This critical analysis of the "cultivation" literature reveals 3 conceptions of the term: (a) George Gerbner's macrosystem explanation of mass media processes and effects, (b) a pattern of operational practices that searches for relationships between television exposure and a wide range of cultivation indicators, and (c) a general forum of explorations of media influence where researchers break away from Gerbner's conceptualizations and boundaries. Using the criteria of heuristic value, empirical support, and precision, this essay evaluates the contribution this large cultivation literature has made to increasing our understanding of the mass media.

Keywords: Cultivation, Theory, Mass Media Effects, Critical Analysis, George Gerbner.

doi:10.1111/jcom.12128

Cultivation as a macrolevel system of explanation about mass media was introduced by George Gerbner (1967, 1969a, 1969b, 1973), who then assembled a research team to help him conduct a series of empirical tests of his system of explanation. By the 1970s other media scholars were attracted to the idea of cultivation and within 4 decades, the cultivation literature grew to well over 500 published studies (Morgan & Shanahan, 2010).

It is the purpose of this essay to provide a critical analysis of cultivation theory and the empirical literature it stimulated by undertaking three tasks. The first task is to analyze the cultivation literature to determine what "cultivation" scholarship means. The second task is to assess the value of cultivation as a system of explanation about the mass media. And the third task is to draw conclusions about the value of cultivation moving into the future.

What does cultivation mean?

While this appears to be a simple question, constructing an answer requires a rather complex process, because the cultivation literature appears to exhibit three different
conceptions of the term. One conception is that cultivation is a mass media theory that was introduced by Gerbner and maintained by him throughout the course of his life. A second conception arises from the pattern of operational practices used by researchers who have published what they presented as tests of various parts of Gerbner’s system of explanation. And a third conception is exhibited by researchers who operate within a general socialization perspective and who largely ignore the conceptualizations of Gerbner as they explore a variety of ways that media exert their influence on individuals.

Cultivation as a theory
When Gerbner introduced his conceptualization of cultivation (Gerbner, 1967, 1969a, 1969b, 1973), he positioned his explanation in contrast to what he considered the dominant form of mass media research at the time, which he characterized as a literature of microfocused studies typically using experiments to determine if differences in particular features of media messages could explain immediate responses in knowledge, attitudes, and behaviors. As an alternative to this type of research, Gerbner presented a macrosystems approach. Gerbner was not interested in the influence of particular message elements; instead he was interested in widespread meanings across the entire media landscape. He was not concerned with how people selected messages for exposure, nor how they processed the information in those messages, nor what effects particular messages exerted on people during exposures or immediately after; instead, he was exclusively concerned with the influence that a much broader scope of messages gradually exerted on the public as people were exposed to media messages in their everyday lives.

Gerbner did not limit his theory to only effects but was also concerned with media institutions and messages in addition to those effects. His systems approach highlighted the interplay of influence across three components: the media institutions, the mass-produced messages, and their cultivated effect on large aggregates. As for the institutional component, Gerbner argued that the “mass production and rapid distribution of messages create new symbolic environments that reflect the structure and functions of the institutions that transmit them” (Gerbner, 1970, p. 69). As for the message component, Gerbner claimed that there were certain mass-produced meanings that were widespread throughout the entire mass media environment. And as for the effects component, Gerbner contended that the widespread meanings presented across all media cultivated public beliefs. Those mass-produced messages form “a common culture through which communities cultivate shared and public notions about facts, values, and contingencies of human existence” (Gerbner, 1969b, p. 123).

To guide researchers in their tests of the major claims he made in his system of explanation, Gerbner outlined three types of analysis: institutional analysis, message system analysis, and cultivation analysis. With institutional analysis, Gerbner directed researchers to look for “changes in the mass production and rapid distribution of messages across previous barriers of time, space, and social grouping” to determine
how those changes "bring about systematic variations in public message content" (Gerbner, 1969b, p. 124).

With message system analysis, Gerbner recommended that the search for these widespread meanings should be undertaken in a scientific manner. "What distinguishes the analysis of public mass-mediated message systems as a social scientific enterprise from other types of observation commentary, or criticism is the attempt to deal comprehensively, systematically, and generally rather than specifically and selectively or ad hoc with problems of collective cultural life" (1969a, p. 141). He also cautioned that content analysts should look for patterns of meaning across the total media landscape and therefore make no distinction between information versus entertainment, fact versus fiction, high culture versus low culture, good versus bad, images versus words, or levels of artistic excellence. Thus the purpose of message system analysis was to identify meanings that (a) could be attributed to institutional practices of mass production of messages, and (b) were found to be disseminated widely across the entire media landscape.

Cultivation analysis, Gerbner (1973) argued, "begins with the insights of the study of institutions and the message systems they produce, and goes on to investigate the contributions that these systems and their symbolic functions make to the cultivation of assumptions about life and the world" (p. 567). Because the mass-produced messages form the culture, Gerbner believed that their influence on the public should be exhibited over the long term by reinforcement as well as change. "The dynamics of continuities, rather than only of change, need to be considered in the examination of mass-produced message systems and their symbolic functions. Such examination is necessarily longitudinal and comparative in its analysis of the processes and consequences of institutionalized public acculturation" (p. 569).

