



In Search of Entertainment-Education's Effects on Attitudes and Behaviors

Donald P. Green

The story of entertainment-education's effects on social behaviors is itself a literary genre of sorts, with memorable legends and alluring theoretical formulations about why dramatized messages work. One of my first encounters with this genre was in 2006, when reading a *New Yorker* magazine article that retold some of the most famous instances of entertainment-education's enormous reach and influence. For example, on the subject of adult literacy:

For this telenovela, called "Ven Connigo" ("Come with Me"), instead of the usual blond, blue-eyed leading ladies, Sabido chose actresses who were dark-haired and spoke with working-class accents. Its main plot centers on Barbara, a teacher from a working-class family, and Jorge, the wealthy man who loves her. Sabido's subplot involved some of Barbara's adult students, including a carpenter, a maid, a single mother who works on a farm, and an ex-con. Posters from the Ministry of Public Education's actual literacy campaign hung in the classroom. In the first episode, an older man sits outside Barbara's classroom without introducing himself. When she finally notices

D. P. Green (✉)
Columbia University, New York, NY, USA
e-mail: dpg2110@columbia.edu

© The Author(s) 2021
L. B. Frank, P. Falzone (eds.), *Entertainment-Education Behind the
Scenes*, https://doi.org/10.1007/978-3-030-63614-2_12

him and asks why he's there, he admits that he is embarrassed, because his grandchildren have finished primary school and he hasn't. The ratings for this series were higher than for any of Televisa's previous telenovelas. In one episode, Barbara's students visit the headquarters of the government's literacy program to pick up free booklets. The following day, more than twelve thousand people converged on the actual headquarters, creating a traffic jam that lasted until after midnight. Close to a million new students signed up for literacy classes. (Rosin, 2006)

At the time, I didn't think to question how the author documented the claim that "more than twelve thousand people converged on the actual headquarters" or that the number of new students was "close to a million." I figured the number of viewers is astronomical; the fan base watches attentively; a certain portion of them fanatically emulate the behaviors they see modeled on TV.

Another anecdote from the same essay paints a similar picture:

The most widely viewed program in the world is not a telenovela but "The Bold and the Beautiful," a CBS soap opera set in rival fashion houses. It is broadcast in a hundred and thirty-four countries, including Bangladesh, Uganda, and Yemen. In 2001, in a major twist, Tony, a young designer, told Kristen, his girlfriend, that he was H.I.V.-positive. The writers had consulted with experts from the Centers for Disease Control, and provided an 800 number that people could call for more information. After one particularly emotional scene, the line received five thousand calls. (Rosin, 2006)

My (admittedly academic) reaction at the time was that tracking the number of calls to an 800 number is an elegant way to measure the behavioral effects of a soap opera. I imagined a study in which one tracks the number of daily calls for a few weeks leading up to the episode in which the 800 number is displayed during the soap opera, then watches the number of calls soar as the episode airs, and assesses how long it takes for the number of calls to return to normal. Eventually, as I describe below, I collaborated on some experiments of this kind. But, to preview what comes later, our numbers didn't soar after the numbers aired on the telenovelas. Maybe our weak effects reflect the fact that our messages aired on ordinary soap operas rather than "the most widely viewed program in the world," but perhaps that is the point. Hundreds of millions of viewers evidently generated a few thousand calls; our programs' viewership was two orders of magnitude smaller, and so were the numbers of people who took some kind of measurable action.

