Television & New Media
2016, Vol. 17(1) 21–43

DOI: 10.1177/1527476415577211
tvn.sagepub.com
SAGE

Scenes from an Imaginary Country: Test Images and the American Color Television Standard

Dylan Mulvin[1] and Jonathan Sterne[1]

Abstract

American analog color television—so-called NTSC color—is likely the most pervasive image standard of the twentieth century, yet it is infamous for its technical shortcomings. Through a history and analysis of the National Television System Committee (NTSC) standard, this article argues that the political presuppositions of engineers shaped the representational capacities of television for nearly sixty years. In particular, the test images used to develop a perceptually satisfying image evince assumptions of a leisurely and white United States as the "normal" subject matter of television. In this way, test materials coalesce abstract assumptions about the normal and the exceptional at the level of both form and content. This article concludes that NTSC color served as a model for how normed cultural sensibilities about image quality, perceptual ability, and the representational imaginary have been built into subsequent technical standards. A Scalar version of this paper, with more pictures, is available at http://colortvstandards.net

Keywords

visual culture, perception, test images, media standards, whiteness, norms, the United States, television, technology, media infrastructure

> . . . Colour television systems do *not* directly register the world; a whole technology intervenes.
>
> —Brian Winston (1996, 41)

[1]McGill University, Montreal, Quebec, Canada

Corresponding Author:
Dylan Mulvin, Department of Art History and Communication Studies, McGill University, 853 Sherbrooke St. W., Montreal, Quebec, Canada H3A 0G5.
Email: dylan.mulvin@mail.mcgill.ca

Figure 1. NTSC color television test slides (Fink and NTSC 1955).

A young woman clutches a kitten, two others play table tennis, some boys canoe, and some play tug-of-war. There are pumpkins, fishing tackle, a lampshade, a tulip garden, a houseplant, and a southern manse. These are just a few excerpts from the imaginary America that color television engineers and the Eastman Kodak Company compiled in 1951. These excerpts are all taken from the twenty-seven test slides (Figure 1) that were an integral part of the testing regime established by the National Television System Committee (NTSC) for its 1953 color standard—which became the most ubiquitous moving image standard of the twentieth century, when it was chosen as the standard for American color television.

Considering the relatively short half-life of communication standards, the NTSC standard enjoyed considerable longevity, lasting for fifty-six years in the United States, until June 12, 2009, when full digital conversion took place. NTSC is still used in some countries and remains a color calibration option in many computer operating systems. Although it was chosen on the basis of a series of local concerns and peculiar aesthetic criteria, the NTSC standard has tuned the hue of a great deal of televisual culture over the last half-century, and through a host of remediations—video tape, CRT screens, YouTube, JPEG, MPEG—a great deal of media culture besides.

In this article, we consider the test slides used to develop the NTSC color standard as well as the contexts in which they operated: the testing scenario itself, the technical infrastructure of monochrome television into which color television was to be inserted, the cultural values circulating in the tests and through the images themselves, and their subsequent media legacy. Test objects are a necessary part of a media standardization process. Test images, as a kind of test object, act as prototypical proxies for how a viable visual medium will handle its materials of transmission. Although standards strive to hide the traces of their testing regimes, the many shortcomings of the NTSC standard point us back to these documentary artifacts for clues about the affordances and failures of the imaginations that brought it into existence. We argue that the NTSC predicted how its standard would work by projecting two components of the future of American color television onto its testing regime. In the form of its test subjects it imagined the future viewers of TV; in the form of its test materials—its test slides and a single filmstrip—it imagined both the future subject matter and the formal characteristics of American color TV content. The content of the test materials, however, evinces a concern for the system's capacity to reproduce formally complex images and, at the same time, give primacy to a circumscribed vision of American television as a domain of whiteness, leisure, and pastoral ease. Together, the test subjects and test materials negotiated the gaps between the universal and the particular. By making a choice about what, in particular, could be reproduced and tested (creating a fixed set of materials and testers), the engineers speculated about the realm of all possible TV images. As we conclude, the NTSC testing regime is illustrative of the ways all media aesthetics work and the ways all media formats contain wagers on their own representational potential. We thus follow a tradition of understanding the very form of technologies as crystallizing representations and social relations at the moments they come into existence (Spigel 1992; Williams 1974).

We read our history through the few artifacts that remain. We only have remediated versions of the test images themselves: four in color, twenty-three in grayscale, printed in a book from 1955, scanned, digitized, and compressed again—their original color characteristics lost to history. We also have detailed accounts of the testing process, gleaned from meeting minutes, the edited collection of papers provided by the committee, and a handful of articles. We offer this history as a complement to existing histories of color TV's reception and governance (Abramson 2003; Boddy 1990; Fickers 2010; Kane 2014; Murray, forthcoming; Spigel 1992) and as a contribution to the project of understanding the cultural practices that go into making media standards (though our empirical material is limited to the United States). Existing accounts of the early years of television describe the trials and provisional successes of early color television demonstrations and the adoption of TV in the home. But little attention is paid to the cultural dimensions of the standards and infrastructures that subtended the television system. Within the practical domains of television engineering lay a politics of how television would work and what it would show. We argue that these engineering decisions shaped the look and feel of television for over half a century. Hence, our focus is on how the NTSC's test images and testing regime enfolded the representational prospects of a new color broadcast technology as well as the perceptual capacities of potential viewers. As

we argue in the conclusion, much of the epistemology behind the NTSC tests continues to shape new image standards. Although analog color television recedes into memory in many parts of the world, its legacy lives on in digital image and video standards still in wide use.

Media standards sometimes show their seams. When we adjust our television to a set of color bars or a projector to an image of a diagonal axis, we calibrate our communication media to the specifications of a standard. When we do this, we also conform to an aesthetic based on the extensive testing and research that went into creating that standard. Sometimes the seams appear elsewhere, as in the unforeseen byproducts and failures of representation endemic in compressed media standards. These moments of "infrastructural inversion" can uncover the epistemic and aesthetic politics of the engineering imaginary (Bowker and Star 1999). In this way, the NTSC color standard is a very fruitful subject. Not only is it the most ubiquitous moving image standard of the twentieth century, but it is also infamous for its failure to faithfully reproduce skin tone: *No True Skin Color* and *Never The Same Color* were some of the system's more common nicknames. This is especially striking because earlier color TV systems were rejected because the "inability to accurately reproduce skin tones is a particularly serious handicap" (Fink and NTSC 1955, 17). The NTSC standard held a functional aesthetic desire—represent white skin tone—in dynamic tension with the technical limits of an infrastructure. At the same time, it was an ongoing monument to the failures of its designers to fully predict its use and reception (Murray, forthcoming).

