Sound and color

by Edward Buscombe

from Jump Cut, no. 17, April 1978, pp. 23-25
copyright Jump Cut: A Review of Contemporary Media 1978, 2005

The last issue of Film Reader[1] devoted half of its total space to examining the relations between industry, technology and ideology in the cinema. Film Reader's initiative is a welcome sign that film theory is paying more attention to economic and technological determinants and that film history is increasingly moving out of the era of mere facts and figures towards consideration of more substantive matters.

However, an article by J. Douglas Gomery in this issue,[2] though providing valuable detailed information on the introduction of sound into Hollywood, raises some problems concerning the extent to which economics can assist our understanding of the cinema. Gomery claims that "economic theory can explain the coming of sound."[3] Gomery has in mind the theory of technological innovation. [4] This theory seeks to explain the factors governing the invention, innovation and diffusion of new technology in any given industry: in what circumstances new techniques or products are first invented and then introduced as practical and commercial propositions subsequently adopted by the industry as a whole. A considerable literature exists on this subject, but we may take as representative the work of one author cited by Gomery. Edwin Mansfield, in his book Technological Change,[5] lists several factors governing a decision to innovate once an invention has been produced:

"To begin with, the firm should estimate, of course, the expected rate of return from introducing the new product or process. In the case of a new product the result will obviously depend on the capital investment that is required to introduce the innovation, the forecasted sales, the estimated costs of production, and the effects of the innovation on the costs and sales of the firm's existing product line. … In addition the firm should estimate, as best it can, the risks involved in innovating." [6]

Mansfield also enumerates those factors affecting the rate at which an innovation will become diffused:

"1) the extent of the economic advantage of the innovation over older methods or products, 2) the extent of the uncertainty associated with using the innovation when it first appears, 3) the extent of the commitment required to try out the innovation, and 4) the rate of reduction of the initial uncertainty regarding the innovation's performance" (Mansfield, p. 88).

Mansfield also suggests that a number of factors might be expected to affect the speed of any single firm's response to a new technique:

1. the size of the firm: one would expect larger firms with more resources to be quicker at innovating.
2. the degree of expectation of profit from the new technique.
3. the rate of growth of the firm: expanding firms might innovate more easily.
4. the firm's profit level: prosperous firms would have the necessary
capital or credit.
5. the age of the firm's management personnel: younger management might be more receptive to new ideas.
6. the liquidity of the firm: the more liquid the firm, the better it might be able to find finance.
7. the firm's profit trend: firms with declining profits might look harder for new profits or techniques (Mansfield, p. 93-95).

There is nothing very profoundly "theoretical" about Mansfield's formulations, yet they do have some explanatory power in relation to the coming of sound. Gomery has shown that Warner Bros. did pay careful attention to the question of costs and to the problem of finding the necessary capital. [7] Furthermore Mansfield's four factors affecting the rate of diffusion help to explain why the changeover to sound was so rapid. The economic performance of the new product and the speedy reduction in the uncertainty regarding that performance more than outweighed the original uncertainty itself and the high costs of installing new equipment.

The seven factors characterizing those firms most likely to innovate should provide an explanation for the fact that it was Warners, one of the smaller companies, which led the way in sound. Unfortunately, the theory of technological innovation breaks down at this point, since Mansfield can find no statistically significant correlations across a range of industries for factors 3, 4, 5, 6 and 7. The only factors known to affect a firm's willingness to innovate are its size (bigger firms do innovate faster) and the expected rate of profit. The latter point seems fairly obvious, while the former shows Warners to be an exception to the rule. Gomery is forced to look elsewhere for an explanation of Warners' actions, which were, he claims, the result of the farsightedness of Waddill Catchings, the entrepreneur who masterminded the firm's strategy.

Gomery's theoretical position therefore ends up not so far as he thinks from that of the film historians he takes to task. He sets out to prove that sound was introduced as the result of an economic law which "theory" can explain. But instead, sound turns out to be the result of one man's initiative. The only substantial difference between Gomery's explanation and that of previous historians is a dispute over which individuals should get the credit, Catchings or the Warner brothers themselves.

