The Truth About Numbers:
Interpreting the DHCA Neighborhood Survey

The Alliance to Improve Emory Village has several concerns about the credibility of the "neighborhood" survey the Druid Hills Civic Association has conducted. As noted below, the survey violates nearly every principle of survey research, is fraught with bias and error, and the credibility of its findings regarding generalization to the neighborhood as a whole is suspect.

As Charles Backstrom and Gerald Hursh-Cesar note in one of the leading textbooks on Survey Research, "Survey research differs from informal techniques because it tends to be a more systematic and impartial means of gathering information. Since, in a survey, only a few people must represent the views of many, it is essential to keep our personal feelings out of the process of selecting the respondents. The survey is thus a formal procedure, a way of getting information that is more or less insulated from the personality—the values, beliefs, predispositions—of the researcher."

In late November the DHCA sent out their Emory Village survey by mail to approximately 4,000-4,500 households in the Druid Hills area (numbers provided by DHCA officials). As Backstrom and Hursh-Cesar note, "a major disadvantage of mail surveys is that response rates are notoriously low, and such surveys are often biased toward people with great interest in the issue who are more likely to respond." According to the figures posted on the DHCA web site, 426 surveys were returned, though the total number of responses varies widely by question, falling as low as 352 on one question.

Calculating a response rate for the DHCA survey is difficult because we do not know how many surveys were actually mailed out (though assuming the 4,000-4,500 numbers are accurate that would yield a response rate of about 10 percent, which would place the DHCA survey at the very low end for a general mail survey according to several leading textbooks on survey sampling). Leslie Kish, author of the most widely used textbook on sampling, *Survey Sampling*, notes that response rates for mail surveys range from "less than 10 to almost 100 percent." He adds that "**low responses to one or even two mailings should not be accepted because they often contain severe selection biases**." (emphasis added).

An additional concern is that the DHCA neighborhood survey was also distributed at a public meeting held by the DHCA on November 28, raising the possibility that some individuals may have submitted more than one return.

Both forms of survey distribution—mail and distribution at a public meeting—yield what is known as "volunteer samples." That is, since the researcher does not control who returns a survey we end up with a nonprobability sample, one in which the researcher does not know what the chance is of each respondent being selected for the sample. As Backstrom and Hursh-Cesar note, with volunteer samples "obviously the resulting

sample consists mainly of those people intensely interested in the subject, and it is probably quite unrepresentative of all listeners, readers, or constituents. As in all nonprobability samples, **we cannot generalize the results of volunteer samples to whole populations**." (emphasis added). A general rule in survey research is there should be less sampling error (i.e., a more representative sample) when a larger proportion of the population is surveyed, though even in large samples other sources of error may be introduced through the survey process that bias the survey results.

What is especially problematic about the DHCA survey is that we know absolutely nothing about the characteristics of the respondents. Every credible survey reports a table that compares the characteristics of the sample (e.g., age, race, gender, education, income, housing type, etc. of the respondents) to the population to whom the survey is intended to generalize. By comparing the sample profile to the population profile, one can better gauge the extent of survey error, and in some instances, statistically correct for the sampling imbalances to obtain more reliable estimates of the population characteristics of interest (i.e., in this case opinions about the revitalization of Emory Village).

For instance, we do not know who responded to the DHCA survey. Were they residents of Druid Hills? Were all sections of the Druid Hills neighborhood proportionately represented in the survey returns or were some streets over-represented and others under-represented in the sample? Were respondents home owners or renters? Long-term residents or short-term residents? Frequent or infrequent users of Emory Village?

In addition to significant bias and error introduced by the sampling plan and the means by which the survey was administered, there is substantial error in the DHCA survey as a result of question wording. As Herbert Weisberg, Director of the Center for Survey Research at Ohio State University, reports in his recent book, *The Total Survey Error Approach*, there are five characteristics that good survey questions should have: "that the questions be consistently understood, the questions be consistently administered, the questions consistently communicate what is an adequate answer, all respondents have the information needed to answer the questions, and respondents be willing to provide the answers."

As you well know, the Emory Village revitalization plan and its accompanying zoning overlay and design guidelines, are extremely complex. Many of the elements, such as building height, parking, open space, and developable land, are closely interrelated with one another and also vary substantially depending upon specific land parcels within the zoning overlay boundaries. Though the DHCA did include a "summary table" with their survey, the information was presented in tabular form with no accompanying narrative summary, making it highly unlikely that a "typical" resident would "have the information needed to answer the questions" and that the questions would be "consistently understood."

Unlike a face-to-face survey where the interviewer can interact with the respondent by providing clarity to questions that are not clear or fully understood by the respondent, a

mail survey is self-administered meaning that respondents themselves read and answer the questions. The self-administered nature of mail surveys means that they are particularly prone to error because of the respondent's misinterpretation of the question or possible responses, which research has shown, substantially biases the survey results. Weisberg points out that "people will answer questions even when they do not know the meaning of the terms, so it is important to make sure the words in the question can be understood." He adds that "researchers often overestimate the familiarity of the people with the topic they are studying, and as a result they word questions over the heads of respondents."

In summary, the AIEV encourages residents and others interested in the revitalization of Emory Village to heed Weisberg's conclusion that "error is inevitable in all research, but what is important is recognizing its existence and properly taking it into account when interpreting results." When the research process does not adhere to strict procedures, as in the case of the DHCA survey, the integrity of the data gathered is compromised and one must keep that in mind so that we temper our reliance on the resulting data.