users that utilize screen reading software, and the FCC's final rule should also address accessibility for this population.

b. Screen-Reader Verbosity Control

While it is imperative to have access to all of the outputs, controls, inputs, and other functions of IVCS, it can be difficult to understand and interact with multiple audio and/or tactile outputs at the same time. For example, if video conferencing participants are speaking while simultaneously entering information in the text chat, it can be difficult to follow the speech and captions if a screen reader automatically voices the text in chat. To maximize accessibility, it is important to be able to control and customize when and how information is being provided to the user. Blind, low vision, and deafblind IVCS users would benefit from settings both within the IVCS platform and in third-party assistive technologies to customize the verbosity of information and to control how and when visual information is provided audibly or tactilely to maximize their full participation in the video conference.

Several assistive technology tools and video conferencing platforms currently provide some of these features. For example, Zoom allows users to enable or disable alerts when a chat is received, a participant has joined the meeting, or a participant has raised their hand. In addition, an NVDA add-on allows users to activate many of the alert customizations settings that Zoom offers via a keyboard shortcut. JAWS allows users to toggle the screen reader on and off with a keyboard shortcut to control how much information is being announced. It is worth noting that toggling between speech on demand and full speech is helpful for improving concentration, but solely switching to on-demand speech can also reduce a user's access to visual information in the meeting.

We recommend that the FCC require that both IVCS and third-party screen readers allow users to customize the voicing of alerts. That means both the ability to turn alerts on and off and to control how much information is automatically read aloud in the meeting, such as by truncating the reading of long chat messages or URLs. Users should be able to control when and how that information is voiced or displayed to avoid conflicts between spoken audio or captions in the video conference and the provision of visual information by the screen reader or braille display.

c. Shared Documents

Several IVCS, like Zoom, Microsoft Teams, Slack, and Google Meet, allow users to share their screen or documents in video conferences. Typically, screen sharing renders the document or screen as nothing more than a flat image or video with no information that screen readers can interpret. This means that blind, deafblind, or low vision users cannot access the visual information being shared, which puts them at a significant disadvantage compared to other users, especially in an employment context. There are, however, some platforms that have begun to deploy features that make the shared document interactive for other users. For example, if a user properly shares a PowerPoint presentation in Microsoft Teams, viewers can navigate slides independently of the user sharing their screen. This PowerPoint Live feature allows viewers to change the slide's visual contrast and magnification and use a screen reader to get access to the slide content. Such features demonstrate that it is possible to make screen sharing accessible for blind, low vision, or deafblind users, at least with certain file types. To make up for gaps in accessibility, some users rely on third-party for access to shared documents. For example, Pneuma Solutions offers a product called Scribe for Meetings that makes Zoom or Microsoft Teams screen shares accessible to participants using screen readers or magnification software.

The FCC should add a requirement that IVCS must provide document sharing features that allow blind, low vision, and deafblind users to interact with the shared documents, both by customizing the visual display (e.g., magnification and contrast) and by making the content detectable to and readable by screen readers and braille displays. In addition, IVCS platforms should allow integration of third-party services that can make these shared documents accessible to meeting participants.

d. Collaborative White Board Feature

Similarly, many IVCS include a feature that allows users to simultaneously type text and draw on a virtual "whiteboard." These whiteboards can be collaborative, but they are not generally accessible to screen reader users. While some of the whiteboard features do not easily lend themselves to nonvisual accessibility (e.g., drawings), it should be possible to make text accessible to screen readers and create some sort of structure, so that screen reader users understand which information that users add is connected to other information. For example, a screen reader user would need to be able to understand whether the text in focus is associated with a header, such as "Ideas" or "Action Items." In addition to understanding visual information from other users, it is also important that screen reader users and other users who do not use a mouse are able to input information to these whiteboards as well. We recommend that the FCC include collaborative White Board features when specifying the types of information that must be made accessible to users with disabilities.

e. Pre-Determined Reactions

Some IVCS allow users to display "reactions" that visually communicate an emotion or immediate response to other users. For example, Google Meet allows users to select emojis, such as a heart, thumbs up, or hand clap, that floats up the screen to express that emotion or reaction without coming off mute or interrupting a speaker. These reactions are not always identifiable or announced by screen readers, so blind, low vision, deafblind users tend to miss out on these reactions. That can be especially problematic if a blind, deafblind, or low vision user is being congratulated or responded to with one of these reactions. In other cases, these pre-determined reactions are used to vote in a verbal poll, and users do not have access to the outcomes of the vote or poll like sighted users do. We recommend that the FCC include these pre-determined reactions among the features that must be made accessible to

blind, low vision, and deafblind users. However, per our discussion about verbosity and chat features, it is important to customize how and when these alerts are identifiable or announced via their screen reader.

7. Artificial Intelligence (AI) and Audio and Visual Image Descriptions

Additionally, we urge the FCC to consider that advancements in AI might allow for future automated audio and visual image description features, such as machine-generated descriptions in videoconferencing sessions. However, we believe that it is crucial to balance innovation without implying a preference for AI descriptions over human-provided descriptions. The ability for an IVCS to support automatically generated descriptions should never preclude a host's ability to incorporate live, human-generated audio descriptions, if the host chooses to do so.

Rules for incorporating automatic descriptions into the main audio channel would need to follow long-established conventions for human-generated descriptions. First and foremost, the descriptions must be accurate and must not introduce falsehoods or misrepresent visual details. Recognizing that automatic descriptions in an IVCS session would be generated in real time and without the benefit of human review, accuracy of these descriptions will be paramount. We, therefore, encourage the FCC to set guidelines for the accuracy of automatically generated audio descriptions, similar to caption quality standards that are established at 47 C.F.R. § 79.1(j)(2), *Caption Quality Standards*.

It is of equal importance that AI-generated descriptions should avoid competing with the main audio. Conciseness is also a crucial aspect of human-written descriptions, and AI-generated descriptions must not be overly verbose or give listeners more information than is necessary. AI systems for generating descriptions would need to be trained to create accurate, succinct, and precise descriptions to convey important visual details relative to a given context, omitting unnecessary details to maintain clarity and reduce aural clutter.

We commend the FCC for its leadership in addressing accessibility in video conferencing and appreciate the opportunity to provide comments. Should you have any questions or require further information, please contact Claire Stanley (cstanley@acb.org) at ACB, Sarah Malaier (smalaier@afb.org) at AFB, and Kim Charlson (Kim.Charlson@perkins.org) at Perkins.

Sincerely,



Sarah Malaier
Senior Advisor, Public Policy and Research
American Foundation for the Blind



Dave Power
President and Chief Executive Officer
Perkins School for the Blind



Kim Charlson, Executive Director
Perkins Braille & Talking Book Library
Perkins School for the Blind



Claire Stanley, JD
Director of Advocacy and Governmental Affairs
American Council of the Blind