THE UNIVERSITY OF MICHIGAN
INSTITUTE OF SCIENCE AND TECHNOLOGY
Acoustics and Seismics Laboratory

Final Report

RESEARCH HOUSE INTERVIEW EXPERIMENT ON APPLIANCE NOISE

Norman E. Barnett
Bruce E. Erickson
Robert Hefner
Stuart J. Schultz
Donald B. Thrasher

ORA Project 04790

under contract with:

WHIRLPOOL CORPORATION
RESEARCH LABORATORIES
ST. JOSEPH, MICHIGAN

administered through:

OFFICE OF RESEARCH ADMINISTRATION ANN ARBOR

July 1966

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF TABLES (Concluded)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Concluded)

ABSTRACT

An exploratory psychophysical experiment dealing with the bothersomeness of appliance noise in the home is reported. Rigorous survey techniques were used to contact respondents in the vicinity of Ann Arbor, Michigan and a subsample of 192 respondents participated in an in-depth interview conducted in the living room of a normally-furnished house. They responded to 150 appliance-like acoustical stimuli and a variety of questions relating to appliance noise in home environments. Moreover, standard audiometric tests were performed on these respondents.

The successful accomplishment of this experiment demonstrates the practicality of obtaining precise psychophysical data away from artificial laboratory environments and an appropriate methodology is described. One major result is presented in graphical format relating cumulative bother engendered by appliance noise with acoustical level and sound type. The response to a white noise stimulus was very different from the response to more realistic appliance-like stimuli. Other reduced data relate demographic and other variables of our sample with national or local census data, and demonstrate the meaningfulness of acoustic inquiry to the public, the consistency with which the public can respond to acoustical questions and that physically-similar appliance-like sounds are significantly different to the public. Loudness is demonstrated to be an insufficient psychoacoustic function to explain the bothersomeness of stimuli for the frame-of-reference employed. Finally some insight is provided into a variety of parameters which contribute to the public's dislike of appliance noise.

1. INTRODUCTION

This is the fourth technical report originating from a continuing research program concerned with consumers' subjective evaluations of sounds produced by home appliances. It is concerned primarily with that portion of the research which was described as Phase Two in our original research proposal (see Reference 1). As such, it concentrates on the description of a research house interview experiment, the methodology employed to collect and analyze the psychophysical data, and the presentation of the results derived from this exploratory research. The present report is addressed strictly to the Research House Interview Experiment and completes the reporting of experimental studies to date.

The previous three technical reports (References 2,3, and 4) provide much of the background information which ultimately led us to undertake the present experiment. In addition to these formal technical reports, several verbal reports of progress throughout the program have been made both to members of Whirlpool's Research Laboratory staff and, at larger formal briefing meeting, to Whirlpool's Directors of Engineering. These verbal presentations have permitted more detailed justification of underlying research philosophy, as well as certain arbitrary decisions, than can be indulged here.

Another type of comprehensive reporting has been accomplished. During late June 1964, two members of Whirlpool's Research Laboratory staff spent a week in our offices learning the details of the digital data processing applied to the research house experiment. These data constitute a veritable library of psychophysical information which can be researched on a multitude of questions not encompassed by the present program reported here. The verbal transfer of how to read our code sheets, the analysis procedures, and the computer output format was the only feasible way to communicate this information to Whirlpool.

1.1. STATEMENT OF PURPOSE

The central problem of the entire research program is consumer subjective response to the sounds produced by home appliances under the natural environmental conditions encountered in homes. The ultimate goal is to provide the sponsor with adequate quantitative information from which to establish realistic design and quality-control noise-criteria for his products.

In order to proceed within a convenient conceptual framework, the appliance noise problem has, from the outset, been formulated in a functional equation (see Reference 1). The dependent term, consumer subjective response to appliance noise, was equated to the continued product of three independent functions, namely, the physical-acoustical characteristics of the appliance, the architectural-acoustical transfer function, and the psychophysical response function of the person hearing the appliance noise.

The research house study, reported here, was aimed at exploring the psychophysical term of the functional equation, the term about which least is known. Previously reported portions of this research program have provided relatively large amounts of information about the other two terms in the equation. Those terms were easier to investigate because a larger fund of basic knowledge about them already existed both in the form of practical experience and of reported research.

Building upon the experience gained from our first exploratory psychoacoustical experiment (see Reference 4), in which interviews were conducted in the respondents' own homes, we have undertaken a companion exploratory experiment using a rented, furnished house. It preserves a home-like environment which we consider essential and, in addition, it permits a more positive control over many of the physical parameters of the experiments.[1]

We are reporting here, consistent with the goals stated in our research proposals, tentative conclusions on the following topics:

(1) The relative nuisance values of extraneous noise in a home for a sample of the American population.

(2) The relationship of such subjective judgments to the characteristics of the sounds and their origins.

(3) The consistency of respondents in reporting their subjective reactions to sounds in a scaled manner.

(4) The relationships of these judgments to deductively selected variables dealing with demography, financial status, and psychological constitution of response groups.

The interpretive applications of these reported results are subject to important reservations deriving from the sample actually participating in the experiment. These reservations are so fundamental both to interpretation of these results and to the planning of future research that they are discussed in Section 2 before presenting the other details and results of the study. The collection of psychophysical function data based on a large national sample

[1]A journal article is now being prepared which presents the philosophy of a research house interview approach to psychophysical investigation of environmental problems. It is planned for submission to Public Opinion Quarterly.

was beyond the scope of our present research program but such corroborative data are essential before information of the type reported here may be applied confidently to a national market situation.

1.2. GENERAL OUTLINE OF THE EXPERIMENT

The final form of the research house interview experiment, with data collection occurring during July and August 1963, can be explained in broad outline with the aid of Figure 1. Beginning at the top of this diagram, a random sample of residents was drawn in a specifiable manner from the population in and around Ann Arbor, Michigan. The detailed description of sample selection as well as other aspects of sampling are discussed in Section 2. The majority of the addresses constituting the sample were contacted by our field interview staff and the field interview questionnaire administered (see Appendix A). The rationale for the questions used in this questionnaire is discussed in Section 2. Essentially, however, the field interview questionnaire is an abridgment of a portion of the interview used in our earlier experiment in the respondents' own homes and reported in Reference 4. It starts with selected questions of a general nature, divulges our interest in appliance sounds, and invites the respondents to participate in the research house interview.

Those respondents, who were willing to participate further, were scheduled for the research house interview. At the research house, they were asked to complete another questionnaire concerning the specific appliances which they owned and the location of these in their house (see Appendix A). Pure-tone audiometric tests were performed on each respondent at some convenient point during the research house interview. A psychological frame-of-reference was described to the respondents, that is, they were asked to make their judgments about our sounds as if these sounds came from an appliance performing a useful task in our research house. This frame-of-reference played a rather crucial role in the design and conduct of this experiment and it is discussed in Section 3.3.

The major portions of the research house interview, called Part I, Part II, and Part III, consisted of presenting to our respondents for their judgments a sequence of sounds having a variety of frequency characteristics, intensities, and coming from several locations within the house. Two different types of acoustical signals were presented. The first type of signals lasted about 20 seconds each and were referred to as <u>absolute</u> stimuli. The respondent was asked to make an independent judgment about the degree of bother associated with each such stimulus. The second type of signals consisted of five-second bursts of paired, stimuli presented in an A - B - A time sequence and referred to as <u>ABA</u> stimuli. In this case, the respondent was asked to make a relative judgment about whether A or B was most bothersome. For convenience, all of the associated questions were organized into a three-part questionnaire (see Appendix A).

SAMPLE: Drawn by Survey Research Center
Approx. 675 households

FIELD SAMPLE: Field interview; general questions plus request to come to research house

Approx. 472 respondents

RESEARCH HOUSE INTERVIEW: (2 hours)

Approx. 192 respondents

1. Floor plan of own home with appliance locations

2. Audiometric test; 6 frequencies; each ear

3. Instruction in frame-of-reference

4. Acoustic Questionnaire
   Part I : 60 absolute judgments
   Part II: 40 ABA judgments

----------- coffee break ----------

   Part III: 20 ABA judgments
   30 absolute judgments

ARCHITECT'S ESTIMATES:

Approx. 150 houses

Value of respondent's house

Type of house construction

Neighborhood noise

Figure 1. Outline of research house interview experiment.

The acoustical stimuli were assembled in an appropriate time sequence on a two-channel magnetic tape which ran about one hour, 20 minutes in length. However, with all features included, the research house interview required a minimum of two hours for successful completion. Our schedule was arranged to have several respondents take the interview simultaneously and the interview sessions were distributed approximately equally among morning, afternoon, and evening.

The research house itself was a wood-frame ranch style building with full basement, moderate in value, and located in a quiet urban neighborhood. The house was furnished in a normal manner and the respondents sat in the living room, comfortable and relaxed, during the interview.

During the data analysis activity following the research house interviews, it was decided that some supplementary information about our respondents' own homes and neighborhoods would be helpful. To this end, the staff architect visually inspected the exteriors of the respondents' homes and estimated house value, construction type, and neighborhood noise: the lower box in Figure 1.

Thus Figure 1 represents, in very condensed form, the structure of the research house interview experiment. The following individual sections of this report develop the nature and limitations of the sample of people used, the background and concepts leading to this experiment, the details of the various portions of the experiment, the methods employed for data analysis, and, of course, the analysis and interpretation of the results.

1.3 NATURE OF RESULTS OBTAINED

Throughout the interview sessions at the research house, some form of the word bother was used in phrasing questions to the respondents about their reactions to the acoustic stimuli. (A subsample of 17 respondents received the same questions except that the word annoyance was substituted for bother.) Bother was selected as an appropriate term which would be understood by the respondents but one which implied a minimum of negative and/or secondary connotations. Other terms, such as annoyance and noisiness have been employed by different investigators and consequently such terms have acquired vague but delimited scientific definition. Our intent in using bother was to select a neutral terminology and then permit the reduced data to divulge from its functional characteristics if any of the more technical terminology would have been sufficient in the mathematical sense. In particular, loudness, a specifically defined acoustical engineering function, was tested in this manner and found insufficient.

It was found possible, within the limits of sample, frame of reference, and types of sound stimuli employed, to represent our results in graphical format with the intensity level of the stimulus as abscissa, per cent of respondents as ordinate, and degree of bother as parameter. The curves appear S-shaped ranging from most of the respondents reporting not bothered at some relatively low intensity level to extremely bothered at some correspondingly high intensity level. Often one or both ends of these curves are poorly defined because the goals of this experiment made it unnecessary and undesirable to cover the entire intensity range of the stimulus. Also, it was necessary in this exploratory study to aggregate many additional parameters, such as loudspeaker location, in order to enlarge the effective sample size but this aggregation process may also have carried along some unrecognized diverging trends.

The graphical format, mentioned above, was selected as probably being most directly interpretable by Whirlpool in terms of the ultimate engineering application for this type of information. The S-shaped curves are somewhat different from one stimulus type to another, at least, for several of our sounds. When, however, these same data are submitted to a variance analysis using a metric for the degree of bother and the corresponding fiducial limits are taken into account, it is demonstrated that, to a highly significant degree, the functions for the several sound types are, in fact, distinctly different.

It is interesting to note that our results using white noise as a stimulus indicate clearly why this particular type of stimulus has enjoyed a preferred status in laboratory-type psychoacoustical experiments. Its use results in a steeper and more clearly defined function; the scatter among data points is smaller and the consistency of responses the highest. Apparently it is easier for people to form definite opinions about white-noise stimuli than for many other types of sounds. However, within the scope of stimulus types investigated in this experiment, the function for white noise is an extreme one; the most different of all from the corresponding functions for realistic appliance sounds.

An analysis of variance using appropriate metrics also demonstrates that replacement of the physical intensity level of the stimuli by the corresponding calculated loudness of the stimuli is not sufficient to coalesce the results. The loudness calculations, which primarily make adjustment for spectral distribution, yield a tighter grouping of results but only account for about one third of the magnitude of the observed differences. Thus loudness, as currently defined, is certainly not the same as bothersomeness within our frame of reference. It appears that the routine calculation of appliance noise in terms of loudness for quality control purposes is a waste of effort in the majority of cases; the original underlying physical measurements would serve as well.

The apparent locations of the sound as well as the spectral filtering which occurs, say, as the sound propagates from basement to living room, both appear to affect the bothersomeness function to a moderate degree. More work is necessary, however, to fully understand the details.