Now you're probably wondering whether this is going to be one of those cranky academic essays that devotes its pages to puncturing inflated claims made by non-academics. It's not. This essay is instead a cranky account of how inflated claims backed up by poorly designed impact assessments (often by university researchers) have caused entertainment-education's effects to be misunderstood. More scientifically rigorous evaluations show entertainment-education (EE) does work but not in the transformative ways that proponents claim. When properly assessed, the effects of EE are found to be narrow. They often change specific ways in which audiences think and act, but much to the disappointment of funders or policymakers, EE seldom changes broad worldviews or core values. One possibility is that we have set our expectations for EE's effects too high. Another possibility is the lack of deep-seated change in the wake of exposure to narratives reflects the limitations of the content or presentation, in which case the producers and evaluators need to collaborate during the writing and production phases to produce more influential programs. The message of this essay is that if we want to do good in the world via narrative entertainment, we need to not only learn whether specific media content is effective, but also, more generally, understand the conditions under which it is effective. Funders do not seem to be embracing the latter goal, and as a result, evaluations tend to be disjointed, one-off affairs that focus on a specific donor-supported messaging campaign. This state of affairs slows the development of effective entertainment-education programming, which is arguably a tragic missed opportunity.

This essay is organized as follows. I begin by briefly summarizing the key findings from the burgeoning literature on narrative education conveyed through mass media. Here I call attention to the kinds of results one encounters when reading recent experimental evaluations of television and radio programs, both multi-episode dramas and short format messages. It is clear that narrative-based media content often has meaningful effects on attitudes and behaviors, but rarely does one see profound and enduring changes. Interestingly, although there are good theoretical reasons for thinking that narrative messages are more influential than non-narrative messages, few head-to-head comparisons have been conducted, and these focus exclusively on health behaviors rather than on controversial social issues. In other words, although I suspect that narrative messaging is an especially effective way of shaping policy-related attitudes and actions, direct evidence on this point is lacking. Next, I discuss my own experiences developing and testing narrative messages. In particular, I call

attention to some of my blunders and missteps, commenting on how they changed the way I think about what to study and how. Finally, I briefly sketch out a new model for developing effective messages that departs from the one-off evaluation approach that dominates the literature. I propose testing more immersive and sustained narratives whose ingredients are developed in conjunction with an evaluation team.

A BRIEF OVERVIEW OF THE ENTERTAINMENT-EDUCATION LITERATURE

A large and growing list of studies evaluates the effectiveness of entertainment-education in the form of video or audio narratives. This literature can be roughly categorized into four groups, based on the scientific rigor with which the evaluation is conducted. The first category comprises qualitative accounts. These reports typically rely on interviews with audience members or observe audiences as they view or listen to narratives. The strength of these studies is that the quotations and descriptions they provide give a sense of how audiences interpret and perhaps internalize the messages to which they are exposed. The main weakness of this approach is the lack of a well-defined control group that would allow us to assess whether and to what extent the audience changed as a result of exposure. Other problems include the obtrusive manner in which outcomes are measured—respondents are aware that they are being asked about the narrative that they were exposed to, and sometimes the questions invite them to describe how the drama changed their attitudes or beliefs. This heavy-handed approach risks encouraging respondents to say what they think the interviewer wants to hear, which presumably is some sort of praise for the narrative program. For what it's worth, such studies typically draw upbeat conclusions about entertainment-education programs.¹ In part, that reflects the leeway that researchers have when recounting audiences' reactions; there may be a temptation to cherry-pick the most positive quotations in order to placate funders or media partners. Anyone who has seen presentations of evaluation findings knows that vibrant anecdotes draw appreciative reactions—perhaps itself a

¹For example, the handbook *Edutainment: Using Stories & Media for Social Action and Behaviour Change* (Perlman, Jana, & Scheepers, 2013) is essentially a compendium of success stories. None of the interventions that are presented as case studies are evaluated via randomized experiment.

commentary on the effects of performance art. But from the standpoint of a detached observer who wonders what became of the qualitative responses that ended up on the cutting room floor, studies of this kind cannot be considered reliable guides on questions of cause and effect.