Although color reproduction was a goal for television from its very beginnings in the nineteenth century and color standardization was always part of TV standards (Kane 2014, 56–57), American industry histories often frame color as an afterthought, because the color standard came after the monochrome standard. The usual story goes like this: the first NTSC negotiated a set of industrial compromises to produce an acceptable monochrome standard for American television, which the Federal Communications Commission (FCC) adopted in 1941 (Boddy 1990, 34; Seldes 1956). A second NTSC convened in 1950 to develop a color standard that would work within the parameters of the existing standard—meaning that existing set owners would still receive the monochrome signal, and new color set owners could receive both signals. Therefore, the second NTSC had to figure out a way to cram a new tricolor signal into the bandwidth that had previously just transmitted a monochrome signal (Fink and NTSC 1955). Standardization implies "the necessity of relations" among institutions, technical protocols, practices, and people (Fuller 2005, 96). If the relations between technical protocols and institutions had to be relatively undisturbed, the relation to people would have to change.

The lever for that change was applied psychophysics. The NTSC built the limits of color perception (as they were then understood) into its color TV standard, rendering them economically useful by treating the presumed perceptual limits of the normal human eye as an exploitable efficiency in the system. In today's language of bandwidth, the NTSC system could transmit less information if it knew which information was less likely to be seen by audiences (Sterne and Mulvin 2014). But the committee's perceptual judgments were made in the artificial context of testing and built around

specific aesthetic materials: Eastman Kodak's twenty-seven color images. Andreas Fickers (2010, 99) has described the psychophysical research behind the European color standard as rooted in "the *habitus* of the engineers and technicians, who believed in the rationality of their profession and neutrality of their behavior as men of science and technology." We take Fickers's provocation seriously for the American context: what values and worldviews were bound up in the *habitus* of the NTSC engineers?

We should also note that the core value that subtends the psychophysical work—the separation of transmission from content—is drawn from information theory, a specific conceptualization of communication (Fink 1951; Shannon and Weaver 1949). This split is as much a political choice as an aesthetic or practical one. It is also immensely powerful as an engineering practice. Even while we challenge the separation of content and medium, we find ourselves forced into using the distinction because it pervades the NTSC engineers' own thinking. With a plea to read us at least a little ironically, we retain the distinction for the sake of legibility and a good story. When looking at the NTSC's test slides and their strictly limited representation of the United States, in addition to the homogeneous makeup of the test subjects, we see that contextual biases around whiteness and around the commercial purposes of television are baked into its media standards. The standard's technical shortcomings are clearly connected not only to the committee's inability to produce a truly functional representational norm but also to the broader dimensions of the tests.

The Greening of America

When the first gathering of the NTSC met to create an American television standard for monochrome broadcasting, their work took nine months. The color standard, by comparison, took thirty-two (Fink and NTSC 1955). The undertaking involved in developing a color standard was larger because the engineers knew they needed a system that was compatible with the existing monochrome standard. By the early 1950s, Americans already owned thirty million television sets and (in a move that might seem counterintuitive now) *not* forcing Americans to buy a new television or an adaptor was a requirement of any potential standard (Seldes 1956). The majority of the testing for color work took place between 1951 and 1953 across laboratories in the United States.

The NTSC engineers completed preliminary, subjective tests on themselves by the summer of 1951. For these initial tests, they used a suite of existing test images and test patterns to estimate the likely point of diminishing returns of bandwidth use: that is, they roughly determined the minimum amount of bandwidth that was necessary to produce a pleasing image by testing their own responses to basic shapes and colors. Test patterns are basic and elementary images. To know whether a visual technology is working, engineers use them to reduce the signal to a series of color blocks and gradients, the building blocks of more complex depictions. Test patterns, like the Society of Motion Picture and Television Engineers' familiar color bars (Figure 2), are usually used as calibration tools for downstream technologies; once a television signal is honed and broadcast and once a television is manufactured, sold, and setup in a

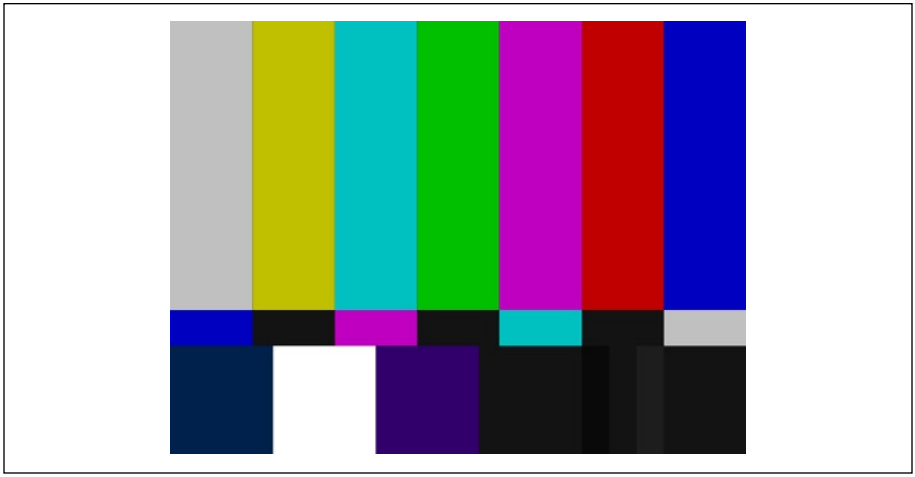

Figure 2. SMPTE Color Bars.

home, test patterns are the final level of quality control, as they can be used to verify that the saturation, hue, and contrast are (subjectively) accurate and/or pleasing. While the simple test patterns served to estimate the general parameters of a pleasing television signal in their initial tests, the NTSC needed more versatile and varied testing materials to more precisely determine the aesthetic requirements of a national television signal. As actual TV shows are rarely composed of stationary blocks of color, the NTSC agreed in August 1951 that Eastman Kodak would produce images for the tests and distribute them to the member firms: "some being 'average' pictures and some being 'exceptional' pictures" (NTSC 1951–1953a, 3). In total, Eastman Kodak produced 68 sets of the slides for the various firms involved, and for the FCC's records (Fink and NTSC 1955, 65).

Test images can portray anything, but once they portray *something*, that something becomes a shared set of fixed points across all testing environments. Because of the longevity of color TV, the choice of using new, unique images, and the peculiarity of the images themselves, the process of picking these fixed points becomes a rich subject for inquiry. Aside from their reproduction in a compiled set of the committee's reports—and their re-reproduction here—these slides are lost to time. However, we do know a few things about them. We know that Eastman Kodak produced the slides for free—one of the company's managers served on the NTSC testing committee—and that the committee's chairman, Alfred Goldsmith from RCA, "laid before the Eastman Kodak Company suggestions as to the general nature of the subject matter which should be included in a set of color-television test slides" (NTSC 1951–1953b, 6). And we know that partway into the testing regime, "three additional slides stressing subject matter predominantly green were added" (NTSC 1951–1953b, 6). As a new monochrome signal would be based off of the level of green in every frame of the color signal, testing green images was of primary significance.