The high consistency of the respondents, both with respect to absolute judgments and ABA comparisons, indicates that our questions about bothersomeness were conceptually meaningful to them. Thus functional relationships of the type we were seeking can be found within accuracies statistically describable. Nevertheless, there exists a significant minority of respondents who give inconsistent answers. Thus, in each situation, it appears necessary to possess a rather complete knowledge of the distributions so as to be able to make appropriate calculations for locations well out on the tails of the distribution curves. Finally, it is possible to demonstrate relationships between the bothersome response and certain deductively selected variables such as education level. In the present experiment, these relationships were not revealed clearly by the powerful routine techniques of regression analysis; only tedious hand manipulations of the reduced data were found effective. This situation is thought to have resulted from a conglomeration of factors such as possessing little or no a priori idea of what function, magnitude or scaling to expect, the degrading effects of missing data, and the general muddiness of data from a first stage exploratory experiment.

Several such deductively selected variables have been so investigated. Level of education, spontaneous appliance noise complaint and location of washing machine have been found to be clearly related to the bothersome response while income, hearing loss, and number of respondents at a session appear insignificantly related. (see section 4.4). There are numerous other variables of similar type which have not yet been subjected to a sufficiently detailed scrutiny to confirm or reject functional relationship to the bother response. These less completely investigated variables may not be discarded yet. Further analysis of the existing data, however laborious, can provide answers. Future experiments can be devised, or even the raw data from the present experiment can be regrouped, so that regression analysis will clearly display the functional dependences. Variables, of the type considered here, do not seem as powerful as, say, sound intensity or frequency distribution, but their aggregated effect can account for a significant part of the observed variance. Thus it remains important to acquire complete information about them so that the principal psychophysical functional relationships can be confidently formulated in terms of all consequential parameters.

1.4. FUTURE RESEARCH OF THE RESEARCH HOUSE TYPE

From a study of this research house experiment, several refinements and extensions applying to future research have become evident and a few of these will be mentioned below. However, the more general scientific question of future research appropriate to the central problem of realistic noise criteria has a much broader scope and has been excluded from this report. The several ideas set forth here apply specifically to the conduct of a similar research house type of interview experiment.

The present research house interview experiment was exploratory and consequently involved features and inefficiencies that should not be propagated into future experiments on this same subject. The general approach, involving a fairly small, random, but statistically significant sample, is thought to be valid for investigating environmental problems. It seems to avoid many of the uncontrollable biases which plague laboratory investigations in this area although a specific correlative laboratory experiment is needed to verify this point. Of course, this type of experiment can itself be more highly developed as additional experience is gained in applying it's methodology. Some specific recommendations are as follows:

1. A full-time manager should be employed to handle the scheduling and the myriad of details involved in operating a research house. Such assistance would permit the scientists involved to concentrate their attention on their appropriate technical problems instead of diluting their efforts.

2. The research house interview should be shortened. The two-hour duration of the present experiment was burdensome to respondents. Future experiments can be directed at fewer and more specific acoustical problems and tailored to roughly an hour or at most one and one-half hours in duration.

3. Once the initial exploratory experiment has been accomplished, the complete format of the interview, the method of coding answers, and the analysis techniques to be employed should be completely worked out before starting to interview. Coding of interview data and certain of the simpler statistical data reductions should progress simultaneously with the interviewing in order to shorten the time interval between conception of the experiment and final reporting.

4. The appropriate physical characteristics of the research house should be completely investigated and all calibration factors should be available for immediate complete reduction of physical data before starting to conduct interviews.

5. The method of generating the sample appears satisfactory. However, extreme emphasis should be placed on maintaining the highest possible response rate. Field interviewers, just as salesmen, vary widely in their effectiveness and only the best should be used for this work. Training and selection of new interviewers ought to be relegated to some less demanding experiment.

6. Since multiple regression analysis techniques are to be applied, missing data causes much wasted effort and drastically reduces the effective sample size. In this case also, extreme emphasis must be placed on securing complete data from each respondent.

7. An improved method of respondent scoring is needed to verify that each answer is correctly correlated with its stimulus, and that all respondents at a session have answered completely before proceeding to the next stimulus. The numbered response form used in this experiment (see Appendix A) allows purely mechanical errors to occur and some elaboration such as push-button scoring and a rapid start-stop tape machine for stimulus presentation might be better.

8. A rented house without physical alteration was employed in the present experiment, however, certain limitations were encountered which it would be desirable to eliminate. Either a house specifically designed and constructed for such experiments (and perhaps adaptable to other environmental studies later) or a contractor's development house adapted for research purposes and later offered for resale might provide a superior research facility. It would be desirable to have the house constructed from acoustically pretested materials to yield more information about the architectural acoustical function. Likewise, a rented house usually does not provide a convenient traffic pattern for respondents and experimenters, for concealment of wiring, for location of instrumentation out of sight of the respondents, and for independent control of background noise.

9. In at least some future experiment, it may be wise to broaden the scope of the inquiry in order to avoid an over-concentration of the respondent on the acoustic aspect.

10. The stimulus program used depends, of course, on the particular problems under investigation. The nature of such future problems is relegated to a future report. However, the present experiment has demonstrated a definite need for high-quality tape recording equipment and sound-dubbing facilities. Commercial facilities such as are used by the radio and recording industries do not appear to have completely adequate equipment, thus probably ruling out the commerical preparation of the master tape programs. However, numerous commercial components may be used in the assembly of a satisfactory master-tape facility. Such a facility appears most desirable if another experiment is to be undertaken and absolutely essential if a series of such experiments are planned.

2. THE NATURE OF THE SAMPLE: IMPLICATIONS FOR THE CURRENT SURVEY AND FUTURE RESEARCH

2.1. INTRODUCTION

Due to the special complexities arising in connection with this survey, the sampling problems, which even in ordinary circumstances require considerable attention, were amplified. In the process of surveying one is attempting to select a small portion of a total population which is representative. The smaller the sample, the more precarious and limited are the interpretations of the data; however, the larger the sample, the more costly and unwieldly the process of surveying becomes. The attempt is always made to compromise to best advantage within the possible extremes.

Since the Whirlpool Research House Study was fundamentally a large scale pilot test, it was in concept never intended that data reported be interpreted as expandable to the entire American public. The sample was restricted to areas within a workable proximity of the research house location in Ann Arbor. The sample, therefore, is restricted to Ann Arbor and environs (which included nearby parts of rural Washtenaw County).

No U.S. county could possibly be representative of the nation in all relevant social and psychological characteristics since the prevailing climatic, cultural and economic features vary widely in different sections of the nation. The area chosen was selected only on the basis of convenience for research personnel and no even vague attempt at national representation was made. The sample area is in fact the center of a university complex and is, therefore, especially atypical. Although the sample was successfully relieved of the institutional population (college and university students); it must be recognized that this is an area of disproportionately high income and educational levels. In addition, selection of a sample which ignored tracts where students resided also to a large degree reduced the number of youthful (age 18 to 24) respondents interviewed.

As the reliability of survey date is contingent upon achieving a high percentage of successfully completed interviews from the original sample selected, every attempt was made to achieve a high response rate. The degree of success has an effect on sampling error which must be minimized if the sample is to predict correctly about the population it was selected from. We not only desired contact and field interviews; but also to entice interviewed respondents to participate further with a research house interview. Such a two part technique is called a "panel interview" and such a type of interview necessarily involves a lower response rate as some mortality between the two parts is to be expected. Not only was there a two part interview, but we requested that the second portion take place in a place outside and usually some distance from the respondent's home. In this type of interview it is important that the data be analyzed to indicate any bias that may exist because of respondents refusing the second portion.

Specific problems, not present in most surveys, were also present in this one. It is to be remembered that this program was multipurposed; not only did we wish to gain a sample from which survey data results were extractable, but we also wished to gain insight into the special problems involved in determining psychoacoustic responses from survey samples.

2.2. SAMPLE DESIGN

Two separate sample selections were drawn by experienced members of the Survey Research Center Sampling Section. Sample one, in June, was primarily utilized to conduct free interviews with very limited structure designed to help construct a better formal questionnaire. The nature of the interview was to depart completely from previous studies; it was felt such an approach was sensible to insure later success. Here we were attempting to define techniques for evoking and recording reactions of sound perception. To our knowledge, and after considerable research, this is a procedure never before utilized in a rigorous survey design.

In order to maximize the use of time in the relatively expensive program the numbers of respondents were bolstered during the early phases of the full scale July-August program by the addition of respondents from two sources outside the expressly selected July Sample List. These were from the June sample who were agreeable to further interview and some selected by a house to house interview program with no specified sample list. The former case represents a good sample; but not one directly comparable to the major July sample. The latter is a serious disregard of precision sampling technique and was used to a very limited extent during the earlier part of the summer when respondents were not otherwise available for our research house. Despite the limited number of respondents involved, a few other exceptions are noteworthy. In some instances more than one respondent from a field interviewed household arrived for the research house interview. The second was either the spouse or in rare cases a mature offspring. In a rigorous sense, only one of these persons is acceptable within the sample design and it is always the person interviewed in the field who is assigned priority. Field interviewers were instructed to interview the wife wherever present and possible but could accept the husband in some instances. A final exception are those respondents who returned from the long depth interview of the Fall of 1962. It was hoped that the social and psychological information could be used more fully in connection with better acoustic and audiometric data. This group, though selected with an original strict sample design are also not directly comparable to the July sample.

Comparisons within these groups and in connection with Census data on a National and Ann Arbor basis are useful in determining the limitations to be placed upon interpretation of results. There are two major response classes, quality and non-quality. Quality indicates priority respondent from the selected July-August sample; non-quality are all others. This really forms four major groups of sample strata to be considered; quality and non-quality interviewed in the field; and those portions of these two groups who attended the research house. Combinations of the above especially comparisons of the full field sample (quality plus non-quality) and full research house sample are also discussed as widespread use of such groups is made in the data analysis. The related tables follow the text directly.

	Frequencies of Sample Strata		
	Quality	Non-Quality	Total
Research House	121	71	192
Field	188	92	280
Total	309	163*	472

*Components of the total non-quality sample:

Not selected from a specified sample listing:	25
Selected from June sample listing:	109
August reinterviews:	12
Non-priority interviews:	17
Total	163

The discussion will first relate the internal elements of the sample to each other, then relate various sections and the overall sample to national and regional data; and last, discuss the implications of the sample for interpreting the data presented in terms of reliability and sampling error while also noting what types of sample results might be expected if the same approach were applied to a significantly larger number of cases. In the case of each type of comparison, the relative importance will be noted at the outset.

It is recommended that prior to interpretation of the tables in this section one becomes familiar with the footnotes that follow Table 1; many of the footnotes on the other tables represent repetitions and refer back to that table for the sake of brevity. Tables have been organized for maximum efficiency in terms of presentation; however, for some of the more complex comparisons it will be necessary to refer to columns on more than one table.

2.3. ANALYSIS OF THE CONTRIBUTIONS OF THE VARIOUS SAMPLE STRATA TO THE OVERALL SAMPLE

Examining the differences between the quality sample resulting from interviews conducted at the home of the respondent (the field sample) and the same interviews of the non-quality sample is of major import. The addition of the 163 non-quality cases would substantially enlarge the total number of cases to be utilized in the analysis of the field material. Such an enlargement enables a more detailed and possibly informative breakdown of the data since it allows for larger cell sizes for use in more detailed tabular presentation of variables and also introduces greater stability into statistical estimates. However, the quality sample has been randomly selected with a high response rate achieved in order to insure representativeness. Thus this data may be trusted for fidelity whereas no such assurance is present for the non-quality respondents. If non-quality cases do differ substantially, such peculiarities must not be ignored if the faithful representation of the population by the sample is required. Knowledge of the nature of such differences enable the use of the larger sample size while being aware of and compensating for the non-representational aspects which have been introduced.

The non-quality field sample has several elements within it that are exerting forces in different directions. The 25 respondents selected without a sample list turned out to be an extremely homogeneous group. This might have been expected, however, since for ease of interviewing at that point all of those non-sample list respondents were selected from a housing development containing predominately young married university students. The June and July listings avoided specifically housing tracts containing such units in order to exclude the unmarried university population who represent by census terminology the " institutional." Such groups are non-representative of the nation at large and would also make impossible any comparison with national and regional census data which excludes such persons. Interviewers who worked without previously specified samples were free to ignore single students and select the married population only. These 25 respondents were younger than the remainder of the sample and, therefore, had other characteristics more closely associated with younger married couples. This youth was the most important characteristic differentiating the non-quality field group from the quality.

Other groups associated with the non-quality sample tended to reduce differentials. A small number of non-priority respondents originated from the July sample and showed distribution tendencies similar to the remainder of that sample. It too was orginally intended as a representative sample unto itself; it however, is not as representative as the July sample because no attempt was made to keep the response rate of this sample group high. Despite the lower response rate there did not appear to be any marked differences between characteristic distributions. The last group of the non-quality sample is the 12 reinterviews from the Fall 1962 sample where again the 15% response rate resulted from no effort to improve it. This group was selected as a small subsample chiefly intended for a separate small study of the depth psychological interview and did not have any special impact.