In summary, the key characteristics of Gerbner’s theory were that it focused attention at the macrolevel of broad scale institutional practices, widespread meaning, and long-term acculturation. It was naturalistic, that is, it did not manipulate exposures or messages but instead acknowledged individuals’ typical patterns of media exposure in their everyday lives. And it took a systems approach by emphasizing the importance of tracing how the media’s institutional practices shaped meanings in the mass production of messages that were then widely disseminated and thereby shaped public knowledge and beliefs over the long term.

**Tests of cultivation theory**

A second type of conception of cultivation can be inferred from the pattern of practices used by researchers who have operationalized one or more of Gerbner's theoretical claims in order to test them. This set of operational practices was created by Gerbner working with his closest colleagues beginning in the later 1960s. When other media scholars began conducting their own tests of cultivation, they typically followed this set of operational practices.
The operational practices established by Gerbner and his team exhibit seven characteristics. The first two of these characteristics refer to the patterns of practices within the message system analysis literature. These are the practices of limiting their sampling frame to TV messages and the equating of frequency with meaning.

In their initial message system analysis studies, Gerbner and his research team designed a series of content analyses that narrowed the focus down from all mass media to the single medium of television. To justify this reduction, they argued that “commercial television, unlike other media, presents an organically composed total world of interrelated stories (both drama and news) produced to the same set of market specifications.” They further limited the scope of their content analysis by examining only the programs presented on primetime and children’s weekend morning programming provided by the three dominant commercial television networks at the time, reasoning that these programs had the largest viewership and if there were particular meanings widespread across the entire media landscape, those meanings should be in evidence in their sample. They further limited their focus by examining only entertainment programming but provided no reasoning to support this decision (e.g., Gerbner et al., 1977). While they gradually expanded their content analyses from looking at the topic of violence to looking also at other topics such as gender roles, marriage and family, aging, race/ethnicity, occupational status, affluence, the environment, politics, and mental illness (Morgan & Shanahan, 2010), they consistently stayed with the same sampling frame of mainstream television entertainment.

The second pattern of operational practices in message system analysis was the using of frequency counts as a surrogate for meaning. For example, when they found a high frequency of occurrence of violent acts from their yearly content analyses, they concluded that the television world was mean and violent (e.g., Gerbner et al., 1977).

When operationalizing cultivation analysis, the cultivation team made a series of design decisions that reveal five key features: use of national probability samples, cross-sectional surveys, assumption of stable TV viewing, focus on beliefs as a cultivation indicator, and categorical analyses of the relationship between TV exposure and cultivation indicators.

The cultivation team typically used data gathered by a commercial polling firm that each year conducted a national telephone survey of households randomly sampled so that it represented the adult population of the United States. The survey design was cross-sectional where respondents were measured once, rather than longitudinal where the same respondents would have been measured year after year to determine if there were changes in respondents’ TV exposure, their cultivated beliefs, or the relationship between the two.

Television exposure was measured by asking respondents how many hours of television they watched on an average day last week. This reveals their assumption of habitual viewing, that is, that people watch TV habitually and that their exposure
patterns varied little from day to day, week to week, or year to year. They routinely divided these continuous distributions of hours viewed into three categories: light viewing, moderate viewing, and heavy viewing.

Cultivation indicators were measured by presenting respondents with two choices of beliefs (one reflecting the television world and the other reflecting the real world) and asking them which of the two they believed to be more accurate. “For example, one cultivation question asks: ‘During any given week, what are your chances of being involved in some kind of violence? About one in ten? About one in a hundred?’” (Gerbner, Gross, Jackson-Beeck, Jeffries-Fox, & Signorielli, 1978, p. 195). While one of these answers is closer to the television world figure and the other is closer to the real-world figure, neither answer is accurate. The providing of inaccurate answers was an operational decision that was motivated by their belief that the accurate answers were too extreme and would therefore look like trick questions (Shanahan & Morgan, 1999, pp. 53–54).

When computing the relationship between television exposure and cultivation indicators, they used two procedures to look for evidence of cultivation effect. One procedure was to compute the degree of relationship between viewing level and selection of an answer on the cultivation indicator (TV answer vs. real-world answer). Thus the larger the coefficient, the stronger the evidence for heavy viewers picking the TV world answer compared to the lighter viewers picking that answer. The other procedure was to calculate what they call a cultivation differential, which was the difference between the percentage of respondents in the heavy viewing group who selected the TV world answer compared to the percentage of respondents in the light viewing group who selected the TV world answer. Thus when the cultivation differential was positive, it indicated that a larger percentage of the heavy users compared to the light users had selected the TV world answer.

While the cultivation team was consistent in staying within the boundaries of Gerbner’s conceptualizations and maintaining the same operational practices, they did make one major addition to the theory in 1980 by adding the constructs of mainstreaming and resonance. Gerbner, Gross, Morgan, and Signorielli (1980) explained “the ‘mainstream’ can be thought of as a relative commonality of outlooks that television tends to cultivate. By ‘mainstreaming’ we mean the sharing of that commonality among heavy viewers in those demographic groups whose light users hold divergent views” (p. 15). Resonance is characterized as a “double dose” of meaning that comes from both real-world experience as well as TV messages. To test for a resonance effect, researchers look for “a feature of the television world (that) has special salience for a group” (p. 23). For example, people who live in dangerous high-crime neighborhoods and watch a lot of television get a double dose of exposure to crime. Therefore, these people should exhibit a higher than average degree of cultivation—a resonance effect—from the television exposure because those TV exposures resonate with their real-world environment.
Other media effects researchers