Test images are the crash test dummies of a visual media standard. Like other forensic media, they are used to bring a new technology's failures into relief. They test the formal limits, potential catastrophes, and likely outcomes of compression, transmission, and perception (Siegel 2014). The NTSC engineers used the Eastman Kodak slides in two ways: to portray and test the "general subject matter" of color television and to test the boundaries of the "average" and "exceptional" colorimetric and formal information of the signal. These technical and representational layers were balanced against other concerns like perceptual clarity for the viewer, the economics of bandwidth use, and the feasibility of reproducing dynamic images. The NTSC tested its possible standards in dark rooms, with dozens of test subjects, trying to determine the lowest amount of information required to produce a pleasurable image. For each of these experiments, the Eastman Kodak test slides were the common denominator, serving as the yardstick for judging all the other variables.

Conspicuously, however, the test images for the most ubiquitous moving image standard of the twentieth century *didn't move*. The NTSC tests predated videotape recording by several years (Remley 1999), and the test film that Eastman Kodak produced and screened a single time cost more than $150,000. This meant that color slides were the only practical, affordable, and reliable means of testing the television signal with anything approaching consistency across lab settings. The NTSC clearly realized the paradox of basing their image standard on still images, and in an aside, the committee describes how actors performing live scenes would have been ideal, though impractical: "Manifestly, live-talent performances would in general be too elaborate, costly, and difficult of exact duplication to be suitable for routine laboratory or field tests . . ." (Fink and NTSC 1955, 64). Other than the use of rudimentary kinescope recordings to time-shift programming from the east coast, television in the early 1950s was a live medium. Although the idea of actors reenacting the same scenes thousands of times for the duration of the committee's tests may sound nightmarish, it is a nightmare that would have actually reflected the production methods of television at the time. But because live action was not feasible, and videotape still a few years away, the NTSC's standard for the moving television image is based almost entirely on twenty-seven slides.

These images confront us with one of the truisms of media studies: that the content of a new medium is an old medium (McLuhan 1964). The standardization of one media format was explicitly calibrated to a set of decisions made for a previous media format. By referring back to Kodak's color process, the NTSC engineers put their work in direct dialogue with a history of media aesthetics that was more about the very possibilities of representation—which colors could and could not appear and under what conditions. Their referent was not color in nature but color in other media. Kodachrome film stock was both an industry standard and a commercial touchstone for Americans in the 1950s, and its undergirding of television's moving images was chromatically consistent with a vision of the world already processed through it. Color photographs that appeared in *National Geographic* used Kodachrome, which was originally sold to Americans through the uncredited advertising photographs of Ansel Adams (Garner 2007).

The slides the NTSC initially used appear to reflect a *LIFE Magazine* or *National Geographic* photographer's vision of American life (Garner 2007). But there is a reason for this. Although we were unable to clarify the origins of the images themselves (for instance, if they were specifically made for the test or selected from an image bank at Kodak based on their suitability for testing), we do know that the images had an afterlife, appearing as test images in visual perception tests in the 1950s (Sziklai 1956). Therefore, we can read the test images anachronistically as examples of "stock photography," though that industry did not really cohere until the 1970s (Frosh 2003, 35, 38). Paul Frosh outlines several characteristics of stock photography that these images clearly share. They are conservative in what they depict, which is to say they work within stylistic idioms already believed to be commercially viable (Frosh 2003, 59). The images are selected based on their "apparent social significance" to advertisers and consumers "as interpreted by professional cultural intermediaries" (Frosh 2003, 69)—in this case, Kodak employees and television engineers. Their meaning is polysemic and generic, contextually determined both in relation to other images and without enough specificity to connote *a particular* brand, story, or set of objects or people without other contextual information (Frosh 2003, 72, 77). In other words, the images' relatively free floating status made them ideal test subjects for television: not too heavily invested with a particular set of meanings or interpretive communities, using widely available visual rhetoric, and not depicting anything in particular, as determined by a set of professional cultural intermediaries. They were meant to call attention to color capacities of television, not to themselves.

Watching the Future Watchers

The test scenario itself was fluid and subject to reconsideration. A year into the color television testing, the NTSC requested information from Eastman Kodak on optimal viewing conditions for the slides. In response, Eastman Kodak recommended a completely dark room and a matte screen: "Ambient light of the same color temperature as the projector light is usually not disconcerting, but daylight would certainly change one's opinion of the color quality on the screen" (NTSC 1951–1953b, 2). These conditions were meant to reflect an ideal laboratory site. But as the rejection of daylight suggests, an ideal laboratory situation was not the same thing as real-world viewing conditions.

As Trevor Pinch (1993, 25–26) writes, "test data are usually thought of as providing access to the pure technological realm, a means by which the immanent logic of a technology can be revealed." This is clearly a rhetorical situation of the test, but it also helps us understand the conditions under which engineers can cordon off a realm as "purely technological" or "immanent" to the technology, beyond human intervention. For us, the testing scenario is an epistemic scenario. By studying testing, we better understand how knowledge—of vision, of the televisual apparatus, and of the viewer and context of viewing television—was produced and managed in the early days of color TV technology. The history of the NTSC's color tests is at least in part an epistemic history. Of course, the test scenario is also an impossible situation, a context that

attempts to transcend all contexts (Sterne 2012, 152). It has a Kantian dimension in that the test seeks to overcome the interestedness of the participants to reach a technically mediated, disinterested conclusion about the "best" system for color reproduction, a system that can work for any image or any user. Yet the definition of *what works* and the criteria to measure success are the stakes of constructing a test scenario.

There were three nodes in the NTSC's constellation of test materials and subjects: the test viewers, the equipment being tested, and the slides. Each combination of two factors tested the third: images and equipment tested the viewers, viewers and images tested the equipment, and so on. "Because the testing context is carefully circumscribed, it is not always clear whether it tests the technology or the user" (Sterne 2012, 153—the following discussion draws heavily from Sterne 2012, 153–60). For instance, in early tests of the color TV apparatus before they had the Eastman Kodak slides, the NTSC measured subjective responses to different color slides with names like "watering can," "triangle thins," "basso," and "tea cups" As the engineers adjusted a filter, more and more viewers found the image satisfactory. Given that the goal of the color tests was to create an acceptable color standard that would work within the limited bandwidth available for television transmission, a test like this reveals two factors for consideration: (1) the point at which a certain measure of image quality becomes acceptable to a given number of viewers and (2) the proportion of viewers who at any given point rate an image as acceptable. These two factors were weighed against the available bandwidth in a cost-benefit analysis: how many viewers could be satisfied (or at least not annoyed) within the constraints of the infrastructure? The test scenario thus reproduces the conditions of infrastructural limitations and aesthetic compromise that animated the development of American color television. Not only did it test for the point at which image quality became acceptable to a majority of viewers, but it also presented gradations of the viewing population, offering choices for what proportion of the viewing population to try to please.