Table 1[1] demonstrates the differences between the quality and non-quality sample respondents with regard to demographic distributions. It shows notably the larger percentage of younger non-quality respondents. In addition, educational levels are largely higher while income levels are lower. Table 2 deals with psychoacoustic features and reveals slightly higher sensitivity (possibly due to keener hearing) and a smaller per cent of spontaneous complaints against appliance noise (due to the smaller number of appliances owned).

The non-quality category does not appear to disturb seriously the overall field sample and on the contrary adds a group that would otherwise be extremely undersampled. The importance of this group of young married respondents is not to be underestimated for this type of study due to the probable role they will play in the consumption of major appliances in the near future.

The reasons for examining the differences between quality and non-quality field respondents are identical for examining the differences between those two groups who later took part in the research house interview. Because this group of young married couples had their field interviews accepted only on the basis of their agreement to attend the research house, the same 25 respondents on a percentage basis play a larger role where the total number of cases involved is less than half the original sample size. As would be expected the youthful effects noted previously in the field are amplified here. The single difference here is a change with regard to educational differences between the quality and non-quality respondents; this will be discussed in another section as it is largely due to the type of quality field respondents who attended the research house from the original field sample.

The comparison of the respondents who attended the rented research house and those who refused enables us to characterize the refusals and to state explicitly what group the research house sample represent. As the non-quality sample is composed of so many diverse elements it is not possible to reveal much about this group in this respect. However, it is important to observe this relationship in the quality group to determine peculiar differences both for the purpose of data presented in this report and any anticipated future research along these lines. Furthermore, comparisons of the overall field and research house sample (including non-quality respondents) is to be discussed because of their usefulness for practical purposes as extensive use is made of the two total samples in other sections of this report.

---

[1] The tables which figure in the discussions of section 2 are organized in sequence at the end of this section starting on page 16.

First, the quality sample will be analyzed in this respect as its far greater theoretical importance for pure sampling differentials are clearer here. The first primary difference between quality respondents who did attend the research house and those who did not was the factor of age. The research house interview did not attract an equivalent number of older persons (age over 64) as is noted in Table 5. It was for this reason that larger families were oversampled. The second primary difference, and of extreme interest, is the larger percentage of the highly educated who attended the research house. This is to be partially expected for all survey approaches, as such persons are more familiar with the value of this research method and are more willing to exert some effort to comply (and in this type of research house interview an abnormally high degree of effort is involved for the respondent). These two primary differences result in the higher income of the research house respondent; education is directly related to income and in addition the lower response rate of older persons reduces the weight of the low income retired. The last demographic difference is the lower response rate of the non-white respondent; this is probably a function of the reluctance of this group to visit a white home in an institutional situation.

With regard to the psychological variables (Table 6) there is a much higher sensitivity noted on the psychoacoustic index of those who attended the research house. This difference is partially explained by income since one of the index component questions concerns the willingness of respondents to spend additional money for quieter appliances and since such willingness is, of course, affected by ability to pay. It is noted, as additional support for the last inference, that in the area of spontaneous noise complaints there are no significant differences.

Comparing the first columns of Tables 1 and 3, and 2 and 4 respectively, enables the analysis of differences between the entire field and the entire research house sample. As noted before this comparison is of practical value for the analysis of this report. Because of the influence of the non-quality sample; both in terms of those who attended the research house and those who did not, the differences between the house and field sample are reduced when examined at this level. The undersampling of older persons in the non-quality field sample reduces the age differential. The 25 non-sample respondents still exert forceful influence on the house sample in terms of a higher per cent of youthful families and the resulting larger typical families. The reduction of differences where present improves comparability but in so doing has reduced the precision of comparison to the entire population. The non-quality sample by mere chance has yielded some benefits beyond larger sample size that was not even hoped for. Because of the lower incomes of these 25 respondents who continually have a great impact on the research house sample, the income differences between the house and field are reduced by the addition of the non-quality respondents. Since the non-quality respondents who attended the research house had high educational levels the addition of them had practically no effect and the comparison of house and field sample for this variable are the same whether or not they are present.

TABLE 1

COMPARISON OF SELECTED DISTRIBUTIONS FROM THE TWO RESPONDENT SAMPLES IN WHICH INTERVIEWS WERE CONDUCTED IN THE FIELD
(Demographic and General Background Features)

Characteristic	All Field Sample	Quality Field Sample[1]	Non-Quality Field Sample[2]
Age			
18 - 24[3]	9%	7%	12%
25 - 34	26	23	30
35 - 44	26	28	21
45 - 54	15	14	16
55 - 64	13	13	13
65 and over	12	15	8
Total	100%	100%	100%
Number of Cases[4]	472	309	163[5]
Per cent of Sample[6]	100%	66%	44%
Response Rate[7]	n.a.[8]	70%	n.a.[8]
Marital/Children Status			
Married, children[9]	57%	55%	59%
Married, no child	26	27	26
Single	15	16	14
Other[10]	2	2	1
Total	100%	100%	100%
Number of Children			
None	44%	46%	40%
One	14	12	18
Two	19	20	17
Three or more	23	22	25
Total	100%	100%	100%
Age of Youngest Child[11]			
Under 2 years	26%	24%	30%
3 - 4 years	21	21	22
5 - 8 years	24	24	22
9 -13 years	16	18	13

TABLE 1 (Continued)

Characteristic	All Field Sample	Quality Field Sample[1]	Non-Quality Field Sample[2]
Age of Youngest Child[11]			
14 years or over	13%	13%	13%
Total	100%	100%	100%
Education			
1 - 8 years	10%	11%	10%
9 -11 years	11	13	4
12 years	33	30	40
Over 12 years[12]	46	46	46
Total	100%	100%	100%
Income[13]			
Under $5000	18%	16%	22%
$5000 - $7499	29	27	33
$7500 - $9999	26	28	23
$10,000 and over	27	29	22
Total	100%	100%	100%

---

[1]The Quality Sample consists of those respondents from our July sample listing who were priority respondents. By priority respondent it is meant that, in cases where more than one respondent per household attended the rented house interview (thus being included in the total field sample), only one respondent was selected. Priority went to the wife when present, otherwise to the head of the household. In some cases where the wife could not be interviewed, while the husband could, the husband was accepted.

[2]Includes second respondents in the household where present in the total sample, Fall 1962 reinterviewees, respondents from the June sample, and non-randomly selected respondents recruited on a house to house basis. (In the last group field interviews were accepted only after the interviewer ascertained that the respondent was willing to take the rented house section of the interview.

[3]The components of this group are as follows: All cases 18-24 plus the few rare cases of those interviewed who were under 18 (these were in all cases non-priority respondents who were offspring of a priority respondent).

[4]These are substantially the same for all characteristics enumerated within this table; however, for clarity of presentation, cases where the information

TABLE 1 (Concluded)

required was not ascertained were deleted. Therefore, the number of cases from one characteristic to another may vary slightly.

[5]The components of this group of sample types are as follows:

June Sample	109
Fall 1962 sample	12
Non-Priority respondents	17
Non-Sample respondents	25
Total	163

[6]These percentages may vary slightly from one characteristic to another due to the deletion of those respondents in each case where information was not ascertained.

[7]Number of sample selections/number of successful interviews.

[8]Due to the inclusion of respondents selected on a house to house basis (see Note 2 above) no meaningful percentage is presentable.

[9]Children under age 21 residing at home.

[10]Includes divorced and widowed (some may have children living at home).

[11]Reported only for those who have children under 21 residing at home.

[12]Excludes those with technical school training who are included in the 12-year group.

[13]The Survey Research Center has discovered that this is a very sensitive subject with respondents. To establish maximum rapport in order to induce participation of respondents in the research house portion of the interview this information was not solicited. Instead, our trained interviewers were asked to use the neighborhood, home and environmental clues in making an estimate of income level. Therefore, interpretation of these distributions is to be made with extreme caution.

TABLE 2

COMPARISON OF SELECTED DISTRIBUTIONS FROM THE TWO RESPONDENT SAMPLES IN WHICH INTERVIEWS WERE CONDUCTED IN THE FIELD[1]
(Psychological Features)

Characteristic	All of Field Sample	Quality Field Sample	Non-Quanity Field Sample
Psychoacoustic Index[2]			
Highest Sensitivity	5%	5%	5%
Moderate - High Sens.	38	35	43
Low - Moderate Sens.	31	33	28
Lowest Sensitivity	26	27	24
Total	100%	100%	100%
Spontaneous Noise and Other Appliance Complaints			
Noise Only	3%	4%	2%
Noise and Other	2	2	3
Other Only	32	28	38
None	63	66	57
Total	100%	100%	100%

[1] Data presented on number of cases, per cent of sample, and response rate on Table 1 are identical in this case. All relevant footnotes (where captions are the same) from Table 1 also apply here.

[2] For a full discussion on the methods employed to construct this index see section 4.4.5.

TABLE 3

COMPARISON OF SELECTED DISTRIBUTIONS FROM THE TWO RESPONDENT SAMPLES IN WHICH INTERVIEWS WERE CONDUCTED IN THE WHIRLPOOL RESEARCH HOUSE
(Demographic and General Background Features)

Characteristic	All of Research House Sample	Quality Research House Sample[1]	Non-Quality Research House Sample[2]
Age			
18 - 24[3]	13%	7%	24%
25 - 34	28	27	28
35 - 44	27	31	21
45 - 54	12	13	11
55 - 64	13	14	10
65 and over	7	8	6
Total	100%	100%	100%
Number of Cases[4]	192	121	71[5]
Per cent of Sample[6]	100%	63%	37%
Response Rate[7]	n.a.[8]	27%	n.a.[8]
Marital/Children Status			
Married, children[9]	66%	68%	62%
Married, no child	21	20	24
Single	11	10	13
Other status[10]	2	3	1
Total	100%	100%	100%
Number of Children[11]			
None	35%	34%	37%
One	17	13	24
Two	25	28	24
Three or more	23	25	21
Total	100%	100%	100%
Education			
1 - 8 years	5%	3%	8%
9 -11 years	8	9	4
12 years	30	31	29
Over 12 years[12]	57	57	59
Total	100%	100%	100%

TABLE 3 (Continued)

Characteristic	All of Research House Sample	Quality Research House Sample	Non-Quality Research House Sample
Number of Appliances Owned[13]			
One - two	11%	2%	27%
Three	16	16	19
Four	20	25	15
Five	25	30	19
Six or more	28	27	20
Total	100%	100%	100%
Age of Youngest Child[14]			
Under 2 years	25%	21%	31%
3 - 4 years	25	23	29
5 - 8 years	24	26	21
9 -13 years	15	21	5
14 years or over	11	9	14
Total	100%	100%	100%

[1]The Quality Sample consists of those priority respondents selected from our July sample who attended the Whirlpool Research Home interview. For a fuller discussion see Table 1, footnote 1.

[2]Includes second respondent per household (non-priority) when present and others attending research house including non-randomly selected respondents, Fall 1962 reinterviewees, and respondents from the June sample.

[3]Includes a few cases under 18; see footnote 3 of Table 1 for detail.

[4]Frequencies may vary slightly due to deletion of not ascertained information; see footnote 4 of Table 1 for detail.

[5]Components of this sample group are as follows:

June sample	17
Fall 1962 sample	12
Non-Primary R's	17
Non-Sample R's	25
Total	71

[6]Percentages may vary slightly due to impact of n.a. deletions.

TABLE 3 (Concluded)

[7]Number of sample selections/number of successful research house interviews. Forty-one per cent of those interviewed in the field attended the rented house interview (includes quality and non-quality sample). For the quality sample only, this fraction is reduced by a very slight quantity.

[8]Inclusion of non-sample respondents allows no basis for presentation.

[9]Children under age 21 residing at home.

[10]Includes the divorced, widowed (some may have children living at home).

[11]Appears to disagree with marital/children status characteristic distribution; this is because of inclusion of some children in "other."

[12]Excludes those with 12 years plus technical schooling (who are included with the 12-year group).

[13]Those who rent are considered to own appliances located directly within the rented unit; excluded, however, are such items as washing machines and dryers located in central utility rooms. (For our definition of an appliance, refer to Field Interview, Page 3, Question 6.).

[14]Distribution of those who have children under 21 living at home.

TABLE 4

COMPARISON OF SELECTED DISTRIBUTION FROM THE TWO RESPONDENT SAMPLES IN WHICH INTERVIEWS WERE CONDUCTED IN THE WHIRLPOOL RESEARCH HOUSE[1]
(Income and Psychological Features)

Characteristic	All of Research House Sample	Quality Research House Sample	Non-Quality Research House Sample
Income[2]			
Under $5000	19%	11%	34%
$5000 - $7499	25	21	31
$7500 - $9999	23	29	14
$10,000 and over	33	39	21
Total	100%	100%	100%
Psychoacoustic Index[3]			
Highest Sensitivity	10%	10%	9%
Moderate-High Sens.	42	43	42
Low-Moderate Sens.	30	30	32
Lowest Sensitivity	18	17	17
Total	100%	100%	100%
Spontaneous Noise and Other Appliance Complaints			
Noise only	6%	6%	6%
Noise and Other	3	3	4
Other only	34	26	51
None	57	65	39
Total	100%	100%	100%

[1]Data presented on number of cases, per cent of sample, and response rate on Table 3 are identical in this case. All relevant footnotes (where captions are the same) from Table 3 also apply here.