By the late 1970s, the cultivation system of explanation started attracting the attention of media effects scholars who essentially used the cultivation team’s operational practices and ran their own tests of the theory (e.g., Christiansen, 1979; Fox & Philliber, 1978). As this literature of tests of cultivation theory grew, there were examples of researchers who tested some extensions to the established operational practices, particularly with cultivation indicator measures. These studies appeared to accept Gerbner’s original conceptual definition for a cultural indicator then sought to test alternative measures in order to fill in the gaps across the range of meaning laid out by that conceptual definition. One example of this is the development of cultivation indicator measures on additional topics. For example, these researchers added to the variety of cultural indicators by testing people’s beliefs about mental illness (Diefenbach & West, 2007), beliefs about substance abuse (Minnebo & Eggermont, 2007), acceptance of homosexuality (Calzo & Ward, 2009), and beliefs about the environment (Holbert, Kwak, & Shah, 2003), to name a few of these extensions of topics.

Another line of research that expanded the operationalizations of cultivation indicators examined different types of measures (knowledge, attitudes, beliefs, emotions, and behaviors). For example, Hawkins and Pingree (1980) tested different types of cultivation indicators to see if knowledge (where respondents were asked to provide estimates of the prevalence of occurrences such as crime) was related to beliefs. Hawkins and Pingree (1980) and others (Gerbner, Gross, Morgan, & Signorielli, 1986; Potter, 1991; Sparks & Ogles, 1990) found that the different types of measures were not related to each other very strongly. In reviewing this research, Shanahan and Morgan (1999) argued that when testing for cultivation indicators within the topic of violence researchers should consider at least three types of measures — knowledge (perceptions of the amount of violence in society), beliefs (about one’s likelihood of being victimized by violence), and emotion (fear of being victimized). Furthermore, they speculated that there are perhaps even more than three types of measures that would serve as valid operationalizations of cultivation indicator as originally conceptualized by Gerbner.

There has also been a pattern of some media effects scholars who essentially accept cultivation theory but are driven by a desire to see if another variable or two could increase cultivation’s predictive power over and above that of using only TV viewing. These tests have added variables such as perceived reality (Busselle, Ryabalova, & Wilson, 2004; Potter, 1986), transportation (Bilandzic & Busselle, 2008), and distance (Bilandzic, 2006; Hetsroni, Elpariach, Kapuza, & Tsfoni, 2007; Van den Bulck, 2003). Also, other researchers have tried to increase the explanatory power of cultivation by incorporating constructs and propositions from other theories, such as knowledge gap theory (Niederdeppe, Fowler, Goldstein, & Pribble, 2010), theory of reasoned action (Beullens, Roe, & Van den Bulck, 2011; Nabi & Sullivan, 2001), spiral of silence (Shanahan, Scheufele, Yang, & Hizi, 2004), elaboration likelihood model (Schroeder, 2005; Williams, 2006), and mental models (Roskos-Ewoldsen, Davies, & Roskos-Ewoldsen, 2004).
Boundary-crossing alterations

In contrast to the literature described above where scholars accept the claims of cultivation theory and attempt to add something to it, there is a growing number of studies where scholars reject at least one of Gerbner’s most fundamental claims and replace them with their own claims. Thus these studies have moved outside the boundaries of cultivation theory as established by Gerbner, particularly in two ways. One of these boundary-crossing trends has been a move away from a macro focus into a micro focus and the second has been a move away from regarding the locus of meaning in the media messages toward regarding the locus of meaning in receivers.

Movement into micro

One of the core characteristics of cultivation theory is its macrolevel focus. While Gerbner did not reject the value of microlevel research, he declared that it was outside the boundaries of his theory. In all his publications, he maintained this boundary, saying, "the comparison of responses of those who claim to prefer or view this and that type of programming (one particular genre), instead of measures of total viewing, is likely to yield confusing, contradictory, and misleading results" (Gerbner, 1990, p. 257). And in 2002, he wrote, "We do not minimize the importance of specific programs, selective attention and perception, specifically targeted communications, individual and group differences, and research on individual attitude and behavior change. But giving primary attention to those aspects and terms of traditional media effects research risks losing sight of what is most distinctive and significant about television as the common storyteller of our age" (Gerbner, Gross, Morgan, Signorielli, & Shanahan, 2002, p. 44).

Despite this clear boundary for cultivation analysis, many researchers have moved away from using total TV viewing as a predictor of cultivation indicators and moved toward testing more micromeasures of exposure. For example, some researchers argued that evidence for a belief in a mean and violent world should be more attributable to exposure to violent programming more than to total TV exposure, and their empirical tests generally confirmed this expectation as genre level exposures (crime drama and news) were found to be stronger predictors of cultivation than total TV viewing exposure (Grabe & Drew, 2007; Hawkins & Pingree, 1980; Potter, 1988). Other researchers testing additional cultivation indicator topics also found that when they used more specific level television exposure measures, they were able to predict respondents’ beliefs better. These studies have tested exposure to talk shows (Glynn, Huge, Reineke, Hardy, & Shanahan, 2007), make-over programs (Kubic & Chory, 2007; Nabi, 2009), romantic shows (Segrin & Nabi, 2002), medical dramas (Chory-Assad & Tamborini, 2003; Van den Bulck, 2002), and local news programming (Romer, Jamieson, & Aday, 2003). Also, some studies claimed to test a cultivation effect from exposure to a single TV series such as Grey’s Anatomy (Quick, 2009) or from playing one particular video game (Williams, 2006).