There is also a story about expertise here. The tests' viewing subjects in the tests stood in for all possible future viewing publics of color television. They were chosen from the ranks of professions concerned with television, and the list is revealing for its repetitive nature: engineers who designed TV receivers, broadcast engineers, engineers in "communications equipment design," television research engineers, television transmitter design engineers, and editorial writers for technical journals. The tests also had observers: teachers (what kind is unclear), television service engineers, and consulting engineers (DeCola, Shelby, and McIlwain 1954). On one level, the list is highly exclusive, consisting primarily of different kinds of engineers and the people who write for them or teach them; on the other, the sheer proliferation of types of engineers in the list shows that they hoped to approach the problem from all angles within the engineering profession. That panoply of engineers was the basis of the test group's claim to universality. Although we do not have demographic information on the participants beyond their last names, the list of professions combined with the roster suggests a group biased toward the white and male (Fink and NTSC 1955, 101). The expert viewers in the tests are therefore some distance from the population for whom they stand in. While they may be more discerning around image quality, they

also differed from the eventual audience in many ways, for instance, in their gender, ethnic, and class makeup, as well as their understandings of the inner workings of television technology. At the same time, the "normality" of the test group should be less of a foregone conclusion than we might at first assume. As Mara Mills (2010) has shown, a great deal of telephone research was conducted with the help of D/deaf and hard of hearing organizations. Without knowing how many of the NTSC's engineers required corrective eyewear, or had some kind of color blindness, it is difficult to claim that they were automatically normal, able-bodied users. The NTSC did give cursory attention to the affordances of the television system but were equally quick to dismiss them. In a questionnaire circulated to the committee's members, they asked and answered:

> *Question:* Taking into account the various possible *deficiencies* of *normal vision*, what types of color distortion or change may be anticipated as a result of each of these deficiencies?
>
> *Answer:* It was considered that the effect of color anomalies in vision would be confined to a small group, and that in any event there was little that could be done, in the design and operation of the system, to accommodate such anomalies. (Fink and NTSC 1955, 100, emphasis in original.)

Given this acknowledgment and dismissal of "anomalous" kinds of vision, the engineers' definitions of "normal" and "exceptional" images (discussed below) say more about their own cultural biases and the way those biases would get built into the aesthetics of a durable television format. This is a moment when a media-related disability is actually being created. Whereas color blindness may or may not have affected viewers of black and white TV, the NTSC engineers are writing off this population. Although they aim to exploit the gaps and absences in normal human vision (as they understand it), other gaps and absences simply go too far.

In creating a constant testing scenario, the engineers were asking a familiar question. "All else being equal," what is the most efficient way to satisfy a viewer with a pleasurable television image? Or more accurately, at what point does diminishing image quality actively annoy viewers? "Marginal" and "unacceptable" were shorthand for points at which the image quality might interfere with enjoyment of the program. Of course, "all else" can never truly be equal. But a testing scenario would be impossible without reducing the number of traceable variables. The constant, reproducible testing scenario creates the illusion that the contrast of two images can be reduced to a pure comparison. In describing the necessity of test images, the NTSC wrote,

> In carrying out tests in the field of color television, it is always convenient and sometimes essential that substantially identical subject matter be used by various laboratories and that ready comparisons be made between the original material transmitted and the received and reproduced image secured over a color television system. (Fink and NTSC 1955, 64)

When the slides are modeled through different compression standards, they are compared both to the original and to other compressed images to find the point of just noticeable difference. Once the results are consistent across enough testing scenarios, the engineers can wager that they have found the effective sweet spot for transmitting a satisfactory image.

When the Committee decided that it needed a standard that could fit *both* the color image and the monochrome image in one signal and that it would use knowledge of the human eye's low acuity for blue to do so, it was clear they needed a testing regime that could pinpoint the moment of just noticeable difference. By establishing the threshold between *satisfactory* and *marginal* television viewing, engineers could determine the bare minimum quantity of information required for inclusion in the signal. This measure, in turn, gave them a way to conserve bandwidth, to figure out how to shoehorn a color signal into the portion of the electromagnetic spectrum allocated to monochrome television. The testing scenario gave them the epistemic authority to make that call.

The NTSC in the Garden

Once the NTSC decided that still slides would constitute their test materials and that engineers would be their test subjects, they needed to decide how the slides would simulate the future of American broadcasting. In other words, what would they look like? The slides needed to simulate the television signal in two ways: at the level of form (their color, their shapes, their constitution) and the level of content.

At the level of form, the images are rather plain. They feature many people, a range of levels and types of shot, and a variety of shapes, textures, and contours. Only four of the images remain in color, and from this small sample, we can see that the test image featured large swaths of basic colors. As the committee requested both average and exceptional images, we can speculate as to which images are meant to be exceptional.

One clearly exceptional image was a picture of two boys in swimming trunks, one pulling the other by the hand. The image is titled "Motion" (Figure 3) and was apparently meant to stand in for the lack of moving pictures. The image's title points us back toward the futility of the exercise and the impossibility of capturing motion in a still slide. As with the discussion of the preference for live performers as test materials, "Motion" marks a fissure in the tests of a medium that, in the early 1950s, was defined by claims to liveness (Spigel 1992, 137–38). The presence of "Motion" reveals a faith that motion can be separated from action and represented through photography, and the remaining task is making sure that action is intelligible. The entire test was predicated on this impossible but necessary separation—a moving image screen technology that could only be calibrated through still images.

A number of other test images also evince visual exceptionality. And, like "Motion," the exceptional images are those that would be difficult for the television system to reproduce. Thus, like many communication engineers of their time, the NTSC enforced a rigid distinction between the content of a transmission and its potential *meaning.* For

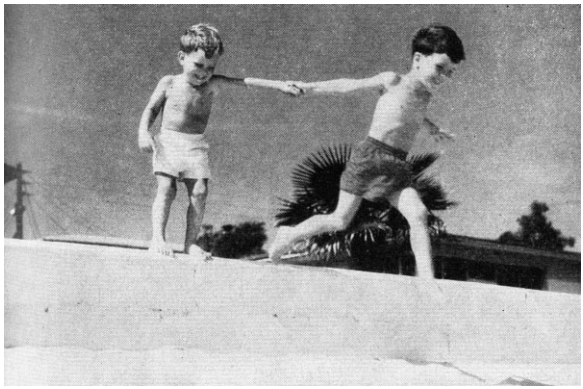

Figure 3. Motion. (Fink and NTSC 1955).