A second group of evaluations might be classified as “point-blank lab studies.” In contrast to the qualitative studies mentioned above, lab studies (usually) randomly assign participants to experimental conditions in which one group is exposed to a narrative message while the other group receives either some other kind of message or none at all. Random assignment makes for a fair comparison between a treatment group that is exposed to narrative messages and a control group that is not. However, two features of these studies undercut the value of the evidence they generate. First, participants are exposed to entertainment-education under artificial conditions—for example, they are often paid to participate in the study, and they watch the program because they are following the experimenter’s instructions. Second, outcomes (attitudes, beliefs, behavioral intentions) are typically measured a few minutes after participants are exposed to the media narrative. The compressed timeframe between exposure and outcome measurement may exaggerate the apparent effectiveness of the drama. For example, a study I conducted in Uganda compared the apparent effects of exposure to a three-part video about teacher absenteeism. One experiment was conducted in a lab-like environment in which subjects viewed the videos on a tablet computer and expressed their opinions immediately afterward; the other experiment tested the same videos by embedding them in a multi-week film festival and assessed outcomes two months later. Both tests showed that the ads substantially increased viewers’ willingness to take actions to address absenteeism, but the effects found in the lab were more than twice as large as the effects found in the field (Green, Cooper, Wilke, & Tan, 2020a).

A third category includes nonexperimental studies that track changes over time or between locations, as regions gained access to mass media or soap operas. These studies often present arresting examples of how changes in attitudes or behaviors coincided with the rollout of mass media. One of the most prominent studies tracked the consequences of cable TV access in Indian villages and found that villages that acquired cable access showed abrupt shifts in women’s subsequent attitudes toward violence against women (VAW) and their preference for having male children (Jensen & Oster, 2009). Another influential study found that Brazilian women aged 25 to 44 living in regions that were exposed to popular soap operas

subsequently experienced declines in fertility (La Ferrara, Chong, & Duryea, 2012). These intriguing studies and others like them raise two concerns. The first is whether, short of conducting an experiment, we can draw reliable inferences about the effects of mass media by comparing the groups that did or did not receive it. The concern is that even if media truly had no effect, there might appear to be an effect due to unmeasured differences between the treated and untreated groups. A second concern is publication bias, or the tendency to write up splashy results that show large effects. Had the comparison of geographic regions shown no effect, would the researchers have bothered writing up the results knowing that academic journals were unlikely to be impressed? If studies that fail to find effects remain in file drawers, the academic literature may exaggerate the influence of media or narratives.²

The final category includes randomized trials that are conducted in real-world settings. I recently collaborated on an effort to assemble a fairly comprehensive literature review of entertainment-education, and we found 30 such studies. The take-home lessons from this literature can be summarized as follows. First, it is clear that entertainment-education “works” in the sense that it regularly generates effects on beliefs, attitudes, behavioral intentions, and behaviors. Audiences update their beliefs about topics such as the risk of HIV infection in the wake of information that is conveyed through narrative. Attitudes, or dispositions to respond positively or negatively to stimulus objects, also change. For example, rural Tanzanians exposed to a two-hour radio soap opera about forced marriage of underage girls became more likely to say two weeks later that girls should be able to choose whom they marry (Green, Groves, & Manda, 2020b). And narratives seem to spur audiences to act, such as texting a hotline to report instances of corruption when this behavior is modeled in a feature-length movie (Blair, Littman, & Paluck, 2019). Among the studies that measure behavioral intentions under various scenarios, narratives again seem to move the needle. Green, Wilke, and Cooper (2020c) studied Ugandan villagers who were exposed to videos that stress the importance of preventing violence against women by reporting abuse; eight months later, those who saw the videos were more likely to express

²One potential advantage of studies conducted outside academia is that they are less susceptible to publication bias, since publication is seldom the objective. On the other hand, these studies often fail to include key technical details about the experimental design or analysis that would ordinarily be expected from a peer-reviewed publication.

willingness to report abuse to authorities. Almost every published study reports at least one “statistically significant” (i.e., too large to be attributed to chance) finding on an attitude, belief, behavioral intention, or behavior.