Reviewers of the cultivation literature seem to be conflicted about whether to include these studies as being cultivation or not (Morgan & Shanahan, 2010; Morgan,
Cultivation

W. J. Potter

Shanahan, & Signorielli, 2014; Shanahan & Morgan, 1999). For example, Morgan and Shanahan (2010) raised the question about when "genre-specific effects should be called 'cultivation'" (p. 340) and answered the question with a caution that "such work can fragment the systemic aspects of the overall viewing experience, and observed relationships may reflect selective exposure more than cultivation" (p. 341). But they then went ahead and reported this literature anyway, labeling it "Genre-Specific Cultivation." And more recently they still claimed "The notion that media genre is a critical component in assessing the phenomenon of cultivation goes against one of the basic tenets or assumptions of cultivation" (Morgan et al., 2014, p. 491).

Another indicator of this shift from macro to micro is the growing number of studies that take a short-term focus on effects, such as studies that use an experiment to present participants carefully constructed message elements and collect outcome data immediately after exposure to those controlled treatments (e.g., Bilandzic & Busselle, 2008; Williams, 2006). Although Gerbner objected to experiments as providing adequate tests of cultivation, reviewers of this literature exhibit acceptance of findings of experiments along with findings from surveys as evidence of cultivation (Morgan & Shanahan, 2010; Morgan, Shanahan, & Signorielli, 2009; Morgan et al., 2014; Romer, Jamieson, Bleakley, & Jamieson, 2014; Shanahan & Morgan, 1999). For example, Morgan and Shanahan (2010) said, "Cultivation clearly construes messages as systems, which are by definition as macro as one can get. Some other approaches tend to reduce messages to components, especially if they can be easily manipulated in laboratory settings. Both of these conceptions survived quite nicely side-by-side through the decades. In an era of mass communication, macrolevel conceptions may seem to have the most explanatory power. In an era of mediated interpersonal communication and fragmented audiences, such macrolevel conceptions may seem less relevant" (p. 351).

Shift of meaning locus from messages to receivers

A second kind of boundary-crossing research has been a shift in the locus of meaning. In introducing cultivation, Gerbner placed the locus of meaning exclusively in the media messages and regarded research that examined how receivers interpreted meaning as being out of bounds for cultivation theory. He argued that he was not concerned with how "different individual and group selections and interpretations of messages take place." He continued, "Whether I accept its 'meaning' or not, like it or not, or agree or disagree is another problem. First I must attend to and grasp what it is about" (1969b, p. 125). Thus while Gerbner recognized that there were individual differences in interpretations of messages, cultivation was not concerned about those variations in interpretations; instead, cultivation focused on the dominant meanings that the media presented to the public. Thus, Gerbner regarded message system analysis as an essential antecedent to cultivation analysis; a clear understanding of the dominant meanings in media messages was an essential first step in the process of constructing cultivation indicator items. Gerbner argued, "Survey questioning used in cultivation analysis should reflect the over-arching content configurations
embedded in television's message systems. . . . Helter-skelter and exploratory questioning may be useful for a variety of theoretical and serendipitous purposes but do not test cultivation theory" (Gerbner, 1990, p. 257).

There seems to be a growing literature labeled as cultivation that rejects Gerbner's perspective on meaning and instead regards the locus of meaning as residing in receivers (e.g., Bilandzic & Busselle, 2008; Busselle, 2003; Martins & Harrison, 2012; Nabi, 2009). Thus these researchers employ a reasoning process that moves away from what Gerbner laid out in his theory, which requires the conduct of message system analysis first to identify widespread meanings, then using these meanings to guide the construction of cultivation indicators. In contrast, this new line of "cultivation" research first constructs a variety of cultivation indicators, then tests which indicators are most strongly related to heavy TV viewing (usually of a genre rather than all TV viewing) then presents the stronger relationships as evidence for a cultivation effect of viewing TV. Furthermore, reviews of the cultivation literature (Morgan & Shanahan, 2010; Morgan et al., 2009, 2014) now acknowledge these studies as tests of cultivation without working through the implications of this type of research for cultivation theory.

Another indicator of this shift in meaning locus within the "cultivation" literature is the growing number of studies that take a cognitive approach in examining how people go about remembering media messages and use those memories to construct their responses to cultivation indicator items (e.g., Shrum, 2004). In reviewing this literature Shanahan and Morgan (1999) cautioned, "we believe that to become sidetracked by the peculiarities of how individuals receive, process, interpret, remember, and act on messages can distract attention from the more central questions of cultivation research" (p. 172). However, they devoted an entire chapter to reviewing the literature.

