

# **Failure of Intervention or Failure of Evaluation: A Meta-evaluation of the National Youth Anti-Drug Media Campaign Evaluation**

Stephen Magura, Ph.D.  
The Evaluation Center, Western Michigan University

Research on Evaluation TIG  
26th Annual Conference of the American Evaluation  
Association  
Minneapolis, MN

October 27, 2012

# BACKGROUND

- The national Youth Anti-Drug Media Campaign, mainly an extended series of television and other media advertisements (“ads”), was conducted during 1998-2004 and was evaluated through a national panel study of adolescents.
- The evaluation contractor’s results were unexpected and controversial, finding both no effect overall and a possibly harmful effect, namely inducing initiation of marijuana use (Hornik et al., 2008).
- The U.S. General Accountability Office (GAO, 2006) conducted a meta-evaluation (critique) of the original evaluation to determine whether the results should be used to guide adolescent drug use prevention policy and funding, and if so, to encourage such use by Congress.
- The GAO’s meta-evaluation supported the original evaluation’s major conclusions, but the Campaign’s sponsor, the Office of National Drug Control Policy (ONDCP), contested both the original evaluation’s findings and the GAO’s assessment of them (Walters, 2006).

# PURPOSE OF THIS STUDY

- To conduct an alternative meta-evaluation of the original evaluation, taking into account both the GAO's meta-evaluation and the ONDCP's objections to both.
- To elucidate further the effects, if any, of the Anti-Drug Campaign and particularly to identify "lessons learned" that might improve the conduct of future evaluations of national drug use prevention initiatives.

# STUDY METHOD

- Review publicly available documentation (published articles, reports, commentaries) on the national evaluation.
- Conduct a meta-evaluation of the national evaluation, focusing on an “evaluability assessment” of the program.
- Evaluability assessment - process by which evaluation sponsors, stakeholders and evaluators determine whether an intervention meets the preconditions for evaluation and how an evaluation should be designed to ensure that the evaluation questions can be answered with an adequate degree of certitude (Wholey, 2004).

# RESULTS

# The evaluation design

- Evaluability assessment includes determining whether an evaluation design can be implemented which will answer the evaluation questions with an acceptable degree of certainty.
- The design of the national evaluation was a purely observational study, consisting of a panel (longitudinal) study of age cohort samples of youths aged 9-18 years old (n= 8117 at baseline to 5126 at three year follow-up).
- Controlled evaluation designs were not fully considered; e.g., ONDCP appeared to subscribe to the view that controlled designs are “inappropriate for public service campaigns” (Walters, 2006).

# The evaluation design (cont.)

- For its part, the GAO did not question the observational evaluation design, limiting its meta-evaluation to assessing the evaluation as it was implemented.
- Contractor believed limitations of a purely observational design could be overcome by applying sophisticated analytical techniques to the observational data.
- Unfortunately, this created fundamental concerns about establishing causality between ad exposure and outcomes that plagued the interpretation of the evaluation results.

# Measuring cognitive marijuana-related outcomes

- Evaluability assessment includes determining the availability of valid impact (outcome) measures.
- The main cognitive impacts selected for this drug use prevention campaign for youths were:
  1. beliefs and attitudes about marijuana use consequences;
  2. perception of social norms concerning marijuana;
  3. self-efficacy to avoid marijuana;
  4. intention to use marijuana in the future
- If these measures show observed change, there are at least two competing explanations: “true” change and shift to more socially desirable responding.



# Measuring cognitive marijuana-related outcomes (cont.)

- The latter would be due to continued Campaign exposure resulting in greater reluctance to admit stigmatized or shameful perceptions about marijuana or to admit marijuana use.
- Prolonged exposure to the Campaign could induce socially desirable responding which could be misinterpreted as a positive effect of the Campaign.
- Neither the evaluation report, the GAO report nor the ONDCP response addressed this issue, other than the GAO noting that computer-assisted self-interviewing was used for sensitive questions.