Figure 4. Meshes. (Fink and NTSC 1955).

instance, a woman shown behind a net (Figure 4) is a mundane close-up of a tennis player, but it is a challenge for a compression standard, because the net could easily introduce blurring and distortion. The same holds true for an image of dark plants against a dark backdrop (Figure 5)—the challenge here was picking up the image with such low contrast. Similarly, difficult images include jewelry (Figure 6) hanging over folds of glossy fabric—an image that indicates both an engineering challenge in how to pick out the fine detail of the jewels against a variegated background and a preoccupation with advertising imagery. The system's capacity to reproduce these images without encumbering the viewer was a measure of its capacity to manage unpredictable television programming.

Figure 5. Plants—dark background. (Fink and NTSC 1955).

Figure 6. Jewelry. (Fink and NTSC 1955).

Here, we see the separation of content and medium most clearly in the testing regime. In "Meshes" for instance, the image is meant to be read *not* as a woman behind a tennis net but rather as a contrast problem. The issue, of course, is that it is *also* a picture of a white woman behind a net. This is how racialized, gendered, and normed cultural sensibilities get built into technical standards. Not only are they not named, but the testing process demands that they not be named. Not surprisingly, each new imaging technology replays these gendered and raced norms, for instance, in the Microsoft Kinekt's initial problems in detecting black skin.

For the NTSC, exceptional images tested the limits of the system's ability to make people and things formally legible. They also tested the system's ability to differentiate people from things or animals and humans from nonhumans. This distinction was

essential both for the narrative and representational conventions of television—regardless of genre—and also for television's advertising function, which required the isolation and presentation of objects as desirable commodities. After Friedrich Kittler (1999), we are used to the idea that media apparatuses decompose materials of representation into signals that do not "care" about their content. In test images like those in Figures 4, 5, and 6, we see documents of engineers fighting to manage this technical tendency, to calibrate the machinic to the human, and to encode a humanist and consumerist aesthetic in a technical protocol.

If the test images were meant to include examples of both average and exceptional formal properties, it is safe to say that no such concern was extended to their content. Instead, the portrayal of American life on display suggests that the content was meant to be as normative as possible. "Normal" content, we are left to conclude, would allow the engineers to focus their efforts on abnormal forms. The images, in other words, could not accommodate both formal differences and cultural differences, as this would introduce one-too-many variables in testing the television signal. If this is the case, then an analysis of the content of the images reveals what was considered normal subject matter for the future of TV.

In general, the NTSC slides portray a narrow band of WASP leisure: children row; teenagers sail; adults also sail, though sometimes they go for a sleigh ride. Just as one woman can stare deeply into the eyes of a goose ("Goose girl"), another can grasp a kitten in a pile of hay ("Kitten girl"), or another can finish off a large crescent of watermelon ("Watermelon girl"). Only two images ("Aviator" and "Boat-ashore pair") feature a direct address to the camera, while nearly all the images of people are medium-to-long shots. Labor is almost entirely effaced: in two separate images, we see a pair of women picking flowers ("Tulip garden") and a chauffeur waiting with his horse-drawn carriage ("Southern manse"). Labor might in fact be difficult to portray in these images because electricity is also absent: except for the protruding arm of a propeller ("Aviator"), a lampshade without a lamp ("Lamp shade"), and four isolated dials ("Instruments"), you could confuse the world represented for a pre-electric United States. As such, the images are exemplary of what Michael Schudson once called "the aesthetic of capitalist realism" which, "without a masterplan of purposes—glorifies the pleasures and freedoms of consumer choice in defense of the virtues of private life and material ambitions" (Schudson 1984, 218).

The suburban and pastoral sensibility represented in the images fits nicely with clichés we hear about 1950s American culture and television, and they seem to advance arguments that the content of 1950s television was doing a particular kind of ideologically conservative work that promoted both consumerism and the containment of women (Brunsdon, D'Acci, and Spigel 2007). Certainly, it seems that the ideological content of mainstream American television reached back, beneath its programming, into the very test slides and psychophysical studies that were used to form its technical standards. But this form of capitalist realism did not simply reflect a cultural consensus of the time. As Carol Stabile (2011) has argued, the landscape of 1950s television, with all its faux-wholesomeness, was a product of the broadcast blacklist, where writers who had any social vision remotely challenging to a McCarthyite mainstream were

systematically shut out. The quaint ideological tinge of the NTSC images—while not necessarily themselves shaped by a blacklist—certainly were of the same moment. As Stabile notes (2011, 267), *Red Channels: The Report of Communist Influence in Radio and Television* was released in 1950 (see also Doherty 2003), a year before the second NTSC got to work and right in the middle of the McCarthyite era.

In total, the slides portray thirty-three people (twenty adults and teenagers, thirteen children), along with three kinds of animal (a goose, a kitten, and two horses), and a range of household objects, framed to reflect commercial presentations. Among these objects are close-ups of coasters, jewelry, fishing tackle, a fruit bowl, an advertisement for a lighter, and a pile of pumpkins. There are no radios, or televisions, or paintings; in fact, there are no media artifacts of any kind. Although the work of capitalism is effectively effaced, Eastman Kodak and the NTSC were clearly concerned with guaranteeing the faithful reproduction of ads.

Pointing out these aspects of the slides does not imply a call for more representations of toasters and factory workers, or to argue that a picture of a TV ought to be included in the imaginary future of the medium. Rather, it is to argue that technical tests are also modes of representation and that they have political dimensions even when their content may not be the first concern of engineers. During the years of the Second World War and the period immediately after, broadcasters worked to convince advertisers that television could fulfill the same advertising potential of radio (Schwoch 1990; Sewell 2014). In this period, television's technical capacities were inherently tied to its capacities to represent products and carry advertisers' messages. In other words, representational concerns were forefront in the minds of television's architects.