There appears to be a trend toward this boundary-crossing research. To illustrate, a search of the Communication Abstracts database of scholarly publications from 2010 to 2014 using keywords of cultivation theory and its synonyms reveals 37 articles reporting empirical tests. Five of those are content analyses, and the remaining 32 are effects studies. Of these 32, 13 do not provide tests of any of Gerbner’s claims; instead they mention cultivation theory as a possible benefactor of their findings without linking those findings to a specific claim made in that theory. The remaining 19 publications position themselves as tests of cultivation theory in their introductions, review of literatures, and rationales. Of these 19, only 8 present a foundation for their cultivation indicators in a message system analysis (either where the authors provided the results of their own message system analysis or cited relevant findings from a content analysis), while the other 11 reason backwards from respondents’ selection of their answers on a cultivation indicator as being the meaning presented in the media. Also, of the 19 tests of cultivation, only 5 used a macromeasure of TV exposure with the remaining 14 using only microlevel exposure measures (e.g., exposure to genres or particular TV programs). Thus of the 32 published articles referencing cultivation
published in the most recent 5-year period, only 3 stayed within the boundaries of cultivation theory by using both a macromeasure of exposure and a meaning analysis as a foundation for the cultivation indicators. Thus there is a growing number of published studies labeling themselves as cultivation research that reject one or more of the most central ideas in Gerbner’s theory. This clearly indicates that many media effects researchers have a different conception of cultivation than that presented by Gerbner.

**Assessing the value of cultivation as a theory explaining mass media**

How well has cultivation theory and the literature it stimulated increased our understanding of the media? The answer to this question is structured by the application of three criteria — heuristic value, empirical support, and precision.

**Heuristic value**

The criterion of heuristic value is concerned with the degree to which a theory has attracted the attention of scholars and has stimulated those scholars to conduct tests of the theory’s claims. Cultivation at first glance seems to have done well on this criterion when we look at the size of the literature it has generated. But when we examine the patterns in that literature for completeness and relevance, the heuristic value of cultivation theory appears less positive.

**Size**

From the 1970s until today, the cultivation system of explanation has been attracting the attention of media effects scholars, many of whom have published studies that they label as cultivation. "As of 2010, over 500 studies directly relevant to cultivation have been published — and more than 125 since 2000" (Morgan & Shanahan, 2010, p. 337). Also, several content analyses of the published literature in communication have identified cultivation as one of the most visible media effects theories (Bryant & Miron, 2004; Potter & Riddle, 2007).

**Completeness**

If we are to give credit to a theory that has stimulated a large literature of empirical tests, we also need to hold the theory accountable for how well it has stimulated tests across the full range of its claims. When we divide the cultivation literature into tests of claims for Gerbner’s three components, we can see that the literature exhibits a heavy concentration in cultivation analysis while institutional analysis has been ignored. As for institutional analysis, there is no evidence that Gerbner or his team published any research to test their claim that “the mass production and rapid distribution of messages create new symbolic environments that reflect the structure and functions of the institutions that transmit them” (Gerbner, 1970, p. 69). While reviewers of the cultivation literature (Morgan, 2009; Morgan & Shanahan, 2010; Morgan et al., 2009; Shanahan & Morgan, 1999) continually acknowledge the
importance of institutional analysis in the cultivation system of explanation, they rarely cite any research in this area, although there is a fairly well developed literature on this very topic (for reviews of this literature, see Grossberg, Wartella, & Whitney, 1998; McQuail, 2005). Furthermore, the scholars who produced this literature have displayed little interest in linking their findings with cultivation theory.

Even more troubling on this criterion of completeness are the gaps in the research within message system analysis where that literature has yet to provide convincing evidence that there are meanings widespread across the television landscape, much less all mass media. While there are many studies that have examined content in mainstream commercial television, there are a few studies that have looked at other kinds of television content such as news, informational programming, and especially advertising. This omission is especially glaring given Gerbner’s argument that television has such a strong commercial interest so that we should expect those commercial values to show up strongest and most consistently in the great amount of advertising messages they present. Despite this lack of an adequate research basis, Gerbner et al. (2002) continued to make claims that television’s "drama, commercials, news, and other programs bring a relatively coherent system of images and messages into every home" (p. 44). And this claim is still being made today that multinational media conglomerates “dominate the cultural symbolic environment with stable and consistent messages about life and society . . . despite the emergence of so many specialized new channels and so many different types of programs that are often targeted to smaller and smaller audiences” (Morgan et al., 2014, p. 481).

Furthermore, there has been a lack of attention to synthesizing findings across studies in the existing message system analysis literature. Although there are a number of reviews of this literature (Morgan, 2009; Morgan & Shanahan, 2010; Morgan et al., 2009, 2014; Romer et al., 2014; Shanahan & Morgan, 1999), these reviews are largely descriptive inventories of a growing list of topics, rather than a careful sorting through the findings to identify the meanings that have been found to have limited scope (and where those limits are) and the meanings that have been found consistently across all types of messages and throughout all media. This lack of synthesis work is especially troubling because Gerbner himself argued for the importance of synthesizing findings across content analyses when he criticized the literature of media content analysis at the time he introduced cultivation as being “piecemeal, sporadic, uncoordinated, and rarely comparable over time and across cultures” (1973, p. 557). Now more than 4 decades later, there is a much larger literature documenting patterns of content in the media; however, this literature is still fragmented because it continues to lack the synthesis needed to coordinate findings across genres, dayparts, channels, media, and cultures. Without such a synthesis, there is no basis to make a claim that meanings are widespread across media. While it is not a problem for a theory to present such a claim as an invitation for researchers to test it, it does become a problem when decades of research fail to fill in the gaps to an extent that we can believe that this claim has been adequately tested and found to have consistent support to establish its credibility.
Relevance
If we use a strict criterion of relevance and include only those studies that provide tests of Gerbner’s claims, then the figure of 500 published tests is likely to be greatly reduced. As for the message system analysis literature, the stricter criteria would require content analyses of media messages to construct their arguments for meanings from the web of context in the presentations rather than simply the frequency of occurrence of manifest elements. Also, the content analysis of meaning would need to focus on message characteristics that can be attributed to production practices as documented through institutional analysis. It appears that almost none of the published studies in the message system analysis literature have met both these criteria.