[2]Interviewer's estimate; for discussion see Table 1, footnote 13.

[3]For a full discussion on the methods employed to construct this index see section 4.4.5.

TABLE 5

COMPARISON OF SELECTED DISTRIBUTIONS OF QUALITY RESPONDENTS WHO REFUSED THE RESEARCH HOUSE INTERVIEW WITH THOSE WHO ACCEPTED
(Demographic and General Background Features)

Characteristic	All of the Quality Sample[1]	Quality Sample at Research House	Quality Sample not at Research House
Age			
18 - 24[2]	7%	7%	7%
25 - 34	23	27	20
35 - 44	28	31	27
45 - 54	14	13	15
55 - 64	13	14	11
65 and over	15	8	20
Total	100%	100%	100%
Number of Cases[3]	309	121	189
Per Cent of Sample[4]	100%	39%	61%
Response Rate[5]	70%	27%	43%
Marital/Children Status			
Married, children[5]	55%	69%	46%
Married, no child	27	20	31
Single	16	8	22
Other Status[6]	2	3	1
Total	100%	100%	100%
Number of Children			
None	46%	34%	54%
One	12	13	12
Two	20	28	15
Three or more	22	25	19
Total	100%	100%	100%
Education			
1 - 8 years	11%	4%	15%
9 -11 years	13	9	16
12 years	30	31	29
Over 12 years[8]	46	56	40
Total	100%	100%	100%

TABLE 5 (Concluded)

Characteristic	All of the Quality Sample	Quality Sample at Research House	Quality Sample not at Research House
Income			
Under $5000	16%	11%	21%
$5000 - $7499	27	21	30
$7500 - $9999	28	29	27
$10,000 and over	29	39	22
Total	100%	100%	100%

---

[1] Priority respondents from July Sample.

[2] Includes a few cases under 18.

[3] Substantially the same frequency for all details on this table; varies slightly because of deletion of not ascertained groups.

[4] Will vary slightly because of deletion of not ascertained groups.

[5] Children under 21 living at home.

[6] Includes divorced and widowed (some may have children).

[7] Distribution of those who have children.

[8] Excludes technical school.

TABLE 6

COMPARISON OF SELECTED DISTRIBUTIONS OF QUALITY RESPONDENTS WHO REFUSED THE RESEARCH HOUSE INTERVIEW WITH THOSE WHO ACCEPTED[1]
(Psychological Features)

Characteristic	All of the Quality Sample	Quality Sample At Research House	Quality Sample Not At Research House
Psychoacoustic Index[2]			
Highest Sensitivity	5%	10%	2%
Moderate-High Sens.	36	43	31
Low-Moderate Sens.	32	30	34
Lowest Sensitivity	27	17	33
Total	100%	100%	100%
Spontaneous Noise and Other Appliance Complaints			
Noise only	4%	6%	2%
Noise and Other	2	3	1
Other Only	28	26	29
None	66	65	68
Total	100%	100%	100%
Amount Willing to Spend Extra for Appliance Quiet[3]			
Nothing	32%	32%	31%
Up to $10	14	16	13
$11 - 25	30	41	24
$26 - 50	9	8	10
Amount extra n.a.	4	3	4
All data n.a.	11	--	18
Total	100%	100%	100%

[1] Data presented on number of cases, per cent of sample, and response rate on Table 5 are identical in this case. All relevant footnotes (where captions are the same) from Table 5 also apply here.

[2,3] For a full discussion on the methods employed to construct index and questions used to ascertain willingness to spend extra see section 4.4.5.

As for the psychological variables, the research house sample still shows higher sensitivity with the addition of the non-quality group while the spontaneous noise complaint shows only minor differences with complainers only slightly oversampled.

2.4. ANALYSIS OF THE REPRESENTATIVENESS OF THE SAMPLE TO ANN ARBOR CENSUS DATA AND THE RELEVANCE OF THE ANN ARBOR POPULATION TO THE NATIONAL POPULATION

At this point a discussion of the relationship between Ann Arbor population characteristics and that of the entire nation is in order; for this purpose Bureau of Census data presented in Table 7 is used. Following this is a comparison of our survey sample and its various components with the Ann Arbor census data. The Ann Arbor population is most notably different from that of the nation in that it has far higher levels of educational attainment; related to this are considerably higher levels of income. The number of persons aged 18 to 24 is vastly above the national level relative to the other age brackets.

In comparing the survey data with the Ann Arbor census data we can see how well our sampling technique succeeded; while in comparing the survey data with the national data we can see how well our sample relates to the national population. This latter point is of general interest despite the fact that we cannot reliably depend upon our sample to represent the nation.

First, in comparing our field sample with local census data, the following relationships are observed: the youthful section of the population has been seriously undersampled; this is a result of the avoidance of the university population who resided in tracts where most of the younger families in the community also reside. It is a happy event that 35% of our sample is not age 18-24 because this would have made our sample so atypical that it would have little relevance to the national population. It is unhappy, however, that the end product of our endeavors also resulted in the portion of our sample aged 18-24 being somewhat below the national census data. Fortunately, the inclusion of our house to house sample did improve our sample in this regard. It is to be remembered that the major portion of the non-quality sample are the June selections that parallel in sample design the July selections. In all other respects the addition of the non-quality sample to the field makes little significant difference in comparing census and survey data. Educational levels of the field sample were considerably higher than the national population (which is to be expected); however, even in comparison with the Ann Arbor data we appear to have missed the portion of the population lowest on the educational scale. In all other respects our educational distribution appears in good order. It is to be noted here that the Ann Arbor data is for persons 25 years of age and over, whereas the survey data is for 18 and over. This group between 18 and 24

TABLE 7

DISTRIBUTIONS ON SELECTED DEMOGRAPHIC AND INCOME FEATURES OF THE POPULATION ON A NATIONAL AND ANN ARBOR AREA BASIS

Characteristic	National		Ann Arbor Standard Metropolitan Statistical Area[1]
Race			
White	90%[2]		92%
Non-White	10		8
Total	100%		100%
Age			
18 - 24	15%[3]		35%
25 - 34	19		21
35 - 44	20		14
45 - 54	18		10
55 - 64	14		8
65 and over	14		12
Total	100%		100%
Income			
Under $5000	39%[4]		29%
$5000 - $7499	24		28
$7500 - $9999	18		14
$10,000 and over	18		29
Total	100%		100%
Number of Children Under 19 in Husband-Wife Households			
None	41%[5]		Not Currently Available
One	18		
Two	18		
Three or more	23		
Total	100%		
Education			
1 - 8 years	33%[6]	36%[7]	28%[8]
9 -11 years	18	18	16
12 years	30	28	18
More than 12 years	19	19	38
Total	100%	100%	100%

TABLE 7 (Concluded)

[1]All data listed in this column are from the U. S. 1960 census, published by The Bureau of The Census: 1960 Census of Population, Volume 1; Characteristics of the Population, Part 24, Michigan. The term Standard Metropolitan Statistical Area originates with the Census Bureau and roughly includes Ann Arbor proper and adjacent unincorporated rural areas; this is almost precisely parallel to the sample selected in the Whirlpool Study for July.

[2]Dept. of Commerce, Bureau of The Census; Current Population Reports, p. 20, No. 125.

[3]Ibid., p. 20, No. 122; data redistributed to age 18 and over equal 100%.

[4]Ibid., p. 60, No. 41; data are incomes of families and unrelated individuals.

[5]Ibid., p. 20, No. 122

[6]Ibid., p. 20, No. 121; distribution of the population 18 and over.

[7]Ibid., distribution of the population 25 and over.

[8]Distribution of the population 25 and over; distribution for the segment of the population 18 and over not available.

Note: In all cases where national data are used; that which was most currently available at the time of the writing of this report was selected. In almost all cases this was information relevant to the year 1962 as statistical information for 1963 is still being tabulated and probably will not be available until June of the current year.

probably has attended (or completed) high school with greater frequency than has the group aged 25 and over. This is confirmed by observing the two educational distributions presented for the national data. Our sample represents with precision the distribution of families with children and the number of children present.

Second, is the comparison of the research house sample with national census data. Again differences between the quality and overall sample are not large enough to devote much attention to. Here the sample was younger; we undersampled the oldest age group; dividing age up into three population sections: 18-34, 35-44 and 55 or older we find our sample to be a fair representation of the nation at large. The same comments pertaining to age in the Ann Arbor data for the field is relevant here. The income relationship remains about the same as it was in the field for the quality group, the more powerful influence of the 25 house to house sample respondents is sufficient however, to considerably improve the representativeness of the total research house sample. With regard to education, the undersampling effect of response groups having less than a completed high school education has been reinforced. The research house sample slightly overrepresents respondents with children.

Due to various reasons previously discussed, which include the speed with which our operations needed to be carried out, the available interviewing and field staff, the limited sample and the location of our research house with relation to the city of Ann Arbor's various income-residence areas, this sample is in most instances quite some distance from a faithful representation of the Ann Arbor or of the national population. A good sample of our size for the Ann Arbor population, would be representative within 10-20% sampling error. It cannot be presupposed that our sampling error is even that low. This does not mean our data is useless; it means that certain considerations must be taken into account in projecting results to other populations. It can be stated firmly that this sample is a good representation of higher income and educational groups. The field sample undervalues youthful responses whereas the research house undervalues responses of older and retired portions of the population. These effects are even larger in comparing our data to national data than in comparing it with Ann Arbor data. Furthermore, if the comparisons with the total population are less than precise, this does not detract from comparability of groups within the sample. Differences between 10 to 15% are significant between sample groups that involve the entire research house sample. Differences which involved the total field sample are 7 to 10% when they indicate significance. Thus we can analyze differences in the responses of various sample subgroups with some degree of precision but we cannot maintain that the representation of these subgroups in the sample have the same relative weight in the national population (in particular) or to a lessor extent the Ann Arbor population. This lack of representativeness is compounded when response rates decline. The 70% rate in the field is a fairly good rate (relative to other surveys) the 27% in the research house leaves much to be desired.

Larger scale more careful interviewing and sampling technique can eliminate some of the problems involved in this survey. The elimination of the problematic and diverse sampling sub-groups would be of much help too. Greater amounts of time to conduct interviews could also be effective improvements. In a future study, if such were to be conducted; better representativeness, slightly higher field response rate (about 80%) and coupled with this a higher research house response rate (40-45%) could fairly be expected. The inclusion of some audiometric, possibly earphone data, could improve our knowledge of research house refusals and better enable us to map these refusals into a strictly structured presentation to determine what section of the population was being oversampled in the research house with a high degree of precision. All these matters taken collectively detract from the effectiveness of our current survey; but as indicated could be substantially corrected. Providing the data in other sections of this report lies in the directions of your interest and is revealing in terms of the basic problems involved, it is believed with little doubt by this research staff that a majority of problems arising from sampling considerations could be overcome in the future providing the original sample design utilized was of sufficiently large size. Such a size might be a total of 1500 to 2000 field respondents and 700 to 800 research house respondents who attended research houses located in representative areas of the nation. No attempt was made to construct detailed sample error and sample error-of-the-difference tables here because these errors are obviously in the magnitude where such precise data is unwarranted. The more ambitious survey outlined above would include such tables. The number of locations, in terms of research houses required, varies with the population it is desired to represent; needless to say, two is better than one (if different areas are used). However, a rough estimate would say that 3-4 would be the smallest number which would make it considerably more profitable to become involved in such a survey, whereas 8-10 would without any doubt yield excellent results in terms of representing the national population faithfully.

3. DESIGN AND OPERATION OF THE RESEARCH HOUSE EXPERIMENT

The notion of utilizing an actual house as a research laboratory in which to conduct appliance noise studies had been considered as a possibility almost from the initial study phase. A working procedure, however, was not clearly visualized until just prior to undertaking Phase Two (Winter 1963). The aim of this section is to present the gist of the reasoning which led to this research house experiment and to describe sufficient operational detail to document the experiment.

3.1 INTRODUCTION

The idea of utilizing a typical residential house as a research laboratory did not appeal strongly to us at the outset. Through a gradually evolving deductive process, however, the values of using a research house became more apparent until finally they predominated as representing the only feasible way to proceed. Thus it is pertinent to sketch briefly some of the main considerations which led to the decision to employ a research house.