Thus the theory of technological innovation seems of limited use, and Gomery reverts from a search for economic explanations back to a kind of "great man" theory of history. But could a different kind of economic theory explain the coming of sound? This would depend on what kind of explanation we are looking for. The theory Gomery wants to use could only explain why it is that innovation takes the course it does. It doesn't explain why there should be innovations in the first place, a more fundamental and surely more interesting question. To answer it, we cannot adopt a simple notion of supply and demand, since the public could hardly be said to have demanded sound pictures until it had seen and heard them. True, once sound had been successfully demonstrated, demand affected the rate of diffusion. But the initial investment in research and development had to be made when future demand could only be guessed at.

One must start with the fundamental law that in a free market economy, a firm is motivated by, to use the terms of capitalist economics, a desire to maximize profits; or, in the terms of Marxist economics, a desire to maximize the rate at which it extracts surplus value. In any given economic situation, this can be done in a number of ways. For example, a firm can attempt to develop fresh markets and so achieve economies of scale. In the late 20s, the film industry had no easy way of finding fresh markets — domestic and foreign penetration of the market being near saturation point. (In 1926 U.S. attendances ran at 100 million a week. In Britain, for example, U.S. films had 74¾% of the market at this time.)
Another possibility is for a firm to lower its costs of production. Given that constant capital costs, both fixed and circulation (that is, the costs both of buildings and machinery, and of raw materials), were relatively inelastic, this could be done only by reducing the cost of variable capital, i.e., labor. (I am assuming, though I cannot prove it, that in the film industry in the late 20s, the costs of constant capital were in fact inelastic.) But in a labor-intensive industry such as filmmaking, and one in which automation had at that time gone as far as it could go (another assumption I cannot actually prove), it seems as though there was little opportunity for cutting costs. However, it is worth noting in this respect that Warners’ original motive in developing sound was to use it as a means of recording vaudeville acts and musical sound tracks for silent pictures. In other words, sound was at first intended to increase the productivity of vaudeville performers and theatre musicians. Only subsequently was it seen as a means of creating an entirely new product.

Another way for a firm to increase the rate of surplus value is to increase its share of the existing market at the expense of its competitors. This can sometimes be achieved by price cutting. But the U.S. film industry had evolved by the end of the 1920s into a mature oligopoly in which the sale of the product (i.e., exhibition) was tightly regulated by the major firms dominating the market, in cooperation with each other. Each production company needed the sales outlets (theatres) of the others in order to market its products. Thus none of the large companies could involve itself in a price war against the wishes of the others. The smaller companies, who might have had most to gain from price competition, were in the weakest position to do anything of the kind, because few of them had theatres of their own and because the majors controlled the most important theatres.

Only one way remains in such a situation for a company to secure an advantage over its competitors. It can create a new product. In a sense, of course, this happened all the time in Hollywood, since every picture was unique and its uniqueness was protected by copyright. But precisely because all products were unique, no company possessed a decisive advantage. This required an innovation of a different order. Such an innovation was sound, a wholly new kind of product, which would make all other kinds obsolete. And the possession of this invention did, indeed, for a time give Warners a chance to increase its share of the market. (It seems likely also that it did for a while increase the absolute size of the market, bringing new customers into the theatres. And it may have helped postpone the decline in attendances brought on by the depression.) The profits which a monopoly on a new product make possible are known in Marxist economic theory as “technological rent.” [9] And the search for this monopoly explains why innovation should be a necessary feature of the economic system even when business seems good.

From this perspective, we should not view innovation in the film industry as a rational and sought-for outcome of attempts on the part of altruistic inventors to “improve” film technology, nor as proof of capitalism’s success in combining profit with the satisfaction of human needs. Human needs are many, but capitalism will produce only those innovations from which rent can be extracted, since the whole basis of the system is production for exchange value rather than use value. Sound would not have succeeded, admittedly, had not the public found a use for it; but the public was given “what it wanted” only because sound offered the opportunity for a monopoly. And the same principle applies, mutatis mutandis, to any other technological innovation. The history of the invention of the camera itself is written largely in the patents taken out for each new modification.