This brings us to the most obvious problem: despite being chosen for their demonstration of color, the images are thoroughly, monochromatically *white*. Not only are these images startlingly alabaster but they also shine conspicuously with the acknowledged burden that they were specifically employed to guarantee the faithful reproduction of white skin tone. This was, at one level, an impossible goal as the number of skin tones covered by the racial fiction of "white" was quite broad even in the 1950s. At the same time, the idea of skin being white was well established in other media and cultural industries, from Crayola's "flesh" crayon that was in use until 1962 to the composition of film stock (Dyer 1997; Winston 1996). Earlier, monochromatic television productions of the 1930s used special makeup techniques to "keep white men from looking like 'Uncle Tom'" (Sewell 2014, 133). Standard Hollywood practice at the time involved bouncing additional, reflected light onto African American performers because color film did not easily render black skin. As Brian Winston points out, this was the result of a set of conscious aesthetic decisions that came up with color film and yet again in tests of Kodachrome film that were coterminous with the NTSC. In those tests, color tones for skin were selected for their pleasing character rather than their realism: "Caucasian skin tones are not to be rendered as they are, but rather as they are preferred—a whiter shade of white" (Winston 1996, 41, 43, quote at 56; see also Dyer 1997). As Lorna Roth has shown, both American and European color TV relied on "Shirley" cards that were used for calibration of white skin tone—to the exclusion of

Figure 7. Southern Manse. (Fink and NTSC 1955).

all others. Over time, the various "Shirleys" used for white-skin-tone calibration reflected changing standards of skin beauty (Roth 2009, 112–15).

The exclusive use of white models in the slides means that interpretations of white skin tone were directly encoded in the color standard—or rather, that the NTSC attempted to write them in. Aside from one slide titled "Southern Manse" (Figure 7), in which the two figures are not legible, the images are staunchly white. This may not be surprising given the era, but decades later, we know that black skin is not well reproduced by NTSC color or by traditional film stock, a theme that has been taken up by both television and film producers in recent years (Hornaday 2013). This simple, "practical" decision had tremendous repercussions on the ability of color television to represent large segments of the population of the United States. Unspoken but clearly central was the articulation of flesh color and the politics of race in the United States in the early 1950s. That the NTSC was calibrating "skin" to aesthetics of whiteness just as the civil rights movement was about to take off in the United States only underscores the political ramifications of this decision. Regardless of the intentions behind them, the design decisions in the NTSC's test images effectively biased the format toward rendering white people as more lifelike than other races, at least within the codes of televisual realism.

The Universal and the Peculiar

Often test images are examples of natural phenomena or objects that are meant to stand in for a whole group. The NTSC slides, however, could not be based on "natural phenomena." Instead, they had to be based on the potential world of American television. Test images constitute a standard set of data—a known quantity—and when

testing a new image compression and transmission standard, engineers are interested in how that data will be transformed through various kinds of mathematical and physical manipulation. Combined with the criteria of judgment based on whether test subjects find the images pleasurable or annoying, bright or dark, clear or fuzzy, too red or green, test images form a shared toolkit for detecting and registering qualitative differences between different image standards.

Test images are an example of what Lorraine Daston and Peter Galison call "working objects." The business of making scientific investigation commensurable and exchangeable—the work, in other words, of "collective empiricism"—requires a cordoning-off of the "too plentiful and too various" world of "natural objects" (Daston and Galison 2007, 19). Instead, a limited selection of common objects stands in for the larger, unmanageable set, and forms the communal working objects of scientific research. These proxies may be typical specimens, average examples, or exceptional instances. Working objects are an indispensable part of making scientific knowledge sensible and comparable.

The NTSC images move between the universal and the particular twice. When the engineers choose and use twenty-seven images as proxies for all of American TV, they imagine a universal set through a defined set of particulars. When the standard is finalized and becomes the basis of a massive communication infrastructure, television engineering reverses the process, universalizing from a small but manageable set of data. The act of sampling from the universe of all possible images requires choices about who or what to represent. The practical politics of standards are divided between two processes: "arriving at categories and standards, and, along the way, deciding what will be visible or invisible within the system" (Bowker and Star 1999, 44). For the NTSC engineers, their categories were limited to the "general subject matter of television" and formally "average" and "exceptional" images. The engineers were quick to request more green images when they needed them. But they exercised a limited view of the cultural context of American life in the 1950s. This process of rendering some things—and people—less visible exemplifies the cultural stakes behind technical standards.

Compare this with recent standards for energy consumption in new televisions, for which engineers compiled a ten-minute DVD of the average content of 200 hours of five countries' viewing habits.[1] The montage of "typical broadcast content" was the product of intense negotiation over the test materials. Manufacturers knew that the choice of content would bear greatly on how much energy the TVs used. LCD manufacturers wanted a signal with more white, while plasma TV manufacturers preferred more black: "Neither group of manufacturers wanted to lose the power battle due to a measurement specification that favored one technology over another" (Fairhurst 2009, 474). The DVD was a compromise that resorted to an "average" of current content as its reference point. The contemporary TV engineers were therefore aware of the relationship between the technical and the representational and heavily invested in shaping representational proxies to produce specific technical results.

The NTSC exemplifies a prehistory of late-twentieth century attempts by industrial actors—content producers and technology manufacturers—to control standardization

processes. In the case of the NTSC, and the engineering arms of NBC, CBS, and DuMont, the manufacturers and content-makers were frequently one-and-the-same. Because the color NTSC formed against the wishes of the FCC as a private, industrial consortium, and because of the vertical integration of these firms, its choices about how to simulate American television—its presumptions, in other words, of *how* the audience sees, and *what* the audience sees—deserve close scrutiny.

American color television is a special case in the history of standards. Not only did its standard operate for an inordinately long time—depending on where you live, at least fifty-six years—but it also resulted (albeit acrimoniously) from a coordinated industry effort to gain government approval, a process that was bloodless compared with the format wars of subsequent decades. Despite its extreme dimensions—still pictures for a moving image standard, impossible (for the time) bandwidth restrictions, pastoral content, the outright collapse of content and service provision—the NTSC's task is in many ways exemplary of the kinds of politics that lie behind the establishment of standards for our most common media formats. When RCA set out to create its 45 rpm single, it settled on the three-minute length through a survey of its back catalogue of songs; Columbia performed a similar calculation with its classical music holdings when deciding on the length of the LP record (Papenburg 2012). Standardized fonts and layouts for print are selected for their reproducibility and with assumptions about the reader and the act of reading (Barnhurst and Nerone 2001). The NTSC's work has also had an illustrious afterlife. Beyond its own format's longevity, its basic working assumptions and approach to bandwidth are common for digital images, most of which are also compressed. For instance, a widely cited paper from 1956 ("Some Studies in the Speed of Visual Perception") by George C. Sziklai used "Southern Manse" as one of its test images. Moreover, the approaches to image compression in formats like JPEG and MPEG directly reference the NTSC's earlier work, from the application of psychophysics to the composition of test personnel and images (Cubitt 2011; Mackenzie 2008; Sziklai 1956).