A review of previous investigations into the subject area of psychoacoustics produced the conclusion that those aspects of the appliance noise problem which we wished to investigate had remained almost completely unexplored. In any event, very little information of definitive value could be found to serve as an effective guide for our studies. Thus it was evident that our research would have to rest heavily upon inductive experimental methodology.

Before attempting to devise the primary experiments, it seemed wise to gain insight into the answers to two questions: what percentage of the American population was bothered by household appliance noise, and, how significant was the degree of bother expressed by the group complaining? To this end, a brief questionnaire on appliances generally, and appliance noise specifically, was included in a national sampling survey conducted by the Institute for Social Research at The University of Michigan. The detailed results from this survey have been reported in Reference 2. In substance, it was found that a significant degree of bother due to appliance noise was encountered relatively infrequently, but that such complaints did not seem to originate in that segment of the population composed of systematic complainers. It was concluded that the existing survey methodology, relying on purely verbal questionnaires, was useful in the present case to gauge the general attitude of the American public but was insufficient to get at the root of the problem. Indeed, the survey results suggested that considerable latent dissatisfaction with appliance noise might exist which did appear in response to abstract verbal questioning.

The research team, acting on the basis of considered judgment, rejected the usual psychoacoustical laboratory methods, typified by earphone experiments, as being far too abstract for the problem at hand. Two general experimental procedures were suggested to cope with the problems of evaluating the degree of bother caused by appliance noise. The first of these was to take the experiment into each respondent's own home, and the second, to create the experimental situation within a house possessing normal residential characteristics and bring the respondents to this house.

Initially, the first procedure was developed into the field experiment conducted during the latter part of Phase One of this research program and this experiment has been reported in References 3 and 4. This field experiment approach maximized the familiar environmental aspect for our respondents and most respondents reacted easily and confidently to our questions. However, this Phase One experiment also demonstrated a practical impossibility of obtaining correspondingly good physical and architectural acoustic data to compare with the subjective data. Thus when undertaking the planning of the Phase Two studies, we were led to employ the second procedure mentioned above; that of conducting the experiments in a research house which simulated a typical home environment and which simultaneously allowed a considerable degree of acoustical-laboratory control. The research house experiment appeared to afford an optimum compromise for the simultaneous control of both the psychological and physical parameters appearing in our problem.

3.2 INITIAL PHASE

Before full-scale data collection in the research house experiment could be undertaken, considerable preliminary development and testing was necessary. Two major and related aspects of the experiment required specific attention; the perfection of a program of tape-recorded stimuli and a clarification of the psychophysical problems including the development of a specific frame of reference for this experiment.

An experimental pretest was conducted during June 1963 involving forty-some respondents distributed among seven experimental sessions held at various times of day. These respondents also were selected from a random sample of households. The respondents were contacted at their own homes, a short interview was conducted and then they were invited to our research house for the experimental session. Thus most features of the interview process during the pretest phase paralleled those applied during the main research house experiment. (See Appendix A for details of the several questionnaires.)

During the pretest, the response rate obtained in the field was excellent but the loss in number of respondents coming to the research house appeared as a potential problem. A variety of procedural changes were made to improve the response rate at the research house during the main data collection phase and an appreciably larger response rate was achieved. Nevertheless, a relatively low response rate at the research house remains a problem in this type of sampling survey which has not been solved to our complete satisfaction.

During the pretest at the research house, the interview format and the program of stimuli were not rigidly set. Moreover, the respondents were encouraged to volunteer information of any nature at will. (During the main experiment, discussion was strictly forbidden to prevent one opinionated respondent from influencing other respondents at the same session.) Our respondents did seem to feel "at home" in the research house and they were able to make reasonably consistent judgments about the experimental stimuli. From the point of view of survey methodology, it was encouraging to discover that the respondents were genuinely cooperative in a situation which made substantial demands on their time and patience.

3.3 FRAME OF REFERENCE

In the original concept of the research house experiment, the respondents were expected to "role play" by making judgments as if the sounds they heard actually originated from appliances located in their own homes. Thus with this initial frame of reference, our research house was a stage set intended to provide a convincingly realistic home environment conducive to "role playing." During pretest, it quickly became evident that a variety of factors predisposed against effective and completely unbiased "role playing" by the respondents. Ultimately, a somewhat different frame of reference had to be devised for this research house experiment. Several of the problems which forced the change are discussed below.

One aspect of unintentional bias appears to stem from a tendency for people to become adapted to the familiar sounds from their own appliances. Many respondents remembered their own appliances as being virtually inaudible except when standing close to the appliance under consideration. Sometimes, there was almost total rejection of our premise that basement-located laundry appliances could be heard in the living rooms of their homes. In several such instances, one of our researchers returned with the respondent to her home and both the researcher and the respondent (to her surprise) then verified that the basement-located appliances were indeed clearly audible in the living room. Extemporary measurements with a sound level meter made under these circumstances suggested that the remembered levels of appliance noise were about 10 decibels below reality. This situation was found to hold even when the living room was remotely located with respect to the appliances and also for expensive homes embodying premium-quality construction.

Another problem mitigating against a "role playing" judgment was related to the specific brand of appliance owned by the respondent and the characteristics of the sound made by the appliance. Our tape-recorded sounds of a Whirlpool automatic washer were familiar to the owners of this particular Whirlpool appliance but foreign to respondents who owned other makes. A higher degree of bother was invariably evidenced during pretest by those to whom the sounds were foreign.

Another difficulty appears to be the result of provoking a subconscious social bias. The respondents seemed to wish to profess that they owned quality homes and/or appliances such that noise due to their appliances was not as intense as the stimuli we generated in the research house. And indeed, the research house was less substantial and less elegantly furnished than the homes of the more affluent respondents. Nevertheless, physical measurements of the intensity of appliance sounds indicated that our experimental stimuli did occur at appropriate levels although subtle differences in sound quality may have occurred.

An additional confounding factor is the matter of passive listening versus active, attentive listening. The research house experiment obviously induces the latter while most respondents, when remembering what their appliances sounded like, probably were recalling a passive listening situation when their attention had been concentrated on almost anything but appliance noise. In the realm of appliance sounds, active listening probably occurs relatively infrequently such as, when a housewife wishes to know if the laundry load is completed, or if there is audible indication of an impending (or real) malfunction. During pretest, various attention diverting techniques were tried and, while influence on the degree of bother was noted, none of these techniques seemed suitable for the controlled experiment we were undertaking.

There were several other problems entailed in selecting a workable frame of reference for the experiment beyond those associated with "role playing" as discussed above. For example, a time factor entered the considerations in two ways. One factor is the presumed duration of a normal operational cycle. Obviously more sound might be tolerated from an appliance having a short operating cycle or an appliance operated only occasionally compared to an appliance having an extended cycle or operated frequently. A second temporal factor was the length of stimulus presentation needed to reach stable judgments. We presumed, of course, that the stimulus presentations would be shorter than the normal cycle for most appliances if for no other reason than to collect a multiplicity of judgments about different stimuli in an interview of reasonable length. Distantly related to the first time factor is the matter of a respondent's guess about the function of an appliance represented by its sound stimulus regardless of the accuracy of that guess. Thus, for example, if a respondent guessed a sound was from a garbage disposer it might be judged less bothersome than if it were guessed to represent a central heating system.

The background noise level was another factor explored during pretest. The research house had been selected, among other criteria, on the basis of a relatively low neighborhood noise level. One of the original ideas was to require a very low level of natural background noise so that a realistic level of artificial background noise could be introduced in a controlled manner. Explorations were conducted using white noise and other shaped spectra as background noise projected through a loudspeaker system. These background sounds were obviously artificial when raised to levels appropriate to moderately quiet residential neighborhoods. None of the attempts at introducing a controlled acoustic background seemed even remotely satisfactory and so the idea was discarded for this experiment.

It seemed as if an artificial background noise would have to sound all encompassing rather than originating from one or two specific (loudspeaker) locations; a mode of operation beyond the capabilities of our electronic equipment. Furthermore, it seemed as if such a background would have to fluctuate in level, apparent direction and probable origin in order to be realistic. Considerable research and development would have been required to achieve a satisfactory level of perfection warranting the introduction of a controlled background and time limitations did not permit such refinement.

The idea of utilizing a precisely controlled background for the stimuli so as to render applicable a concept such as <u>sensation level</u> remains a valid, and perhaps crucial point for future investigation.

The research house experiment was conducted with the natural neighborhood noise level as a background for the stimuli. Our research house differed from an average home not only by its location in a low noise neighborhood but also because it lacked most of the internal noise such as would be generated by children and normal household activity. Thus, by contrast, stimulus sounds often seemed quite striking against a background of relative quiet.

The frame of reference ultimately adopted treated the research house as representing a home which the respondent might visit but not frequently enough to be completely familiar with it. Being of moderate value in a middle-class neighborhood, a visit to such a house would not be foreign or inconceivable to any stratum of respondents. The respondents were instructed to consider the various sounds (stimuli) as representing the sound from an appliance within the house which was performing a useful household task. That appliance normally would be expected to operate for a longer period of time than the few seconds allotted to each stimulus presentation, however it did not operate continuously. (A more exact designation of operating period, such as ten minutes, was scrupulously avoided.) Under these conditions, was the appliance sound (stimulus) bothersome? The actual questions about bother were framed as given on the questionnaires in Appendix A. In this frame of reference, the respondent only role-played to the extent of imagining a visit to someone else's living room and forming a judgment about degree of bother there.

Perhaps the most cogent justification of this particular frame of reference is that it worked; that is, our respondents accepted it as an understandable situation and were able to produce reasonably consistent judgments. By contrast, the frame of reference first employed during pretest led to confusion. The frame of reference finally selected for the research house experiment must be considered exploratory in the same sense that the entire experiment is exploratory. As this type of experimentation is developed and applied more exactingly, the frame of reference must be developed accordingly.

The selected frame of reference undoubtedly influences the results in ways which are difficult to assess and which constitute research problems in their own right. For example, it is now reasonably certain that the bother functions found for stimulus R (refrigerator), as described in section 4.1 and presented in Appendix B, must have been displaced toward higher values of sound intensity level than are realistic for refrigerators. (We have rather good evidence that consumers complain about refrigerator noise at significantly lower sound levels.) This displacement probably is a consequence of the frame of reference adopted by the majority of respondents, where from the easily recognized automatic washer stimuli, they may have been led to presume a laundry-like function for all of the stimuli rather than a function like that of a refrigerator or of a central heating plant. We conjecture that the general shape of this experimental function remains correct but if it is to be applied _a posteriori_ to an actual refrigerator situation, then a downward translation by several decibels becomes necessary. On the other hand, all of the functions in Appendix B would seem to be properly placed with respect to laundry appliances, automatic dishwashers, and the like.

3.4 TAPE PROGRAM

It was decided that the main portion of this experiment should consist of a series of acoustic stimuli of different sound characteristics, intensities, and apparent location. These would be presented to the respondents for their absolute judgments about the degree of bothersomeness engendered by each stimulus individually. In selecting a final series of absolute stimuli many choices and compromises had to be made, some of them quite arbitrarily.

The length of presentation time for each stimulus constituted a major decision. Household appliances, such as laundry appliances or refrigerators, operate for many minutes at a time. Realistic time durations for such appliance sounds obviously were impractical in this experiment because only a few such stimuli could be presented at each interview session and also the respondents probably would become bored while listening to them. If the presentation time were very short, the respondents might have difficulty arriving at a stable judgment expecially since we wanted the stimuli judged as if they were appliance sounds having appreciable duration. A stimulus duration of 20 seconds followed by 10 seconds of silence for scoring was selected as the basic time interval.

Now that the experiment has been completed, the 20-second stimulus duration seems to have been a satisfactory choice. The question of whether or not respondents can effectively extrapolate their judgment about a 20-second sound to a long real-appliance situation remains unanswered. However, they seemed to accept instructions to make this extrapolation with few qualms. As an interview progressed and respondents acquired familiarity with their tasks, most of them began to record their judgments within 5 to 10 seconds after a stimulus onset and then appeared impatient for the next stimulus. From a purely operational viewpoint, it would appear that some time could be saved by shortening the stimuli partway through the program; there may be more compelling reasons for doing otherwise.

With the time per stimulus selected, the scope of other parameters which could be accommodated in an interview of resonable length became more clearly delimited. Eight types of sounds presented at seven different intensities and from two different loudspeaker locations were selected. All combinations of these three parameters would have resulted in 112 different stimuli. Actually 90 such stimuli were used in the final tape program. Seven of the stimuli were replications so that 29 of the possible combinations were omitted. The details of the tape recorded stimuli are given in Appendix A, Tables 29, 30, pages 138 and 139 .