# Measuring marijuana use as an outcome

- Marijuana use is an illicit and socially stigmatized behavior.
- Illicit drug use of all types is underreported in studies of high risk populations (Magura and Kang, 1996; Magura et al., 1999).
- Marijuana use has been shown to be greatly underreported in general population of youths/young adults (Magura, 2010).
- Unlike the measurement of cognitions relating to drug use, drug use itself can be measured objectively by drug tests (e.g, urinalysis, saliva analysis).
- There was no evidence that the limitations of self-reports of marijuana use were considered in designing the evaluation or that drug testing was given consideration as a part of the data collection.
- The GAO concluded that, if illicit drug use was stigmatized as found in previous studies, then “increased stigma associated with drug use would lead to decreases in self-reports of drug use over time.”

# Independence of intervention and outcome measures

- Evaluability assessment should also consider the independence of the intervention and planned outcome measures.
- The evaluation believed that it was “rely(ing) on naturally occurring variation in campaign exposure among individuals to estimate the Campaign’s effects, after adjustment for variation in potential confounders, including the amount and type of media consumption” (Hornik et al., 2008).
- Adjustment for confounders refers to observed factors that might cause joint variation in both ad exposure and outcome measures, leading to spurious relations between exposure and outcomes.

# Independence of intervention and outcome measures (cont.)

- What this overlooks is a second potential source of bias, which is that the putative outcomes may be the cause of changes in ad exposure.
- In other words, changes in marijuana use and related perceptions might affect ad exposure, instead of, or in addition to, ad exposure affecting marijuana use and perceptions.
- Thus, when attempting to draw inferences from observational data, controlling for confounders (spurious relationships) is not sufficient; attempting to establish the direction of effect (if there is an effect) between an intervention and a putative outcome measure is also important.

# Adequacy of the analytical plan

- Evaluability assessment must consider whether the analytical plan is likely to yield unambiguous conclusions about the impact of the program.
- Three types of analyses were planned:
  1. Changes over time (trends) for each outcome (measures of marijuana-related cognitions and self-reported marijuana use).
  2. Correlations between non-marijuana users' self-reports of degree of ad exposure and marijuana-related cognitions.
  3. Lagged correlation analysis which examined the relation between degree of ad exposure during an earlier period and outcomes at a later period, with the logic that this demonstrates a causal direction from exposure to outcome, since an effect cannot precede a cause in time.

**How adequate was this analysis  
plan for assessing the results of the  
Campaign?**

# *Adequacy of the “trend analysis”*

- The evaluation report stipulated that this result would be uninterpretable, because “upward or downward trends can be the result of many influences, without the Campaign necessarily being the cause” (Hornik et al., 2008).
- So, why was it done?

# *Adequacy of the cross-sectional analysis*

- Even if degree of ad exposure is assumed to affect the cognitive measures, and even if we control statistically for possible confounders, we remain unable to distinguish between true perceptions and socially desirable responding.
- Any significant correlations could be attributable to a direction of effect opposite to the one assumed, that is, perceptions could affect ad exposure.
- Is a hypothesis which leads to an uninterpretable result even suitable for testing?



# *Adequacy of the lagged correlational analysis*

- The evaluation hypothesized that, if the Campaign were effective, the degree of ad exposure during an earlier period should predict outcomes at a later period.
- This begins to address the issue of direction of causation.
- What is missing, however, is examination of the reverse direction of effect – whether the measures of outcome during an earlier period predict ad exposure at a later period.
- The evaluation did not conduct the full cross-lagged panel correlation analysis.
- Unfortunately, the logic of cross-lagged correlation analysis is only pertinent if the outcome measures are valid, in particular not subject to social desirability bias induced by exposure to the ads themselves.

**Did The Analysis Support The  
Interpretation Of The Results?**

- Most pertinent point: 25 of the 30 correlations in the cross-lagged analysis were *not significant*; thus there is *no consistent pattern of effects involving ad exposure*, whether favorable or unfavorable.
- The original evaluation concluded that more ad exposure was associated with subsequent initiation of marijuana use. This was based on one correlation involving general ad exposure for one time period, out of six analogous correlations for marijuana use (Hornik et al., 2008: Table 4).