Beyond the NTSC's media historical legacy, we can see many of the same issues cropping up in test images today. Test images, along with calibration images, are part of any new image technology and shape new image standards and their capacity for reproduction. Some test images have even formed something of a canon. Below, we see five examples. At least one of these five images is likely to be found in any image processing textbook of the past forty years: the mandrill, Lena, the fishing boat, peppers, and a fingerprint (Figure 8). By far, the most popular picture in image compression, however, is Lena. Lena is torn out of the Playboy centerfold from November 1972, and by some accounts, it was one of the first images digitized for network transmission (Pennebaker and Mitchell 1993; Salomon 2007).[2]

In the end, it turns out that the politics of test images share some important features with the broader cultural politics of images. The NTSC test images vanished from the history of television because like other test materials and technical images they are dismissed as trifling instruments of a larger technical practice. As such, they appear neither technical enough to merit mention in a history of technology nor conspicuous enough to catch the gaze of visual culture studies. If nothing else, we offer in this

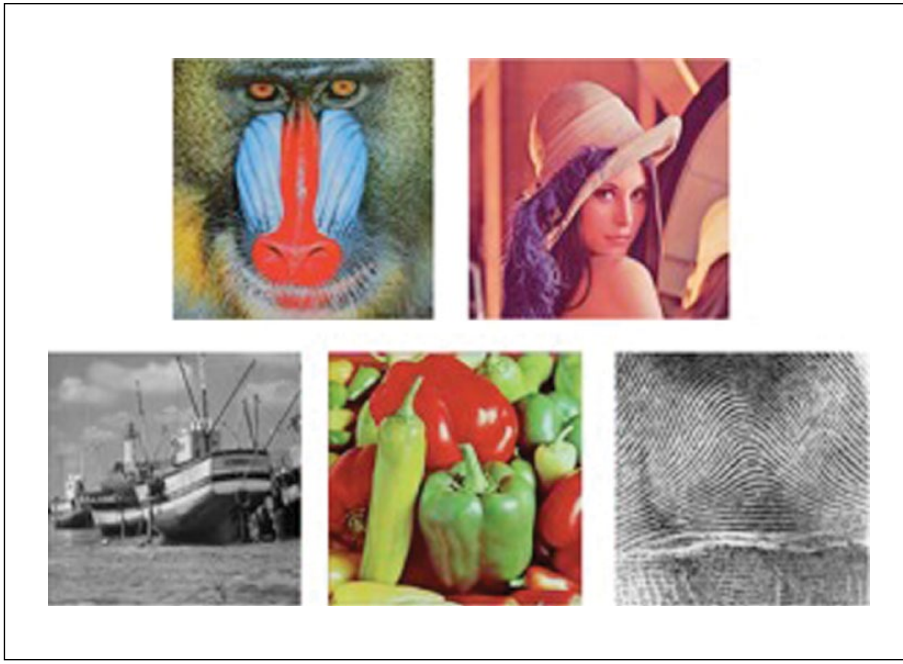

Figure 8. Five standard test images (clockwise from top-left): Mandrill, Lena, Fishing Boat, Peppers, Fingerprint.

article a third option: an investigation of the visual culture of technical practices. Test images become commonplaces that engineers use to compare tests over space and time. Unlike the NTSC, no one mandates the use of most test images, though convention leads to the use of some images more than others, since canonical images can work as a kind of *lingua franca*. Their canonization in textbooks reinforces engineers' more informal choices of certain images as touchstones. These images, as well as their placement and framing in technical manuals, carry with them cultural understandings of the purpose of imaging technologies and the content they are meant—or not meant—to convey. Universal aspirations are rooted in particular examples meant to be stand-ins for future content, where the images function as working objects in Daston and Galison's (2007) terms. Although current engineering practice demands this kind of selection, the subject of the images is political. As a product of the testing situation, the subject of test materials is never neutral.

We have offered a prehistory of American color television through the documentary vestiges of the NTSC. Using the Kodak test slides, some in color, some in monochrome, and the advantage of historical perspective, we know that *something* happened in the development of the color television standard that meant that it failed to consistently reproduce the range of human skin tones—white and nonwhite alike. This failure is a byproduct of the testing regime, combined with the workings of the

technology, the bandwidth allotted for broadcast, and the limitations of both standards bodies and audiences. When we ask how the NTSC could have failed so severely in guaranteeing the faithful reproduction of skin tone—white skin, its goal, and many other skin tones besides—we may want to start with their test materials and test subjects, and work our way up. The engineers were attuned to the shortcomings of the first batch of images with regard to colorimetric values, that is, the need for more green. But in our survey of the NTSC's remaining documents, we did not find anyone who expressed a concern about the monochromatic skin tones of television's subjects or offered commentary on the portrayal of a pastoral, consumerist lifestyle in 1951. "No true skin color" might be an amusing and accurate critique of the NTSC's format. But as a statement about the politics of race, it also signals the inescapable political contradictions faced by all image standards, past and future: there is no *true* skin color.

Acknowledgment

The authors would like to thank Vicki Mayer, Sue Murray, Carolyn Kane, Andreas Fickers, Carrie Rentschler, Nancy Baym, Kate Crawford, Tarleton Gillespie, Jessa Lingel, Kate Miltner, Microsoft Research New England, our audience at the Society for Cinema and Media Studies, and two anonymous reviews for support, conversation, and advice as we wrote and revised this article. We would also like to thank the Social Sciences and Humanities Research Council of Canada for its financial support.

Authors' Note

The full, official version of this article can be found at http://colortvstandards.net. There, we have constructed a version in scalar (http://scalar.usc./edu) that includes a complete set of images discussed in this article and additional comments on them. It is searchable and printable. We have also maintained the pagination for easy citation of the digital version. We consider the scalar version the official version of the article, and this pdf as a placeholder in the journal's system.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Social Sciences and Humanities Research Council of Canada and Microsoft New England.

Notes

1. "The project members measured at least forty hours of typical broadcast content, including a variety of genres from a variety of broadcast stations in Australia, Japan, the Netherlands, the United Kingdom, and the United States" (Commonwealth of Australia 2008, 30).
2. These are two key volumes: one the JPEG handbook and the other a data compression textbook. However, a search of almost any image processing volume will find at least one of these images.

References

Abramson, Albert. 2003. *The History of Television, 1942 to 2000*. Jefferson: McFarland.

Barnhurst, Kevin, and John Nerone. 2001. *The Form of News: A History*. New York: Guilford Press.

Boddy, William. 1990. *Fifties Television: The Industry and Its Critics*. Urbana: University of Illinois Press.

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge: MIT Press.