Naturally, it was desired to experiment with as wide a variety of sound types as possible but the limitation to eight sounds permitted only a start at exploring type as a parameter. Some research on the annoyance caused by sounds has been reported in the Journal of the Acoustical Society of America, starting with the very first volume issued in 1929, and elsewhere. Most of such research deals with artificially created sounds presented in a laboratory environment and, at most, it has contributed only general guidance for the selection of the present stimulus types. The focus of our attention on practical studies of appliance noise dictated that several realistic appliance sounds be included. The sounds from: (1) an automatic washing machine, (2) an electric clothes dryer, and (3) a refrigerator were chosen. These all represent situations in which the appliance sound would be expected to persist for durations of many minutes. Appliances normally presenting more transient sounds such as waste disposers and vacuum cleaners were excluded on the presumption that their acoustical characteristics involved too many additional parameters for this pilot study.

The three sounds selected included an aspect of different controllability. In this sense, the clothes dryer is started at a voluntarily selected time and produces a relatively constant sound; the automatic washer, once started, normally completes a predetermined sequence of different operations before stopping; the refrigerator starts and stops upon demand of its thermostat which, to the casual observer, is an apparently random sequence. From the standpoint of an acoustical spectrum, these three appliance sounds are rather similar, all being predominently low frequency sounds.

Another group of three sounds was selected as stimuli which constitute a progression into the domain of simulated appliance sounds. On the realistic side, the automatic washer sound was filtered to remove its higher-frequency acoustic-quality characteristics. From the artificial side, the same filtering was used to provide a corresponding low-frequency band of random noise. An intermediate stimulus was concocted by superimposing a 120 cps tone on the low-frequency band of random noise.

The eight stimulus types were completed by including two purely laboratory-type sounds which among their other attributes provide a possible bridge between our experiment and the psychoacoustical literature. White noise was selected as one stimulus and a band of noise from 670 cps to 1000 cps as the other (Reference 5).

The abbreviations assigned for convenient reference to these stimuli are given in Table 29, Appendix A.

The eight stimuli were recorded on the master program tape at seven different electrical levels distributed in three-decibel steps, that is, 0 db, -3 db, -6 db, -9 db, -12 db, -15 db, and -18 db. The zero-decibel level was determined as the maximum rms amplitude at which the white noise signal could be recorded with low distortion. (This was the most difficult signal to accommodate as a consequence of the electrical pre-emphasis normal to tape recorder electronics.) All other stimulus types were adjusted to the same rms electrical signal level and then the 3 db steps were obtained in a like manner. Of course, by virtue of adjustable playback amplification, the resulting 18 decibel range of sound pressure levels generated in the research house could be placed at almost any intensity desired. (The washing machine agitate stimulus at 0 db level was adjusted to approximately 65 db SPL in the living room.) All stimulus levels occurring at each seat location for each test condition at the research house were actually measured with sound level meters and these measured levels, rather than nominal levels are associated with each response for much of the data analysis. The variation of stimulus level from one seat location to another is one of the significant complications of the research house experiment.

The 90 absolute stimuli were randomly distributed on the two tracks of the tape and ultimately were radiated from one or the other of two loudspeakers placed in three different rooms of the research house. Our earlier studies (see Reference 4) had led to a concern that respondents might be reacting to the apparent location of the sound source in addition to the amount and character of the stimulus at their ears. The use of these loudspeaker locations assisted investigation of the apparent source location aspect and in all, four different experimental arrangements of the two loudspeaker were used. Roughly one quarter of our respondents experienced each arrangement or mode. One of the four modes was selected for each interview session and not altered during the session.

The combination of loudspeaker locations or experimental modes were as follows:

(1) Kitchen—basement

(2) Kitchen—living room

(3) Living room—basement

(4) Living room (filtered)—basement

The "living room (filtered)" experimental situation involved inserting an electrical filter network into the signal directed to the living room loudspeaker. This filter altered the acoustical spectrum of the stimuli originating from the living room loudspeaker to very closely resemble the architectural acoustical filtering which naturally occurs as the sound is transmitted from basement to living room.

The type of sound, its level, and the tape track on which it occurred were all selected randomly, within the range of each parameter discussed above, by drawing lots. Only two limitations were imposed on the randomness: (1) The first stimulus to occur at the beginning of the tape had to be at one of the more intense levels to ensure that the respondents started by certainly hearing the stimulus. (2) No stimulus was presented more than twice.

In addition to the 90 absolute stimuli, 60 paired stimuli of the ABA type were included on the master tape program. These ABA stimuli were included to provide insight into a respondent's ability to make relative judgments compared to absolute judgments. In addition to the eight stimuli discussed above, three pure tones at 125 cps, 256 cps, and 4000 cps were included as stimuli. These pure tones introduced audiometric possibilities into the experiment and related directly to the pure tone equal-loudness contours as determined by Fletcher and Munson or Robinson and Dadson. (See References 6 and 7.) The equal-loudness contours are the most completely explored and defined psychoacoustic function existent.

The original intention was to adjust the randomly paired ABA stimuli to a precise loudness match and then find out if they would be judged as matched or unmatched with respect to bother. During master tape preparation, loudness matching caused too many difficulties and so ultimately the paired stimuli were adjusted to equal broadband rms electrical levels. As it turned out, the uneven sound distributions in the living room would have undone loudness matching anyhow. This same uneven sound distribution worked to our advantage by presenting stimulus A at different levels with respect to stimulus B as a function of speaker location and respondent's seat location. Thus some of the ABA combinations ranged from A considerably louder than B to the reverse. This range of values encompassed, in some cases, those which were judged matched or almost matched with respect to bother and this matched condition could be directly compared during analysis with the predicted loudness match condition.

Since it was desired that the ABA sounds always be clearly heard, their electrical levels were limited to the four highest recorded levels, namely, 0 db, -3 db, -6 db, and -9 db. Also each ABA combination was replicated at the same level but in reverse order of presentation, that is BAB, somewhere on the tape. Actually then there were only 30 distinctly different paired stimuli.

Subjectively, the relative judgment of bother engendered by the ABA stimuli seemed easier and more certain to make than the absolute judgments. However, it remains for the consistency analysis to reveal any inherent differences in performance of the two distinctly different judgment tasks.

3.5 ARCHITECTURAL ASPECTS OF THE RESEARCH HOUSE

To evaluate appliance sounds in the context of a home environment, a furnished dwelling was rented in Ann Arbor, hereafter referred to as "Research House." The Research House enabled us to deal more effectively with parameters such as building characteristics, background noise, transmission loss, and the distribution of sound in the rooms of interest.

3.5.1 Research House:

The building characteristics of the Research House were as follows: Total property value including lot, $17,500; area, 1,030 square feet; number of rooms, 5; number of levels, 1; full basement; roof construction, asphalt shingle. Table 1 compares the characteristics of the Research House with the respondent's own homes and with the Ann Arbor and National house characteristics. The housing characteristics of Table 8 were obtained as follows: The respondent's home characteristics by a field survey; the Ann Arbor housing characteristics from the U. S. Census of Population (Reference 8); and the National house characteristics from the Department of Commerce, Bureau of Census (Reference 9).

In the field survey concerning the respondent's home characteristics, the effective N[1] included both quality and non-quality respondents with non-quality respondents chiefly residing in multi-dwelling units. This effective N was slightly reduced in the field survey due to the inability to properly identify certain housing characteristics. This reduction occurred mainly in the quality group.

In Table 8 under National Housing Characteristics, the average existing house characteristics were used instead of the new house characteristics. The variance between new house and the existing house characteristics are as follows: The average new home of 1960 consisted of 5-1/2 rooms, of which 3 were bedrooms, with an area of 1,142 square feet. The average existing home (Table 8) was slightly smaller and had a total average area of 1,101 square feet. Also we find that 48 per cent of the existing homes had basements in contrast with only one third of the new homes. The total average for both existing and new homes shows a total property value of $14,102 with an area of 1,122 square feet.

The Research House as shown by Table 8 is comparable to the typical Ann

[1] N represents the number of respondents included in the sample.

Arbor home and the typical National home. In the areas of cost, basement versus no basement, and number of levels, the housing characteristics show a variance. Table 8 shows that the respondents and the Ann Arbor residents have a greater percentage of homes in the $20,000 and plus bracket (53.4%, Research House Interview Group; 39.0% Ann Arbor) while National house characteristics indicate in contrast only 5% of homes in this value range. We also find that Ann Arbor and the Research House Interview Group homes have more full or partial basements than the national average home.

Figures 2A and 3 present the floor plans of the research house and the furniture arrangement employed during the interview sessions. Figure 2B indicates the seat locations used by respondents and the interviewer. Figure 4 identifies the locations at which measurements were taken with sound level meters to determine the sound transmission and sound distribution characteristics of the research house.

3.5.2 Sound Transmission

The average sound transmission characteristics as a function of frequency from basement to living room and from kitchen to living room were needed to define the architectural transfer functions associated with the experiment. Moreover, the basement to living room data were required to guide the design of an electrical filter network to simulate this acoustical loss for the living room-filtered mode of operation. The transfer function from kitchen to living room was sufficiently flat with respect to frequency so that no electrical filtering was necessary to simulate the spectra of kitchen-located appliances heard over the living room loudspeaker.

The transmission measurements were performed with the aid of a special tape recording of warble-tone signals. Sound pressure levels were measured at the various locations indicated on Figure 4 and, for a given room, the decibel readings at the several locations were averaged arithmetically. Figure 5 presents the average sound reductions or the architectural transfer functions for the two paths of interest. The basement-to-living room function clearly resembles mass-law behavior with a 6 db octave slope over most of the range. There is some indication of room mode effects at the lowest frequencies. The electrical filtering selected to simulate this architectural function has simply a 6 db per octave slope.

The architectural transfer function from kitchen-to-living room, as shown in Figure 5, is approximately flat with frequency on the average although appreciable fluctuation does occur. It was decided that for this research house experiment, no electrical filtering as a function of frequency to simulate this transfer function was needed. As indicated on Figure 5, this kitchen-to-living room curve applies to a condition when some of the furniture had been removed from the living room. The slope and level of the curve had been verified with all furniture present but when preparing this report figure, through some in-

TABLE 8

COMPARISON OF THE HOMES OF RESPONDENTS ATTENDING THE WHIRLPOOL RESEARCH HOUSE WITH A CENSUS OF REGIONAL AND NATIONAL HOMES

House Characteristics:	Research House	Survey Group	Ann Arbor Homes	National Homes
Property Value				
Under $10,000		15.5%	4.0%	17.0%
$10,000 - 14,999		19.3	24.0	53.0
$15,000 - 19,999	$ 17,500	11.8	33.0	25.0
$20,000 - 24,999		21.4	17.0	4.0
$25,000 and over		32.0	22.0	1.0
Total		100.0%	100.0%	100.0%
Median		$20,800	$18,300	$13,200
Mean Floor Area	1,030 sq. ft.	n.a.*	n.a.*	1,101 sq. ft.
Number of Rooms				
Mean Number	5	n.a.*	4.7	5.4
Number of Levels				
One level	100.0%	53.8%	90.0%**	87.0%
Two levels	0.0	40.6	2.0**	4.0
Three levels	0.0	0.6	2.0**	2.0
Split level	0.0	5.0	4.0**	6.0
Total	100.0%	100.0%	100.0%	100.0%
Basement				
Full or partial	100.0%	98.2%	70.0%**	48.0%
No basement (slab or crawl space)	0.0	1.8	30.0**	52.0
Total	100.0%	100.0%	100.0%	100.0%
Roof Construction				
Wood Shingle	0.0%	3.8%	1.0%**	n.a.*
Built-up	0.0	1.6	4.0**	n.a.*
Asphalt Shingle	100.0	94.6	88.0**	n.a.*
Other	0.0	0.0	7.0**	n.a.*
Total	100.0%	100.0%	100.0%	

TABLE 8 (Concluded)

House Characteristics:	Research House	Survey Group	Ann Arbor Homes	National Homes
Basic Construction Materials				
Solid Brick	0.0%	0.5%	3.0%	2.0%
Brick Faced Masonry	0.0	14.6	11.0	10.0
Brick Faced Frame	0.0	25.4	41.0	33.0
Wood Faced Frame	0.0	32.4	30.0	24.0
Shingle Faced Frame	100.0	24.9	9.0	9.0
Stucco Faced Frame	0.0	2.2	***	14.0
Other	0.0	0.0	6.0	8.0
Total	100.0%	100.0%	100.0%	100.0%

*Not ascertained.

**North Central Section of United States percentages were used when Ann Arbor housing characteristics were not available (see Ref. 10).