# **DISCUSSION AND CONCLUSIONS**

- Should this publicly-funded evaluation study, as designed, have been conducted at all?
- There was insufficient attention to evaluability assessment in the design of the Anti-Drug Campaign and its evaluation, which led to grave difficulties in interpreting the evaluation findings.
- One plausible and superior evaluation design would have been to compare outcomes for media markets exposed to the Campaign with outcomes for similar markets not exposed, using a randomized design.

- Is drug testing a practical recommendation within such an evaluation design?
- Subsequent to the planning of the Campaign evaluation, it was demonstrated that biological specimens for testing can be obtained from general population samples of adults (Fendrich et al., 2004; Colon et al., 2001) and adolescents (Harrison et al., 2007).
- There are indications that the Campaign was ineffective, consistent with the conclusions of the original evaluation and the GAO meta-evaluation.
- But had the Campaign been effective, the evaluation would have been unable to determine that fact due to lack of a control group and the high chance of socially desirable responding.

- Conclusion: the Anti-Drug Campaign probably failed to achieve its aims and the evaluation as designed should never have been conducted.
- This example illustrates the importance of an iterative, continuing process of meta-evaluation that should occur more often than it does in the evaluation field.

ARTICLE IN PRESS: in the journal *Substance Use and Misuse*.

# REFERENCES



- Colón, H. M., Robles, R. R., Sahai, H. (2001). The validity of drug use responses in a household survey in Puerto Rico: comparison of survey responses of cocaine and heroin use with hair tests. *Int J Epidemiol.* 30(5):1042-9.
- Fendrich, M., Johnson, T. P., Wislar, J. S., Hubbell, A., Spiehler, V. (2004). The utility of drug testing in epidemiological research: results from a general population survey. *Addiction* 99 (2):197-208. Harrison, L. D., Martin, S. S., Enev, T.,
- Harrington, D. (2007). *Comparing Drug Testing and Self-report of Drug Use among Youths and Young adults in the General Population* (DHHS Publication No. SMA 07-4249, Methodology Series M-7). Rockville, MD: Substance Abuse and Mental Health Services Administration, Office of Applied Studies.
- Hornik, R., Jacobsohn, L., Orwin, R., Piesse, A., Kalton, G. (2008). Effects of the National Youth Anti-Drug Media Campaign on youths. *Am J Public Health* 98(12):2229-36.
- Magura, S., Kang, S. Y. (1996). Validity of self-reported drug use in high risk populations: A meta-analytical review. *Substance Use & Misuse* 31(9), 1131-1153.

- Magura, S., Laudet, A., Goldberger, B. (1999). Increasing the validity of drug abuse research through hair analysis. In: Mieczkowski, T., Sunshine I., eds. *Drug Testing Methods: Assessment and Evaluation*. FL: CRC Press.
- Magura, S. (2010). Validating reports of illegal drug use to evaluate national drug control policy. *Evaluation and Program Planning* (33): 234–237.
- United States Government Accountability Office. (2006). *ONDCP Media Campaign: Contractor's National Evaluation Did Not Find That the Youth Anti-Drug Media Campaign Was Effective in Reducing Youth Drug Use*. Washington, DC: GAO.
- Walters, J.P. (2006). Comments from the Office of National Drug Control Policy. In: United States Government Accountability Office . *ONDCP Media Campaign: Contractor's National Evaluation Did Not Find That the Youth Anti-Drug Media Campaign Was Effective in Reducing Youth Drug Use*. Washington, DC: GAO.
- Wholey, J. S. (2004). Evaluability assessment. In: Wholey, J.S., Hatry, H. P., Newcomer, K.E., eds. *Handbook of Practical Program Evaluation*, 2nd ed. San Francisco: Jossey-Bass.

**CORRESPONDENCE TO  
STEPHEN.MAGURA@WMICH.EDU**