As for the cultivation analysis literature, the stricter criteria would require that at minimum the cultivation indicator measures be constructed from the findings of a meaning analysis, as was pointed out above. Also, the stricter criteria would call for the use of macromeasures of exposure over the long term and the use of longitudinal designs to examine the long-term influence of exposure in order to be considered an adequate test of Gerbner’s claim about the influence of widespread meanings that are exhibited in cultivation indicators as a result of people’s everyday exposures over the long term. Recall from the description of the boundary-crossing literature above, there is a trend away from testing the claims in Gerbner’s theory although the authors of these studies still call them “cultivation.”

This raises a question about boundaries, that is, how far can a research study deviate from Gerbner’s conceptualizations and still be regarded as a "cultivation" study? If any kind of content analysis of any set of media messages and any media effect study warrants the label, then the scope is so broad that the “cultivation” label would seem to have become meaningless.

Empirical support
The pattern of findings from studies labeled as tests of the cultivation system of explanation shows that support for Gerbner’s claims is null (institutional analysis), partial (message system analysis), and weak (cultivation analysis).

Support for the existence of widespread meanings across the television landscape is partial, because almost all of the message system analysis research is concentrated on looking for patterns within the narrow sampling frame of entertainment programming, which has been further reduced to focus on mainstream channels and limited dayparts. While there are message system analyses of news and informational programming (e.g., Lee & Niederdeppe, 2011; Romer et al., 2003), the findings from these studies have not been integrated with findings from other types of programming in a complete enough fashion to make the case for widespread meanings.

As for cultivation analysis, the evidence that supports cultivation theory’s explanation of television’s influence on public beliefs has been persistently weak. This was pointed out by early critics of the theory (Hirsch, 1980; Hughes, 1980), and despite the huge growth in the literature since that time, the low ceiling of predictive power has not been raised. A meta-analysis of 5,799 separate findings derived from
W. J. Potter

Cultivation

97 studies/samples of tests of cultivation analysis reports that the average correlation across all those findings is .10 and that the average partial correlation is .09 (Shanahan & Morgan, 1999). This means television exposure predicts only about 1% of the variation in cultivation indicators.

Gerbner was aware of these weak relationships and defended his system of explanation by saying that cultivation’s effect was cumulative over time. However, this defense is faulty because there is very little evidence that cultivation’s influence is cumulative; to the contrary, the evidence suggests that the cultivation effect is not cumulative. In their meta-analysis, Shanahan and Morgan (1999) report that cultivation differentials are lower for older respondents than for younger ones. And longitudinal tests have either found no support (Morgan, 1982) or equivocal support (Morgan, 1987) for a cumulative influence. Although the findings of these two longitudinal studies challenge the validity of the cultivation explanation of media effects, cultivation researchers have let those findings stand unaltered for over 2 decades by not conducting additional longitudinal studies.

Furthermore, it appears that the cultivation effect is not widespread, but that it shows up in only a smaller percentage of the population. When examining the extent of evidence for cultivation in typical cultivation analysis studies, sizable percentages of respondents—even in the heavy viewing groups—do not choose the TV world cultivation indicator. For example, Hetsroni and Tukachinsky (2006) found that only 13% to 28% of their respondents showed evidence of a cultivation effect depending on the topic.

The empirical support of cultivation appears especially weak when we compare it to the level of support generated by other media effects theories. For example, agenda-setting theory is arguably the one media effects theory closest to cultivation in terms of its examination of patterns of meaning in media messages and the influence of those message patterns on large aggregates of people in the course of their everyday lives. Unlike cultivation theory, agenda-setting theory has been shown to have a relatively strong predictive power \( r = .53 \) in a meta-analysis of its empirical literature (Wanta & Ghanem, 2007). Agenda-setting researchers have also added constructs and propositions to the original conceptualization of the theory and have incorporated useful constructs from other media effects theories such as framing and priming (McCombs & Reynolds, 2009; Shah, McLeod, Gotlieb, & Lee, 2009).

The strength of findings in the cultivation analysis literature also appear relatively weak when we compare them to the strength of findings determined by meta-analyses of other media effects such as the third-person effect \( r = .500; \) Paul, Salwen, & Dupagne, 2000); inoculation effect \( r = .430; \) Banas & Rains, 2010); hostile media effect \( r = .296; \) Hansen & Hyunjung, 2011); effect of listening to popular music on mood and attitudes \( r = .210; \) Timmerman et al., 2008); influence of sexually explicit materials on physiological and psychological reactions \( r = .212 \) to .248; Allen et al., 2007); and engagement with media entertainment on knowledge, attitudes, and behaviors \( r = .270; \) Tukachinsky & Tokunaga, 2013).
As for the topic of the effects of violence, which has been the most popular cultural indicator, cultivation analysis has been much weaker at predicting effects than have other systems of explanation. For example, Paik and Comstock (1994) conducted a meta-analysis of 217 studies examining the effect of exposure to television violence on antisocial behavior and report that the average effect size is \( r = .31 \). Also, Anderson et al. (2010) conducted a meta-analysis on 130 research reports that presented over 380 effect-size estimates based on over 130,000 participants and found an average effect size of \( r = .217 \) on experiments testing the effect of playing violent video games on aggressive behavior. They also found an average effect size of \( r = .183 \) from cross-sectional surveys.