Gomery argues convincingly against those film historians who claim that Warners decided to produce sound films in a desperate gamble to ward off bankruptcy. Gomery shows that the decision formed part of a carefully thought-out strategy to upgrade the company’s status to that of
But the case of sound (introduced by Warners and Fox, at that time two of the smaller studios) does not show that technological innovation in the cinema results only from a special set of circumstances. An oligopoly reduces competition in certain areas; it does not eliminate it altogether. Firms continue to compete with each other, but the main form of competition takes the shape of a search for new products. Innovation and technological rent are functions of the system as a whole, not just the result of attempts by small firms to break into the big time. The first three-component Technicolor film, for example, was released by RKO and the first CinemaScope picture by Twentieth-Century Fox, both majors.

II

Economic theories can only partially explain technological innovations; economics cannot say why innovations take the form they do, only why they are an essential part of the system. Economics can explain the necessary but not the sufficient conditions for innovation. No new technology can be introduced unless the economic system requires it. But a new technology cannot be successful unless it fulfills some kind of need. The specific form of this need will be ideologically determined; in the case of cinema the ideological determinant most frequently identified has been realism. Whether the search for greater realism has been welcomed, as in the case of Bazin's discussion of deep focus or Charles Barr's of CinemaScope, [11] or whether realism is subjected to a fundamental critique, as in the case of writings by Comolli and Baudry, [12] theorists appear to agree that realism indeed dictates the formation of the needs which technology satisfies.

But to define "realism is no simple matter. And while we may agree that realism is dominant, it may not always be the only ideological need fulfilled by technological innovations. The history of the use of color in the cinema provides an interesting test case for the precise role of realism. The scientific principles of color, like those of sound, were known long before sound or color films became technically and commercially feasible. With color, as with sound, the delay in its introduction resulted in part from technical problems in producing a system that would work under commercial operating conditions (early color films were very prone to scratching, for example). But again as with sound there was also resistance on aesthetic grounds. Douglas Fairbanks, whose picture THE BLACK PIRATE (1927) was produced in two-component Technicolor, complained that color had

"always met with overwhelming objections. Not only has the process of color motion picture photography never been perfected, but there has been a grave doubt whether, even if properly developed, it could be applied, without detracting more than it added to motion picture technic. The argument has been that it would tire and distract the eye, take attention from acting, and facial expression, blur and confuse the action. In short it has been felt that it would militate against the simplicity and directness which motion pictures derive from the unobtrusive black and white." [13]

Such objections appear rather strange if one supposes that the demand for realism in the cinema has always been merely a question of the literal rendering of appearances. We perceive the world as colored, after all, and therefore an accurate representation of it should also be colored. (Leaving aside the fact that complete accuracy is impossible since color in film only approximates the colors perceived in the real world.) But in fact it has never been a question of what is real but of what is accepted as real. And when it first became technically feasible, color, it seems, did not connote reality but the opposite.

This may in part be for historical reasons, since the very first uses of color involved the tinting of certain sequences in films shot in black and white. Such a usage was extremely conventional, a long way from a
literal representation of the world. And as I suggest below, there may be more important reasons why color was not accepted as connoting reality. At any rate, the objections to which Fairbanks refers are clearly consistent with a realist aesthetic. Color would serve only to distract the audience from those elements in the film which carried forward the narrative: acting, facial expression, "the action." The unity of the diegesis and the primacy of the narrative are fundamental to realist cinema. If color was seen to threaten either one, it could not be accommodated.

It thus becomes possible to understand why color took so much longer to take hold than sound. The technical problems were probably no greater, nor was it simply force of habit. Audiences accustomed to silent pictures adapted to sound practically overnight. Color, on the other hand, has become universal only since the advent of color television, which lowered the relative resale (to television) value of theatrical features made in black and white. Color technology has taken so long to diffuse, we can conclude, partly because unlike sound it could not be instantly accommodated to the realist aesthetic.

Further evidence of color's "unreality" for early spectators can be found in the use actually made of it. For example, in the first few years after the introduction of three-component Technicolor (originally used in the Disney cartoon *Flowers and Trees* in 1932), the great majority of films employing the process were produced within genres not notably realistic in the sense of their being accurate representations of what "life" is "like." It can be argued, of course, that not many Hollywood pictures represent what life is like. But it nevertheless remains true that a kind of hierarchy ranks genres according to the extent to which the world they portray, fictional or not, is close to what the audience believes the world to be like. Thus at one end of the scale, we find newsreels, documentaries, war films, crime films, etc., and at the other, cartoons, musicals, westerns, costume romances, fantasies, comedies. Virtually all the early three-component Technicolor pictures are in these latter genres.