That said, EE’s effects are rarely large. One of the strongest reported effects is that Nigerians shown eight 22-minute episodes of the TV series MTV *Shuga* on consecutive weekends displayed increases in knowledge and HIV testing, with mixed results concerning risky sexual behavior (Banerjee, La Ferrara, & Orozco-Olvera, 2019a). In relative terms, HIV testing is almost twice as high in the treatment group as the control group, but in absolute terms, this effect amounts to a 3 percentage point increase (i.e., a change among 3 out of every 100 viewers). In a similar vein, Ugandans in villages that were randomly assigned three five-minute videos dramatizing teacher absenteeism during the commercial breaks of a free film festival showed clear increases in their support for community action to address absenteeism *vis-à-vis* their counterparts in other villages that received messages on other topics. Eight months later, approximately 59% of the treatment group favored organizing or participating in some kind of collective response to absenteeism in the local school, as opposed to 54% of the control group. Given the large number of villages and participants, the treatment effect is convincing but by no means overwhelming.

With few exceptions, the changes that do occur are specific to the particular messages that are conveyed in the drama. For example, the Ugandans who saw the messages stressing the importance of preventing violence against women became more inclined to report incidents to village leaders but did not change their minds about the more general issue of whether husbands are justified in beating their wives. Similarly, the Tanzanians exposed to a soap opera that made them more skeptical about early/forced marriage did not change their views about gender hierarchy more generally. So far as we can tell, the rapid transformation in core attitudes occasioned by exposure to TV in nonexperimental studies such as La Ferrara (2012) has not been reproduced in randomized trials.³

A final observation about the EE literature is how rarely researchers conduct head-to-head tests of narrative versus non-narrative messaging about

³ One partial exception is Banerjee, La Ferrara, and Orozco-Olvera (2019b), which finds a marginally significant effect of watching *Shuga* (which focused in part on gender-based violence) on respondents’ views that husbands are justified in beating their wives under various scenarios.

social issues. The studies of which we are aware of focus on health-related knowledge and behaviors. Tests of competing videos report conflicting results. Moran, Frank, Chatterjee, Murphy, and Baezconde-Garbanati (2016) found that 11-minute videos more effectively communicated information about cervical cancer when they contained a narrative component, and Murphy et al. (2015) found that narrative content was more effective at inducing women to schedule a PAP test in the six-month period since exposure to the video. On the other hand, Bekalu et al. (2018) tested competing four-minute video clips on pandemic influenza and found the non-narrative version to be better at imparting knowledge. The mixed findings about health knowledge and behaviors suggest that narratives are not necessarily more informative or compelling, at least given an experiment in which participants are expected to look at whatever message is put in front of them. It may be that in more naturalistic viewing environments, narratives have the advantage of attracting and retaining willing audiences.

But even if we accept the mixed results from the health literature, it remains unclear what to conclude about the effects of narratives on widely held social attitudes, such as the belief in rural East Africa that a husband has a legitimate right to beat his wife if she disobeys. Social attitudes invite resistance from those whose viewpoint is challenged by the message. Narratives, according to one leading theory, break down resistance by *transporting* audiences and inducing them to take the perspective of the protagonists (Green & Brock, 2000). Narratives also *model* appropriate attitudes and effective behaviors in ways that lead audiences to update how they are expected to think and act (Bandura, 2004). The hypothesis that underlies entertainment-education is that narrative has the ability to win over an audience that would otherwise reject direct arguments against their social and policy convictions. Unfortunately, we simply do not have direct evidence about whether narratives work as these theories suggest in the domain of social attitudes. At present, the evidence in hand indicates that narratives produce meaningful but not overwhelmingly large effects, primarily in the narrow domain on which the narrative focuses.

EVALUATION AND CONTENT CREATION

My involvement in the entertainment-education literature has been in the role of evaluator. From time to time, philanthropic organizations, research groups, or NGOs have offered me the opportunity to conduct randomized trials to see if dramas can influence attitudes, beliefs, or behaviors on

a wide array of topics: gender-based violence, early/forced marriage, abortion stigma, HIV stigma, teacher absenteeism, or vote-buying. Beyond the study of media dramas in particular, I am also involved in ongoing randomized evaluations of access to radio and TV, with an eye toward verifying the findings of Jensen and Oster (2009) on the effects of media access. Most of my studies have taken place in East Africa, although I have also conducted some research in India and the United States. I have not been paid to conduct these evaluations; I take them on because my scholarly interests include media effects, public opinion, and the conditions under which people change their views about social issues.