When researchers conduct tests to compare the explanatory power of cultivation with another media effects theory typically find cultivation to be the weaker explanation. For example, Diefenbach and West (2007) compared cultivation with the third-person effect and found the third-person effect to be a stronger explanation. Gross and Aday (2003) found that local news exposure accounted for an agenda-setting effect but did not cultivate fear of being a victim of crime. Martins and Harrison (2012) found that social identity theory was a stronger predictor of preadolescents’ global self-esteem than was cultivation theory.

**Precision**

Precision refers to the theory’s ability to articulate its assumptions and propositions clearly enough so that scholars can understand its explanation and so that researchers have the guidance they need to design adequate tests. Theoreticians who exhibit a high degree of precision recognize where their system of explanation is ambiguous or inconsistent, then work to make additions and alterations to increase the clarity of their explanations.

Cultivation theory has continued to exhibit a troubling degree of ambiguity in all three areas of its system of explanation. As for institutional analysis, Gerbner continued to make his argument that television messages are mass-produced “to market specifications of industrial organizations for commodity and political markets” (Gerbner, 1990, p. 251). However, he did not develop his arguments in enough detail to guide researchers to identify what these market specifications might be nor how those market specifications are revealed in the meaning embedded in media messages.

As for cultivation analysis, Gerbner was initially unclear about how the meanings in media messages created or shaped the public’s beliefs. In using a socialization perspective, it appeared that Gerbner conceptualized media influence as exerting a gradual but continuing influence. This is apparent in his use of metaphors like a glacier. Also, Morgan and Signorielli (1990) argued that cultivation focuses on "long-term, cumulative consequences of exposure to an essentially repetitive and stable system of messages" (p. 18). The idea of accumulating influence is also in evidence when reviewers dismiss the persistently weak results in empirical tests by saying that while media influence appears weak at any given point in time, its influence is constant and therefore accumulates into something more substantial.
over time (Morgan et al., 2014; Shanahan & Morgan, 1999). However, findings from cultivation analyses do not support this assertion. To the contrary, the meta-analyses of Shanahan and Morgan (1999) show that elderly people exhibit less evidence of cultivation than do younger people, yet they still make the claim for a cumulative effect (Morgan & Shanahan, 2010; Morgan et al., 2014) and repeatedly show no recognition of this inconsistency between their claim and the findings of the research they review.

As for message system analysis, the problem with precision can be traced more to practices in the tests of the theory rather than the theory itself. Gerbner provided a fair amount of detail to specify how message analysts should begin the task of documenting widespread meanings. For example, Gerbner (1973) was clear that the meaning in messages typically does not reveal itself in a single symbol but that meaning lies in the relationship among the symbols in the messages. He cautioned that “the characteristics of a message system are not necessarily the characteristics of individual units composing the system. The purpose of the study of a system as a system is to reveal features, processes, and relationships expressed in the whole, not in its parts” (Gerbner, 1969a, p. 143). Furthermore, Gerbner (1990) directed message system analysts to look deeper than surface features such as plot configurations and look instead at "the underlying uniformity of the basic 'building blocks' of the television world: thematic structure, interaction patterns, social typing and fate," which lies in the consequences (p. 255). And to help guide message system analysts in this challenging task, Gerbner (1969a) laid out four questions: “(1) ‘what is’ (i.e., what exists as an item of public knowledge), (2) ‘what is important’ (i.e., how the items are ordered), (3) ‘what is right’ (or wrong, or endowed with any qualities, or presented from any point of view), and (4) ‘what is related to what’ (by proximity or other connection)” (1969a, p. 144).

However, the message system analysis literature shows a pattern of counting the occurrence of manifest elements (such as acts of violence, gender of characters) and presenting those frequency counts as evidence of meaning, which is rather like trying to determine the meanings in a novel by counting the occurrence of different words. While this equating frequency and meaning can often work well in constructing an answer to the first question of “what is,” it is far too simple a method for constructing credible answers to the remaining three questions. Constructing an answer to the second question requires analysts to consider the position of information within messages. The answer to the third question requires analysts to recognize key contextual factors then assemble those factors in a way to reveal the meaning of an act or occurrence. And the answer to the fourth question requires analysts to look for patterns across messages.

When we look at the cultivation team’s annual profiles of television content, the reported findings are based solely on arguments about the frequency of certain manifested elements, without considering patterns of contextual factors. While it is a challenging task to examine the web of context within which violence (or any other manifest occurrences) is presented, there are some examples of content analysis studies that have been able to move beyond the simple reporting of frequencies of acts and attempt to engage the more fundamental issue of the meaning that arises from
the way those acts are portrayed. One example is the National Television Violence Study (1999), which examined 43 contextual variables and used patterns among those variables to make claims about meaning in the way violence was portrayed.

This continued lack of precision is troubling because Gerbner (1969b) cautioned researchers that "I must stress that the characteristics of a message system are not necessarily the characteristics of individual units composing the system. The purpose of the study of a system as system is to reveal features, processes, and relationships expressed in the whole, not in its parts" (p. 128).

A future for cultivation?