Thus by the 1930s the original objection to color, that it would detract from the narrative, had given way to the extent that color was permissible in some films, and so therefore no longer totally incompatible with audience concentration on a story. (Of course such an objection as Fairbanks describes must always have been an extreme position since certain uses of color such as tinting became quite common very early on.) Yet it was still considered sufficiently unrealistic to be taboo for films with "realistic" subject matter. [15]

We must now return to the question of why color was not perceived as realistic. Why was its use during the 1930s restricted to unrealistic genres, whereas the use of sound was not? Color must surely have connoted something else. What that something else was could, I think, be demonstrated by an analysis of the color films produced.

But I propose instead to take a short cut and consult an industry manual published in 1957, *Elements of Color in Professional Motion Pictures* [16]. Written by a committee of film industry personnel, it distills the collective theory and practice of color photography in Hollywood up to the late 50s. By this time the use of color was no longer restricted to certain genres; by the date of publication, the authors suggest, two-thirds of all features were produced in color. Nevertheless, certain of their remarks on the relation of color to realism shed some light on why for a long time color was restricted to special uses.

For the authors of this book, one should note first of all, realism is never to be equated with naturalism, strict fidelity to the world as it appears:  

"This psychological factor can be of great importance in creating an atmosphere of reality or verisimilitude on the screen. With the filming of a historical or 'period picture,' for example, research is done not only on architecture and
decoration, but also on the colors in use during the particular period and in the specific country. Yet the use of the actual colors of the period or the country are very rarely employed (sic). Because of psychological factors governing the response of a modern viewing audience, far better results are achieved by the use of a desaturated tonality of the times — that is, a less saturated range or ‘palette’ of color and pattern, but adequately punctuated with authentic identifying colors so that the end result stands to be identified as historically accurate yet believable” (Elements, pp. 41-42).

The colors we accept as real are therefore a compromise between what we are accustomed to and what used to be. The need to make the audience believe in what is depicted on the screen permits, indeed demands, a distortion of what actually is, or was. Such a practice can, of course, be observed in other aspects of Hollywood filmmaking, though the practitioners are rarely so honest about what they are doing.

The authenticity of what the producers know to be false is guaranteed by the other “realities” of the film, principally the narrative. The authors of this textbook are in no doubt that it is to the narrative that color must ultimately be subordinate.

“The objective being to have color ‘act’ with the story, never being a separate entity to compete with or detract from the dramatic content of the picture” (Elements, p. 41).

Such a position is exactly what we should expect. But the book allows, interestingly, for some exceptions to this rule; other values, it seems, may conflict with the necessity of realism. First, there is the value of the star:

“The feminine star, for example, whose appearance is of paramount concern, must be given undisputed priority as to the color of makeup, hair and costume which will best complement her complexion and her figure. If her complexion limits the colors she can wear successfully, this in turn restricts the background colors which will complement her complexion and her costumes to best advantage” (Elements, pp. 40-41).

Thus it is not simply the appearance of the real world (modified to make it “believable”) or the requirements of the narrative which dictate the use of color. The values of stardom must have their place, even if they are in conflict with the dictates of realism (which presumably might demand background colors which did not suit the star.) That the reference is to “feminine” stars alone makes it fairly clear what kind of values are in question here.

But the authors challenge realism most strikingly in their remarks on musical and fantasy pictures. In these genres, it seems, color may escape the demands of realism. It need no longer be subordinate to plot and the appearance of the real world:

“Musicals and fantasy pictures are open to unlimited opportunities in the creative use of color. Here we are not held down by reality, past or present, and our imaginations can soar. Musicals and fantasies are usually designed to provide the eye with visual pleasure in the way that music pleases the ear” (Elements, p. 42).

Thus these genres are privileged. Here the bonds of realism may be slipped and the audience may give itself up to “pleasure. The musical, interestingly, offers another means whereby the dictates of narrative can be avoided, for although musical numbers are often motivated by the plot, they do sometimes succeed in cutting free of narrative altogether and functioning outside it.
Color, then, need not serve realism. It may simply provide pleasure. Yet pleasure in the cinema is never a simple matter. The pleasures cinema offers — the pleasures of realism itself or other kinds — are always within ideology. What ideological forms do the purely visual pleasures of color take? On this point the manual is silent, and we must return to the films themselves.