My involvement as an evaluator takes one of two forms. The first mode might be described as “program evaluation” in both senses of the term. Here, the research team is asked to evaluate the effectiveness of existing soap operas that have already been aired. Since we cannot go back in time, my approach is to assess their effects on new audiences that are similar to the ones that were exposed “naturally” when the programs were aired initially. For example, the Tanzanian NGO UZIKWASA, which focuses on issues of gender equality, produced a multi-part radio soap opera set in the northeast of the country that depicted early and forced marriage in a Muslim family. Many remote villages in this region, however, were outside the catchment area of the radio station that initially aired the series. Our research team selected 30 unexposed villages and interviewed a random sample of villagers at baseline. At the end of this interview, we invited them to attend a two-hour audio drama a few days later, without disclosing the content of the program. After baseline data were collected, we randomly assigned each village to one of the two soap operas: one on forced marriage or another on HIV stigma. Fully 83% of those invited showed up, with almost identical attendance rates in each experimental group. Two weeks after the radio programs aired, 95% of participants were reinterviewed in their homes. (We plan to return to interview respondents one more time once the COVID-19 threat lifts.)

One attractive feature of this experimental design is that we simultaneously test the effects of two dramas on two different sets of outcomes. In this case, we learned that one soap opera increased support for letting girls choose whom they marry, while the other increased support for increasing access to retroviral drugs. Another feature of this type of design is that we get answers back quickly. Follow-up interviews were conducted two weeks after audiences listened to the soap operas, so the results were in hand quickly enough to make timely changes to the content of the programs if

we had needed to do so. From the standpoint of practitioners, this design strikes a useful balance between the speed of a point-blank lab study and the naturalism of a field experiment with long-term follow-up.

That said, this experiment is far from perfect. To condense 20 hours of soap opera content to a two-hour format changes the nature of the listening experience, and our evaluation does not necessarily do justice to the effects that the original series may have had over the many weeks that listeners tuned in. An alternative approach is to bring audiences together repeatedly over the course of several months to have them listen to the complete soap opera, as Paluck (2009) did in her study of an ethnic reconciliation radio drama in Rwanda. Not only did our study change the cadence of exposure to the drama; it also changed the content, as we sought to distill the original plot down to its bare essentials.

The other mode of evaluation is to have the research team coordinate the writing and production of a drama, whose effects are then tested by way of an experiment. One of my first attempts along these lines was in collaboration with a team of scholars led by Elizabeth Levy Paluck, whose pathbreaking work in this field was mentioned earlier. In this study, we worked with a leading Spanish-language network in the United States to weave nine social messages (e.g., put your infant into car seat when driving, register to vote, eat low-cholesterol foods) into the scripts of its nightly *telenovelas*. The experimental design randomly varied *when* each theme was woven into the plot; once this was determined, we worked with scriptwriters to provide information about each issue and guidance about the goals of each message. For example, one message encouraged audiences to put their savings into a bank account rather than save cash at home. The subplot featured a main character working up the courage to visit a bank to open up an account; an actual bank with a large number of local branches was depicted, and the scenes emphasized the way in which the new client was helped by a friendly bilingual bank officer. This bank shared with the research team daily records of how many new accounts were created before, during, and after the airing of these scenes. No apparent uptick in accounts occurred. Similarly disappointing results were obtained for other behaviors, such as registering to vote or visiting a website that featured information about college scholarships for minority students (Paluck et al., 2015). I might add parenthetically that these results changed my view about the effects of entertainment-education. I am now more skeptical that large aggregate behavior changes routinely occur in the wake of subplots among audiences in developed countries. It may be the case that behavior changes

occur when entire series focuses on a given theme, as in “16 and Pregnant” (Kearney & Levine, 2015), although even here the (nonexperimental) evidence is debatable (Jaeger, Joyce, & Kaestner, 2019).