***Less than 1/2 of 1%.

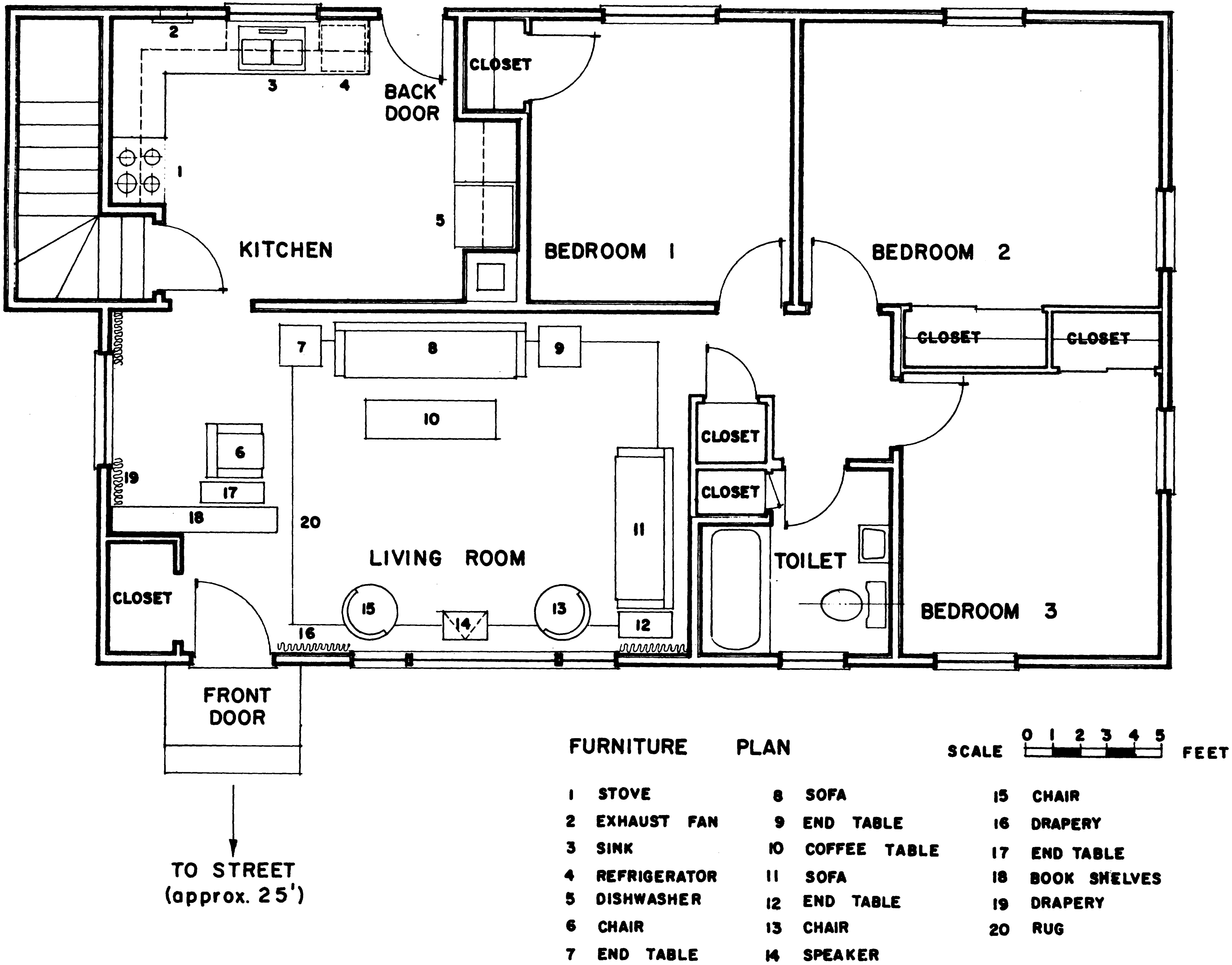


Figure 2A. First floor plan of research house.

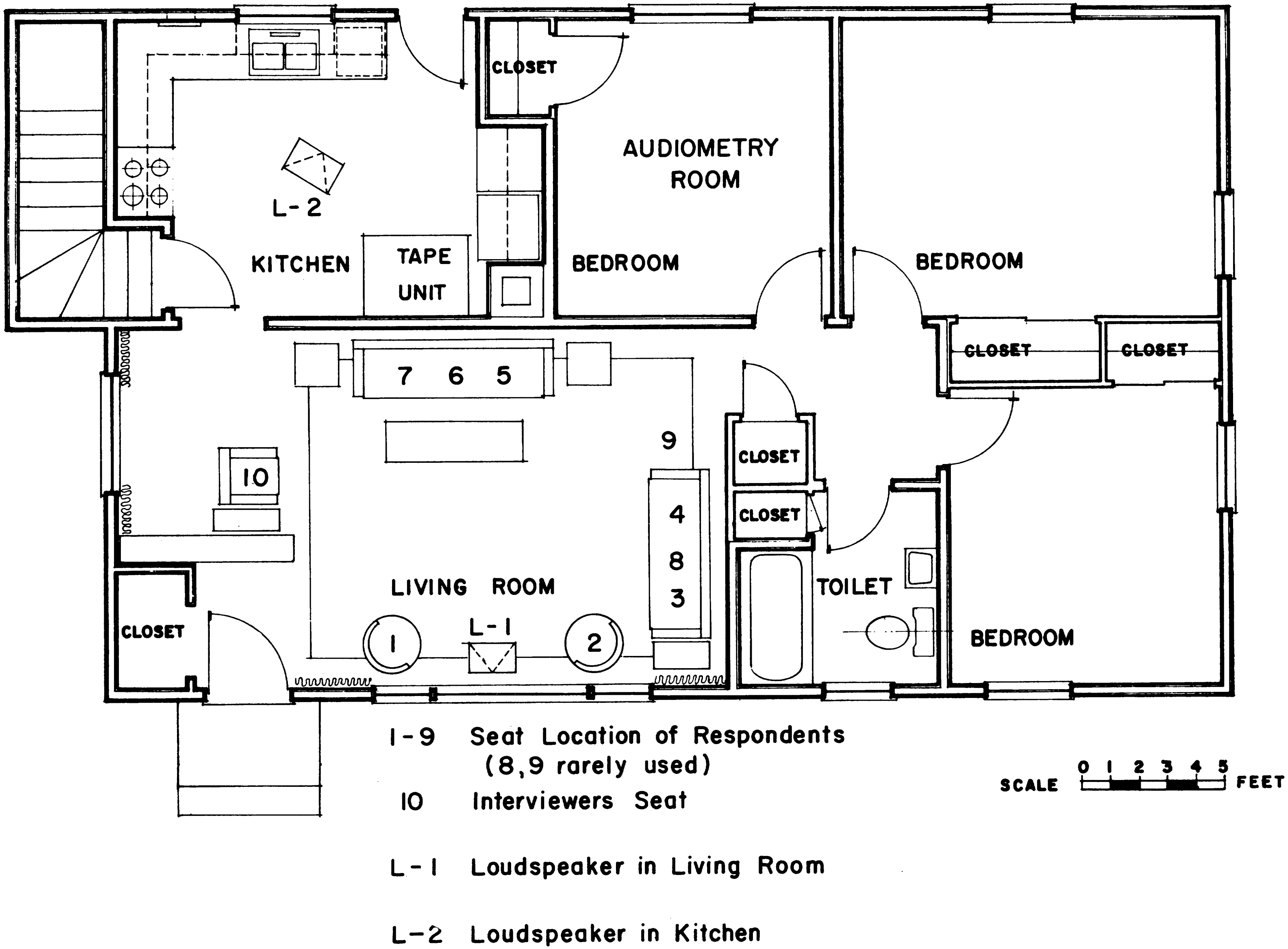


Figure 2B. Respondents' seat locations.

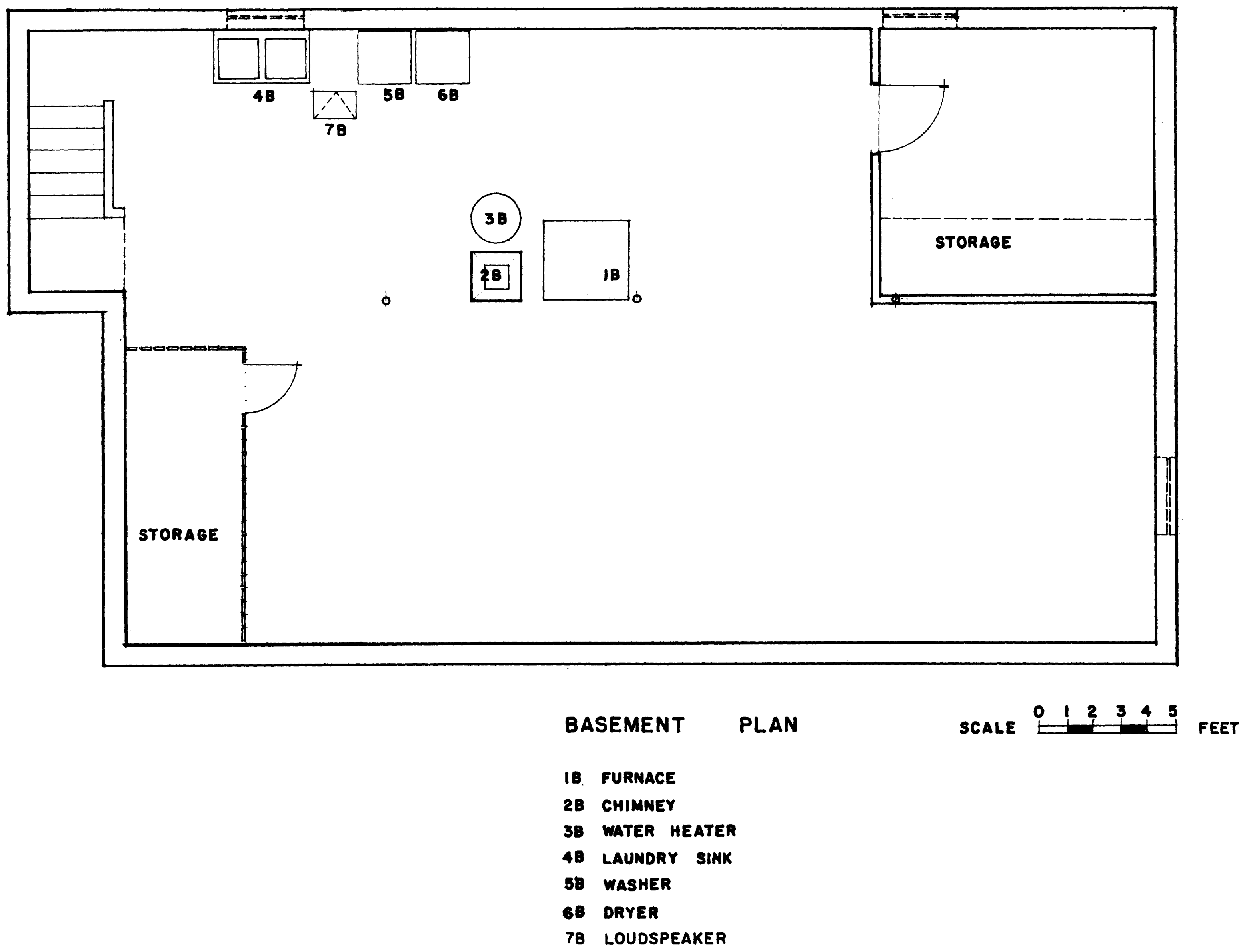
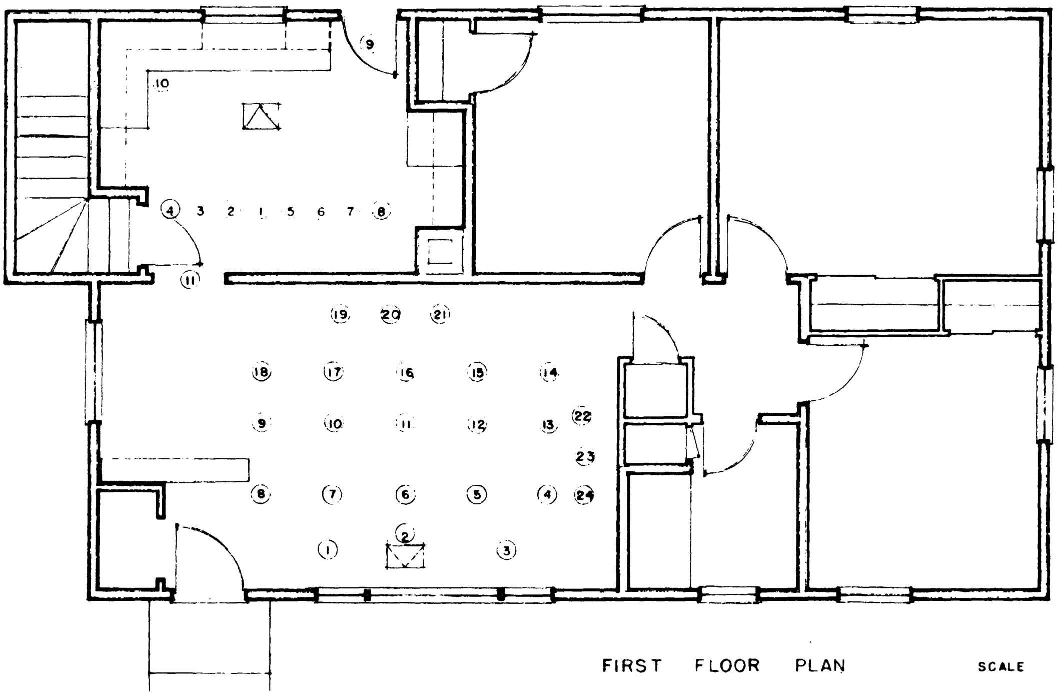
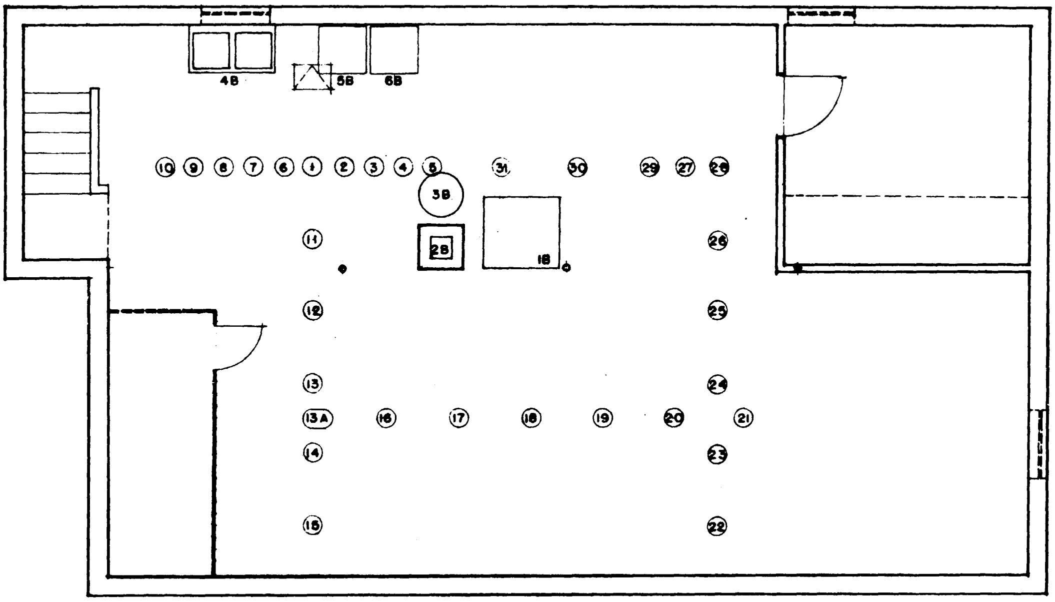