The ideological appeal of color suggests two possibilities. First, color must signify luxury or spectacle. Whether employed in the western to enhance the beauties of nature, in the costume drama to portray the sumptuousness of the Orient or the Old South, or in musicals to render the dazzle and glamour of showbiz, color serves to embody a world other than our own, into which, for the price of a ticket, we may enter. We should not suppose, of course, that color must always signify luxury or spectacle, since such a signification depends in part upon its scarcity value and even on the mere fact of its costliness. Once color has become normal in the cinema it begins to lose these connotations. One should add, though, that in certain kinds of documentaries and even occasionally in features, black and white is still used as a guarantor of truth, which would not be possible unless their opposite, color, signified something other than truth.

Second, color in early Technicolor pictures operates as a celebration of technology: “Look how marvelous the cinema is!” Color, far from providing a recognizable portrait of the real world, lifts us out of that world, above its mundane problems and unreconcilable contradictions into a new world where the limitations of the old are swept away and its difficulties transcended. (Consider, for example, the relation between the black and white and color sequences in THE WIZARD OF OZ.) Early Technicolor functions as a form of self-reflexiveness, which instead of deconstructing the film and destroying the illusion effects a kind of reification of technology. Other forms of film technology function in the same way: Cinerama, 3-D, even spectacular crane or helicopter shots all having the effect, satirized in the Cole Porter song in SILK STOCKINGS: “glorious Technicolor and breathtaking CinemaScope and stereophonic sound.” In this way, we might see color working to confirm Ernest Mandel’s statement,

“Belief in the omnipotence of technology is the specific form of bourgeois ideology in late capitalism.” [17]

That color can function to signify luxury or celebrate technology does not mean that these two uses of it are necessarily subversive of the dominant cinematic ideology. Not everything which is not realism is counter-cinema. Nevertheless, color clearly did function to an extent as a contradiction of realism. Realism, though dominant, could not provide all the things which were in demand. Realist ideology held out against color first by denying its compatibility with narrative and then by confining it to certain genres. Color, however, was able to satisfy needs which realism could not. Were this not so, it is hard to see how, given its unrealistic connotations, it could ever have been introduced at all. Since the 1930s, however, color has become progressively absorbed back into realism, with the result that the audience’s need for spectacle and for technological wonders has had to be satisfied by a succession of further technological developments: wide screen, 3-D, Sensurround and so on. Even wide screen has now (though in a form less wide than the original CinemaScope) been absorbed into conventional technique. It seems at least possible that a similar fate might have befallen 3-D and other marvels had not they been too expensive for a contracting industry.

Notes

1. Film Reader 2 (1977).
3. Ibid., p. 219.

5. Edwin Mansfield, *Technological Change* (New York: 1971) is a shortened version of his work *The Economics of Technological Change* (New York: 1968), which Gomery cites.


9. Ernest Mandel in *Late Capitalism* (London: 1975), states:

   "The continuous and systematic hunt for technological innovations and the corresponding surplus profits becomes the standard hallmark of late capitalist enterprises and especially of the late capitalist large corporations" (pp. 223-224).

10. Gomery, op. cit.


15. I would not wish to assert that the slow diffusion of color technology was solely due to ideological factors. Undoubtedly there were technical problems, possibly greater than those encountered with sound films. And because color was more expensive, there was an economic rationale for reserving its use for pictures which were expensive in other ways and which could be given special treatment by exhibitors (restricted runs in large urban theatres, etc.). GONE WITH THE WIND would be an example. My main point, however, is that economic factors never exist in isolation, and that in the case of color economics and ideology are mutually reinforcing. See the remarks about luxury and scarcity below.


You wish you never gave it all away. No more to see the setting of the sun. Sound and color. Life in Sound and color.

More on Genius. About "Sound & Color." "Sound & Color" is the title track from the Shakes' sophomore album, shares more than an ampersand with the band's debut "Boys & Girls," which marked the Athens, Alabama group with a "retro-soul" label and launched a sparkling tour that included a stop to play "Memphis hits" for the Obamas at The White House.