As we move into the future, what is the value of cultivation for generating insights that will increase our understanding of media processes and effects? Because there are three conceptions of cultivation, the answer to this question needs to be approached through three avenues — theory, empirical tests of the theory, and extensions.

As for moving forward with the theory as laid out by Gerbner, we need to be concerned about the kind of research it has stimulated. While the theory has clearly generated a large number of apparent tests of its claims, those tests have ignored institutional analysis; ignored the systems nature of Gerbner’s explanation; failed to generate much message system analysis that moved beyond simple counts to engage the more challenging task of determining meaning; failed to generate more than a few cultivation analysis tests that moved beyond cross-sectional surveys to engage the more challenging task of examining the role of time; and failed to generate reviews of the literature that moved beyond descriptions of findings to engage the more challenging task of synthesizing patterns about the construction of meaning, the degree to which those industrial produced meanings are widespread, the extent to which the population is cultivated by those meanings, and the shape of those meanings’ influence over time.

Cultivation theory, however, might still have potential as a viable system of explanation of media effects moving into the future if it can deliver a higher level of precision in articulating its core ideas and if those articulations stimulate researchers to test those core ideas. The theory needs to be a better guide to help researchers examine the linkages between institutions, messages, and effects; to guide researchers through the task of establishing claims for meaning based much more on the contextual web instead of simple frequency counts; to show the meanings that are widespread through significant syntheses of the message system literature; to provide a clear conceptualization for how widespread patterns of meaning first shape beliefs then reinforce them over time; and to incorporate additional constructs to increase its power in explaining media influence. Furthermore, it needs to do all this while recognizing the challenges imposed by the new media environment rather than to continue denying that media message production is now highly fragmented, that meanings are rapidly changing rather than stable, and that, while there may still be some widespread meanings, the meaning environment is likely to be dominated by layers of specialized meanings keyed to genres and subgenres.
The empirical literature of established operational practices has produced a very useful finding: Amount of television viewing is consistently related to all kinds of cultivation indicators across all kinds of people and across a wide variety of cultures and countries. However, the degree of that relationship is consistently weak. Because we already have a large enough literature to know that this finding is robust, another study that uses these established operational practices would have very low marginal utility. Instead, if this literature is to increase in value moving into the future, cultivation researchers need to move beyond "picking the low-hanging fruit" of using counts of manifest elements in messages as surrogates for meaning and of using cross-sectional surveys to establish claims for accumulating media influence. Cultivation researchers need to engage the higher-level challenges of identifying institutional practices that shape messages; digging below the surface characteristics of messages to construct convincing claims for meaning; examining messages much more systematically across the full media landscape to construct convincing claims for widespread meanings; and following people's belief patterns over time to plot the occurrence of when a person's belief shifts from a real-world basis to a TV world basis then continues to be reinforced over time. If researchers are unwilling or unable to engage these challenges, then there is little reason to expect that the growth in this literature will deliver anything beyond tests of more topics for cultivation indicators that will show a weak relationship with TV exposure.

It is highly likely that the growing boundary-crossing literature will be able to deliver valuable insights about media messages, processes, and effects. This literature has already delivered interesting insights about specialized meanings keyed to more microunits within the overall media landscape and on immediate effects on individuals who are exposed to various specific content patterns. However, while this research can make valuable contributions to our understanding about media processes and effects, it is puzzling how this literature increases our understanding of cultivation. Furthermore, the labeling of these studies as cultivation is more likely to lead to ambiguity about what cultivation means. If the authors of these studies removed the "cultivation" label from them, those contributions would lose none of their value, and they would increase in clarity by eliminating the confusion generated by expecting to see a test of Gerbner's ideas. Also, if these studies acknowledged their closer affinity to other theories (especially agenda-setting theory, social cognitive theory, knowledge gap, elaboration likelihood, and heuristic processing), they could contextualize their findings into more elaborate and more powerful systems of explanation, rather than trying to fit them into the increasingly amorphous cultivation literature.

Cultivation theory has delivered value in the past by stimulating media effects researchers to think more at a macrolevel, to use more naturalistic approaches, and to reason more systematically across different interlocking components of the broad media phenomenon. In the past 4 decades, the literature of empirical studies that purport to test the claims of cultivation theory has moved away from Gerbner's original ideas and has evolved into an exploratory forum for any type of media
research as long as it looks at some form of exposure to media content and looks for the relationship of that exposure to some kind of effects outcome. When we examine the direction of this empirical literature and use it to assess the value of cultivation theory, we must conclude that the theory has lost its uniqueness as a system theory focusing on macrolevel phenomena of media influence on large aggregates over time; the literature has moved toward a pattern of research practices that test short-term reactions to specific message elements, which characterized the dominant form of research in the 1960s that Gerbner was trying to create an alternative for. Much of what is labeled as cultivation research now is oriented toward testing claims made by many other media effects theories much more than the claims made by cultivation theory. It is time for the authors of this boundary-crossing research to frame their studies more accurately as either tests of other more promising theories under the perspective of socialization or to create new theories that offer a much more clear and unique identity along with a more elaborate system of explanation that can generate a research literature with more powerful support.

Acknowledgments

The author thanks the editor of this journal and the two reviewers for their extremely valuable insights.

References


Shrum, L. J. (2004). The cognitive processes underlying cultivation effects are a function of whether the judgments are on-line or memory-based. *Communications, 29*(3), 327–344.


