Exposure to Smoking in Movies and Smoking Initiation Among Black Youth

Sonya Dal Cin, PhD, Mike Stoolmiller, PhD, James D. Sargent, MD

Appendix A

Explanation of model parameterization

In order to test the proposed hypotheses, the regression model was parameterized somewhat differently from what is usually presented. The hypotheses here were that among black viewers, black-oriented movie smoking (BMSE), but not mainstream movie smoking exposure (MMSE) will predict smoking initiation, whereas among nonblack viewers, both types of exposures will predict initiation. These hypotheses relate to differences in the effects of different types of exposures *within* racial groups, not the differences *between* the groups. Indeed, a test of these hypotheses does not involve any examination of the significance of the difference between the two groups of participants (black versus all other teens).

Other analyses that involve multiple groups are typically designed to test differences between those groups. To do so, analysts typically specify a model that estimates the effect of group membership on the outcome relative to a reference group. In the case of examining whether the associations between BMSE and MMSE are different for black youth than for their nonblack peers, the following parameterization is typical:

Equation 1

$$\log(\text{odds}(\text{smoking}_i)) = b0 + b1(\text{black}_i) + b2(\text{BMSE}_i) + b3(\text{MMSE}_i) + b4(\text{MMSE}_i^2) + b5(\text{black}_i*\text{BMSE}_i) + b6(\text{black}_i*\text{BMSE}_i^2) + b7(\text{black}_i*\text{MMSE}_i) + b8(\text{black}_i*\text{MMSE}_i^2) + b9(\text{gender}_i) + b10(\text{SES}_i) + \text{linear effects of additional covariates.}$$

The asterisk denotes multiplication, and the subscript i indicates the ith subject. This model includes a constant term (b0); a main effect for BMSE (b2); and linear (b3) and quadratic (b4) main effects for MMSE. Black is a dichotomous dummy variable, equal to 1 for a black subject (0 otherwise). This model would give direct estimates of the BMSE and MMSE dose–response curves for the nonblack sample. The terms involving the (dummy) variable black represent differences between the dose–response curves for black and nonblack subjects—not the dose–response parameters for black respondents. This familiar parameterization produces estimates that test different hypotheses than the ones posed in the current paper.

To test the current hypotheses, the following logistic regression model was constructed:

Equation 2

$$\log(\text{odds}(\text{smoking}_i)) = b1(\text{black}_i) + b2(\text{black}_i*\text{BMSE}_i) + b3(\text{black}_i*\text{BMSE}_i^2) + b4(\text{black}_i*\text{MMSE}_i) + b5(\text{black}_i*\text{MMSE}_i^2) + b6(\text{nonblack}_i) + b7(\text{nonblack}_i*\text{BMSE}_i) + b8(\text{nonblack}_i*\text{MMSE}_i) + b9(\text{nonblack}_i*\text{MMSE}_i^2) + b10(\text{gender}_i) + b11(\text{SES}_i) + \text{linear effects of additional covariates.}$$

The asterisk denotes multiplication, and the subscript i indicates the ith subject. Black and nonblack are both dichotomous dummy variables. Black is 1 for a black subject (0 otherwise); nonblack is the reverse (1 for a nonblack subject and 0 otherwise). The model does not include a constant term (i.e., a single intercept for both groups) or main effects for BMSE and MMSE (i.e., effects of the different exposures collapsed across race). Indeed, these terms would be redundant and perfectly collinear with the terms already in the model (e.g., Gelman and Hill[1]).

Parameterizing the model in this fashion results in direct estimates of the dose–response curves for BMSE and MMSE among black subjects (product terms black*BMSE and black*MMSE, respectively) and the corresponding estimates for nonblack subjects (product terms nonblack*BMSE and nonblack*MMSE, respectively).

In the current analysis, both BMSE and MMSE were centered about a value of 200 occurrences (a compromise that avoids the extremes of the distributions) and rescaled by a factor of 400. Thus, raw regression effects are scaled per 400 movie smoking occurrences, and lower-order effects describe the nature of the dose–response curve (level and slope) at the specific exposure of 200 occurrences.