Evaluation teams are sometimes given an opportunity to craft the main storyline, not just a subplot. In partnership with Peripheral Vision International, my research team worked to develop a mini-series of three five-minute vignettes that would be aired to rural Ugandans during commercial breaks in feature-length films. In an effort to make it easy for audiences to be drawn in by these narratives, the vignettes were written by local writers and filmed on location in the local language. Overdubbed Hollywood films were shown each weekend for four to six weeks, initially in 56 villages during the pilot study and a year later in 112 villages during the main study. The experiment randomly assigned different mini-series to villages hosting the film festivals.

The pilot testing phase turned out to be crucial, as it gave us an opportunity to test audiences' reactions to each of the mini-series. One of the series focused on violence against women and sought to articulate and model official norms regarding the topic. Characters that included a visitor from Kampala, a village leader, villagers, and police officers all expressed norms that wife-beating is illegal, immoral, and unjustified in all circumstances. In the end, the abusive husband is arrested by police. Bear in mind that Ugandans, especially rural Ugandans, widely believe that husbands are justified in beating their wives when presented with scenarios that include disobedience, gossiping, unfaithfulness, and the like; thus, the views expressed by characters in our mini-series ran counter to prevailing or at least widespread views. Reports from the field suggested that this message did not sit well with male audience members, and our surveys two months later revealed that the video, which was randomly assigned to some villages and not others, had no apparent effect on attitudes about gender-based violence. (Meanwhile, the other two series on teacher absenteeism and abortion stigma seemed to be well-received and effective in shaping the views of audiences whose villages were randomly assigned to see them.)

Rather than repeat this debacle on an even larger scale in the main study of 112 villages, we developed an alternative mini-series on the same topic but this time building on locally prevailing norms. Depth interviews with villagers suggested that although most rural Ugandans (even women) believed that husbands are justified in beating their wives under some conditions, when pressed to say whether by “beaten” they had in mind a slap or something more forceful than that, respondents overwhelming said a

slap. This apparent norm against extreme violence came through as well in our structured surveys of villagers. So we decided that the new video should feature a case of unacceptable violence, which unfortunately is quite common in rural Uganda, where a large share of women report that their husband has at some point beaten them severely. Our *Tale of Two Cities* narrative begins in a village where there is reluctance to report VAW. The protagonist is an affable woman whose husband beats her severely, despite her sincere efforts to appease him. In a crucial scene, the protagonist's neighbor overhears her screams but decides not to speak out. In the second vignette, which begins with the protagonist's hospitalization and ends with her funeral, audiences learn that her daughter and parents also knew about the violence and now regret their failure to speak out. The final vignette depicts the "disclosure" village. The focal woman in this scene is also beaten by her husband but decides to disclose this information to her parents, who intervene to help mediate. Moreover, the parents share the information with the local women's counselor (*Nabakyala*), who visits the household to provide guidance. The vignette closes with the couple getting along, and the voiceover reminds the viewer of the importance of saying something in order to prevent violence. This video mini-series did not work miracles (it had no effect on views about whether husbands are justified in beating their wives or on attitudes about gender hierarchy) but did have meaningful and persistent effects on viewers' willingness to report incidents to authorities. We also found evidence suggesting a decline in violence against women in the villages that were randomly chosen to receive this message, perhaps reflecting the deterrent effect of being in a village where more people are willing to speak out.⁴

In some sense, it was dumb luck that we stumbled on a video that worked given that we only had two chances. We were guided by theories that emphasized the persuasive effects of narratives that model appropriate behaviors and that draw viewers into the story such that they let down their guard. Yet, it is clear from our first failure that audiences need not follow the models that are presented to them, and that they do not let down their guard for just any narrative. The fact that the second mini-series worked (to some extent) suggests that modeling norms might work when writers meet their audiences half-way, in this case denouncing

⁴Although the combination of production and evaluation costs totaled more than \$300,000, this sum is arguably a bargain for a scalable media intervention that is known to be effective.

excessive violence rather than any violence, but even that is a conjecture rather than an established fact.