Brunsdon, Charlotte, Julie D'Acci, and Lynn Spigel. 2007. *Feminist Television Criticism: A Reader*. Oxford: Oxford University Press.

Cubitt, Sean. 2011. "Current Screens." In *Imagery in the 21st Century*, edited by Oliver Grau, 21–35. Cambridge: MIT Press.

Commonwealth of Australia (Department of Resources, Energy and Tourism). 2008. "Test Method for Measuring Average TV Power Consumption." *In Methods of measurement for the power consumption of audio, video and related equipment: Committee for Safety of Electronic Equipment*.

Daston, Lorraine, and Peter Galison. 2007. *Objectivity*. New York: Zone Books.

DeCola, Rinaldo, Robert E. Shelby, and Knox McIlwain. 1954. "NTSC Color Television Field Test." *Proceedings of the IRE* 42 (1): 20–43.

Doherty, Thomas. 2003. *Cold War, Cool Medium: Television, McCarthyism, and American Culture*. New York: Columbia University Press.

Dyer, Richard. 1997. *White*. London: Routledge.

Fairhurst, Jon. 2009. "How an International Standard Reduced TV Power Consumption." *Proceedings of the IEEE* 97 (3): 474–75.

Fickers, Andreas. 2010. "The Techno-Politics of Colour: Britain and the European Struggle for a Colour Television Standard." *Journal of British Cinema and Television* 7 (1): 95–114.

Fink, Donald G. 1951. "Alternative Approaches to Color Television." *Proceedings of the IRE* 39:1124–34.

Fink, Donald G., and National Television System Committee. 1955. *Color Television Standards; Selected Papers and Records*. New York: McGraw-Hill.

Frosh, Paul. 2003. *The Image Factory Consumer Culture, Photography and the Visual Content Industry*. London: Berg.

Fuller, Matthew. 2005. *Media Ecologies: Materialist Energies in Art and Technoculture*. Cambridge: MIT Press.

Garner, Gretchen. 2007. "Photography and Society in the 20th Century." In *The Focal Encyclopedia of Photography*. 4th ed., edited by Michael R. Peres, 187–98. Boston: Focal Press.

Hornaday, Ann. 2013. "'12 Years a Slave,' 'Mother of George,' and the Aesthetic Politics of Filming Black Skin." *The Washington Post*, October 18. Accessed from http://www.washingtonpost.com/entertainment/movies/12-years-a-slave-mother-of-george-and-the-aesthetic-politics-of-filming-black-skin/2013/10/17/282af868-35cd-11e3-80c6-7e6dd8d22d8f_story.html

Kane, Carolyn. 2014. *Chromatic Algorithms: Synthetic Color, Computer Art, and Aesthetics after Code*. Chicago: Chicago University Press.

Kittler, Friedrich A. 1999. *Gramophone, Film, Typewriter*. Stanford: Stanford University Press.

Mackenzie, Adrian. 2008. Codec. In *Software Studies*, edited by Matthew Fuller, 48–54. Cambridge: MIT Press.

McLuhan, Marshall. 1964. *Understanding Media*. London: Routledge.
Mills, Mara. 2010. "Deaf Jam: From Inscription to Reproduction to Information." *Social Text* 102:35–58.
Murray, Susan. Forthcoming. *Brought to You in Living Color: A Cultural History of Color Television*. Durham: Duke University Press.
National Television System Committee. 1951–1953a. *Report and Reports of Panel No. 11, 11-a, 12–19*. Vol. 2.
National Television System Committee. 1951–1953b. *Report and Reports of Panel No. 11, 11-a, 12–19*. Vol. 3.
Papenburg, Jens. 2012. "Serial Songs and Coherent Works: Technicization of Music Listening in the Single and LP around 1950." The Art of Listening, Berlin, Germany, July 12–14.
Pennebaker, William B., and Joan L. Mitchell. 1993. *JPEG Still Image Data Compression Standard*. New York: Van Nostrand Reinhold.
Pinch, Trevor. 1993. "'Testing—One, Two, Three. Testing!': Toward a Sociology of Testing." *Science, Technology, & Human Values* 18 (1): 25–41.
Remley, Frederick M. 1999. "The Challenge of Recording Video." In *Magnetic Recording: The First 100 Years*, edited by Eric D. Daniel, C. Denis Mee, and Mark H. Clark, 124–36. New York: IEEE Press.
Roth, Lorna. 2009. "Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity." *Canadian Journal of Communication* 34 (1): 111-136.
Salomon, David. 2007. *Data Compression: The Complete Reference*. London: Springer.
Schudson, Michael. 1984. *Advertising, the Uneasy Persuasion: Its Dubious Impact on American Society*. New York: Basic Books.
Schwoch, James. 1990. "Selling the Sight/Site of Sound: Broadcast Advertising and the Transition from Radio to Television." *Cinema Journal* 30 (1): 55–66.
Seldes, Gilbert. 1956. *The Public Arts*. New York: Simon & Schuster.
Sewell, Philip. 2014. *Television in the Age of Radio: Modernity, Imagination, and the Making of a Medium*. New Brunswick: Rutgers University Press.
Shannon, Claude, and Warren Weaver. 1949. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.
Siegel, Greg. 2014. *Forensic Media: Reconstructing Accidents in Accelerated Modernity*. Durham: Duke University Press.
Spigel, Lynn. 1992. *Make Room for TV: Television and the Family Ideal in Postwar America*. Chicago: University of Chicago Press.
Stabile, Carol. 2011. "The Typhoid Marys of the Left: Gender, Race, and the Broadcast Blacklist." *Communication and Critical/Cultural Studies* 8 (3): 266–85.
Sterne, Jonathan. 2012. *MP3: The Meaning of a Format*. Durham: Duke University Press.
Sterne, Jonathan, and Dylan Mulvin. 2014. "The Low Acuity for Blue: Perceptual Technics and American Color Television." *Journal of Visual Culture* 13 (2): 118–38.
Sziklai, George C. 1956. "Some Studies in the Speed of Visual Perception." *IRE Transactions on Information Theory* 2 (3): 125–28. doi:10.1109/TIT.1956.1056814
Williams, Raymond. 1974. *Television: Technology and Cultural Form*. New York: Schocken Books.
Winston, Brian. 1996. *Technologies of Seeing: Photography, Cinematography and Television*. London: British Film Institute.

Author Biographies

Dylan Mulvin is a PhD candidate at McGill University. His dissertation is titled *Reference Materials: The People, Places, and Things of Making Measurements*. He has published on the histories of video and television.

Jonathan Sterne teaches in the Department of Art History and Communication Studies at McGill University. He is author of *MP3: The Meaning of a Format, The Audible Past*, and over fifty articles. His website can be visited at http://sterneworks.org.