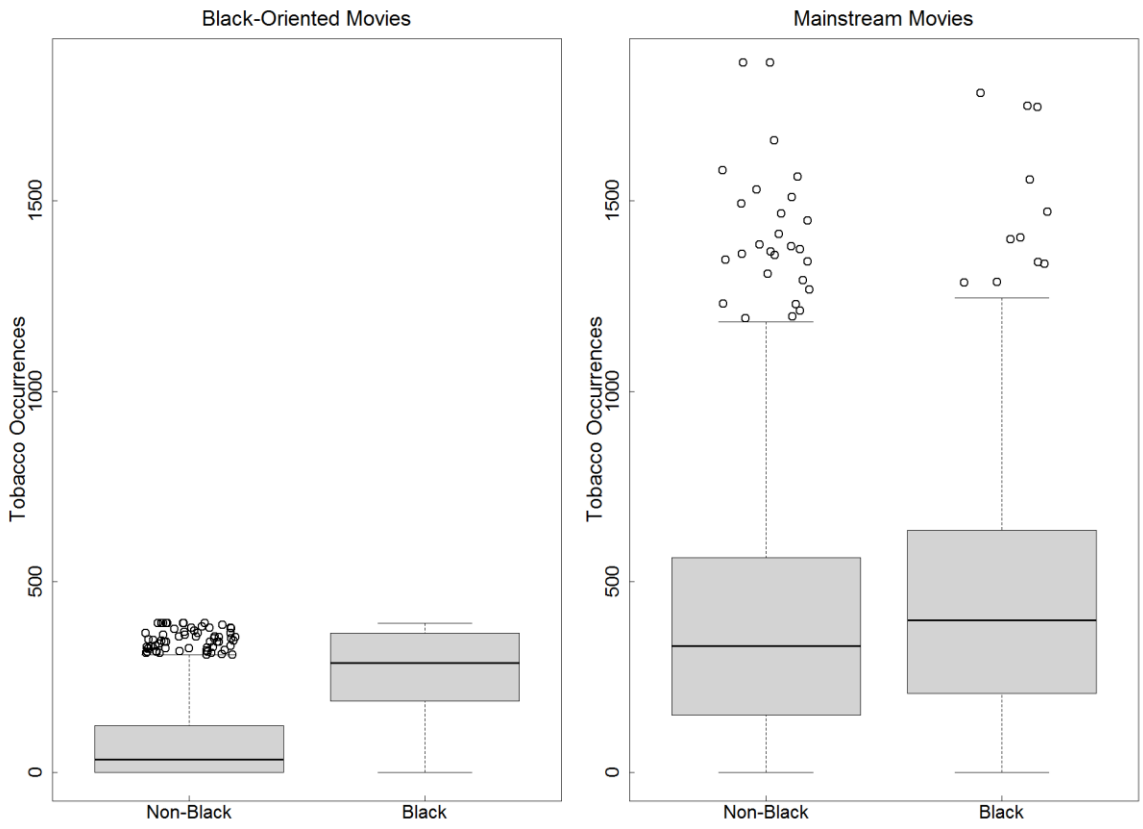
For gender, SES, and the additional covariates, no interactions or product terms with race/ethnicity are included. This means that these variables have the same effects in both subgroups in the current model. Some of the covariates (e.g., gender) were nominal and without an obvious reference group; in such cases, the arbitrarily chosen reference group (e.g., males, for gender) is noted in Table 2 in the main text.

References for Appendix A

1. Gelman A, Hill J. Data analysis using regression and multilevel/hierarchical models. New York: Cambridge University Press, 2007.

Appendix B

Exposure to smoking occurrences in black-oriented and mainstream movies by race of viewer



Gray boxes indicate the interquartile range (IQR); horizontal lines inside boxes indicate the median. Whiskers below the boxes indicate minimum values, those above indicate values more than 1.5 times the IQR above the 75th percentile; circles indicate outliers (jittered).