The experience of developing and testing successive mini-series on the same topic inspired us to dig more deeply into the question of what makes a particular narrative effective. My co-authors, Jasper Cooper and Anna Wilke, led an effort to re-edit the footage from the two mini-series so that the narrative and characters had different combinations of features. In some mini-series, the woman who is beaten is depicted in a way that inspires empathy; in some plot lines, the theme that violence can spiral out of control is stressed; sometimes the storyline and narration emphasize the benefits of intervention; sometimes the script mentions legal norms that declare wife-beating to be illegal. In all, ten different combinations were produced and tested in a lab-like setting against a placebo video about teacher absenteeism. The lab-like setting is less than ideal, especially since participants watched the video alone rather than in a communal setting; what's worse, because we were coming to the end of our grant, the study was far too small (351 participants spread over 11 different experimental conditions). But putting aside the details of how we conducted this small study, the broader point is that head-to-head testing of narratives with different ingredients is an attractive way to discover the most effective entertainment-education package for a given audience. We should have done this sort of test before conducting our field tests. Indeed, this kind of evidence-based R&D should be part of the ramp-up to any entertainment-education initiative.

From a theoretical standpoint, even this kind of retrospective R&D is potentially of enormous value. Leading theories are stated at a level of generality that is detached from the specifics that might guide scriptwriters as they attempt to craft influential narratives. I am told that the development of narrative entertainment content is sometimes guided by focus group reactions to preliminary versions, but I am aware of no rigorous experimental tests that add and subtract narrative elements in an attempt to discern the mechanisms by which entertainment is influential. Suppose we assume that audiences are persuadable when they are drawn into a story and take the perspective of the protagonist. Very well, what narratives, audiences, and contexts contribute to this kind of perspective-taking? Do the persuasive effects increase or diminish when audiences view the narratives collectively (as often occurs in rural East Africa) as opposed to individually (as often occurs in urban and more affluent areas)? Detailed theorizing requires a rigorously developed evidence base that offers robust insights into what works.

Lessons Learned and Best Practices

The education entertainment literature is filled with inspiring examples that make success look a lot easier to achieve than it actually is. A sober reading of the literature suggests that entertainment-education is a promising and often cost-effective avenue for effecting social change, but we currently lack an empirically grounded understanding of the conditions under which narrative entertainment works. To develop this understanding, however, requires a very different approach than what is currently on offer. Philanthropic funders typically focus on evaluating a single entertainment-education product (i.e., the one that the group that they funded produced), not theoretically telling alternative versions of it. Scholars are often quite content with the one-at-a-time method of evaluation, perhaps because it leads to lots of discrete publications, but this disjointed process slows that rate at which theoretical knowledge is accumulated. And when scholars do compare alternative messages, they tend to focus on bite-size media content so that tests can be conducted conveniently within the confines of a lab-like environment. We need to think bigger, on both the production front and the evaluation front. Given the kinds of social outcomes that hang in the balance, we urgently need a more coherent evidence-based partnership between creative artists and evaluators.

The alternative vision is a much more systematic investigation of the conditions under which entertainment-education works. For any given topic or target audience, this investigation requires exploring an array of narratives that vary along an assortment of dimensions. What attitudes and behaviors are modeled, and by whom? What kinds of archetypal storylines work best for attracting sizable audiences and persuading them to think or behave differently? To what extent does the effectiveness of the narrative depend on serialization and repetition of stories and themes? Do the effects of entertainment-education grow reliably with the dosage of drama that audiences receive?

One may think of this kind of experimental study as a bake-off among different recipes, with the aim of discovering the ingredients—or the portion sizes—that tend to produce large and persistent changes. It may turn out that even optimized entertainment interventions fail to produce effects of legendary proportions, yet the exercise will still be worth the effort if it points to reliable ways to make entertainment-education more effective.

Acknowledgements I thank Dylan Groves, Beatrice Montano, and Bardia Rahmani for their comments on earlier drafts and for their help in assembling the entertainment-education literature.

REFERENCES

- Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior, 31*(2), 143–164.
- Banerjee, A., La Ferrara, E., & Orozco-Olvera, V. H. (2019a). *The entertaining way to behavioral change: Fighting HIV with MTV*. The World Bank.
- Banerjee, A., La Ferrara, E., & Orozco-Olvera, V. H. (2019b). Entertainment, education, and attitudes toward domestic violence. *AEA Papers and Proceedings, 109*, 133–137.
- Bekalu, M. A., Bigman, C. A., McCloud, R. F., Lin, L. K., & Viswanath, K. (2018). The relative persuasiveness of narrative versus non-narrative health messages in public health emergency communication: Evidence from a field experiment. *Preventive Medicine, 111*, 284–290.
- Blair, G., Littman, R., & Paluck, E. L. (2019). Motivating the adoption of new community-minded behaviors: An empirical test in Nigeria. *Science Advances, 5*(3), eaau5175.
- Green, D. P., Cooper, J., Wilke, A. M., & Tan, B. (2020a). *Encouraging community action against teacher absenteeism: A mass media experiment in rural Uganda*. Unpublished manuscript, Columbia University.
- Green, D. P., Groves, D. W., & Manda, C. (2020b). *A radio drama's effects on attitudes toward early and forced marriage: Results from a field experiment in Rural Tanzania*. Unpublished manuscript, Columbia University.
- Green, D. P., Wilke, A. M., & Cooper, J. (2020c). Countering violence against women by encouraging disclosure: A mass media experiment in rural Uganda. *Comparative Political Studies, 53*(14), 2283–2320.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701–721.
- Jaeger, D. A., Joyce, T. J., & Kaestner, R. (2019). *Tweet sixteen and pregnant: Missing links in the causal chain from reality TV to fertility*. w25446. National Bureau of Economic Research.
- Jensen, R., & Oster, E. (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics, 124*(3), 1057–1094.
- Kearney, M. S., & Levine, P. B. (2015). Media influences on social outcomes: The impact of MTV's 16 and pregnant on teen childbearing. *American Economic Review, 105*(12), 3597–3632.
- La Ferrara, E., Chong, A., & Duryea, S. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics, 4*(4), 1–31.

- Moran, M. B., Frank, L. B., Chatterjee, J. S., Murphy, S. T., & Baezconde-Garbanati, L. (2016). A pilot test of the acceptability and efficacy of narrative and non-narrative health education materials in a low health literacy population. *Journal of Communication in Healthcare*, 9(1), 40–48.
- Murphy, S. T., Frank, L. B., Chatterjee, J. S., Moran, M. B., Zhao, N., Amezola de Herrera, P., et al. (2015). Comparing the relative efficacy of narrative vs nonnarrative health messages in reducing health disparities using a randomized trial. *American Journal of Public Health*, 105(10), 2117–2123.
- Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3), 574–587.
- Paluck, E. L., Lagunes, P., Green, D. P., Vavreck, L., Peer, L., & Gomila, R. (2015). Does product placement change television viewers' social behavior? *PloS One*, 10(9), e0138610.
- Perlman, H., Jana, M., & Scheepers, E. (2013). *Edutainment: Using stories & media for social action and behaviour change*. Soul City Institute. Retrieved from <https://www.soulcity.org.za/resources/research/published-articles/edutainment-using-stories-and-media-for-social-action-and-behaviour-change>.
- Rosin, H. (2006). Life lessons: How soap operas can change the world. *The New Yorker*, 82(16), 40–45.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