Figure 3. Basement plan of research house.



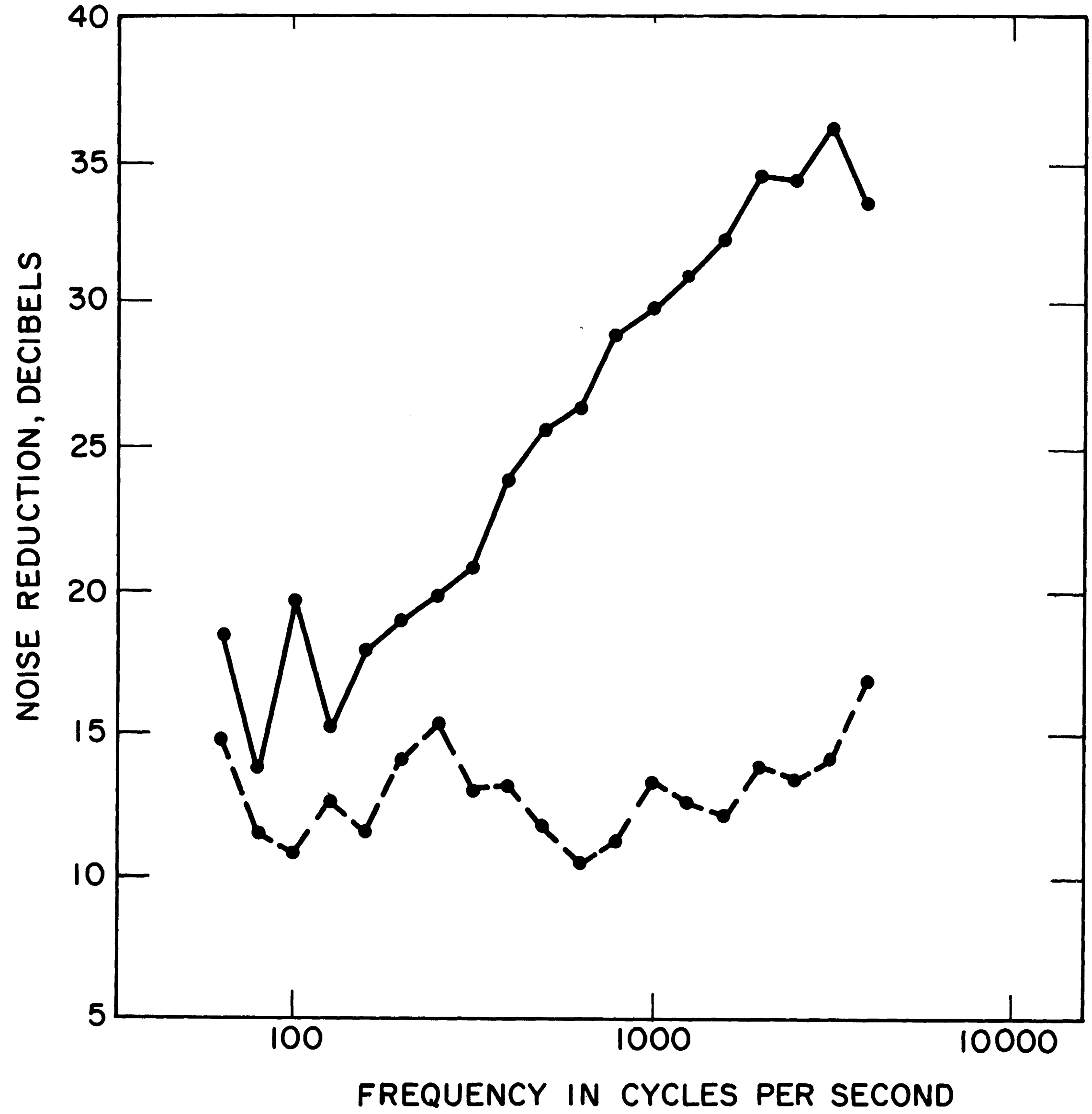
FIRST FLOOR PLAN

SCALE 0 1 2 3 4 5



BASEMENT PLAN

SCALE 0 1 2 3 4 5



Basement to living room transfer function (noise reduction). All furniture present.

Kitchen to living room transfer function (noise reduction). Living room partially furnished. Approximately 1.5 dB more reduction with full furnishing.

Figure 5. Architectural acoustical transfer functions.

advertences, the most complete data were for the condition of partial furnishing. The principal effect of removing furniture is to decrease the room absorption; with complete furnishing about 1.5 db more reduction was obtained across the spectrum.

3.5.3 Sound Distribution

As might be expected, especially regarding the predominately low-frequency stimuli, the distribution of sound intensity was not uniform around the living room. Furthermore, the distribution varied from one stimulus type to another and with loudspeaker location. As a consequence it became necessary to measure the actual sound pressure levels and frequency distributions which did occur for each stimulus for all modes of operation at ear level for all seat locations occupied by respondents. The measurement locations corresponding to such seats are represented in Figure 4 living room floor plan by numbers 1, 3, 19, 20, 21, 22, 23, and 24.

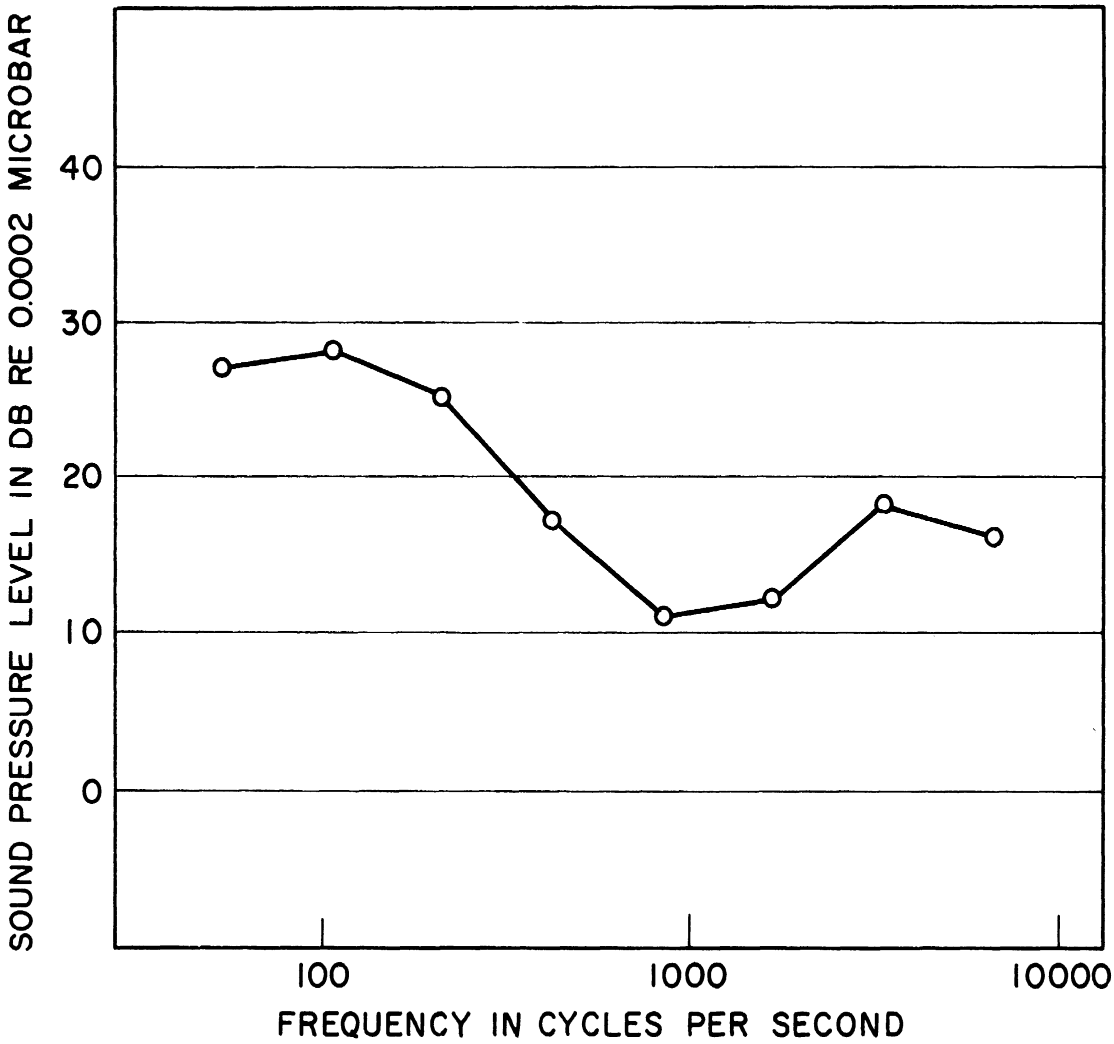
These physical sound data were used to compute the intensity level and loudness level associated with each stimulus presentation at each seat location and consequently the appropriate values of level were associated with each subjective judgment for the purposes of the various data reductions and manipulations. The final values of intensity level and loudness level will not be tabulated in this report because all of the answers to the questionnaires would be needed also to render such tables useful. Actually, all such information is contained on the punched cards and computer printouts which have been communicated to Whirlpool outside the context of this report.

3.5.4 Background Noise

The selection of a research house in a quiet neighborhood and the subsequent decision to conduct the experiment solely with natural background noise have been discussed in Section 3.3 (A convenient methodology does not exist for recording and describing background noise in valid statistical terminology. Background noise constitutes a significant problem area which needs to be thoroughly investigated in the future.) The background noise originating outside the research house was rather low in level most of the time and, by contrast, the occasional noises at higher levels were quite obtrusive. Noise from commercial aircraft constituted the most frequent source of interference but there were also irregular occurrences of noise from light aircraft, motor vehicles, power lawn mowers, children playing, etc. In most instances when such increases in background noise level occurred, the interview sequence was continued even though a few of the low level stimuli became completely masked. Occasionally, high levels of background noise persisted so that an interview had to be delayed as, for example, when a neighbor mowed an adjoining lot with a particularly noisy power mower.

The interview sessions were scheduled to avoid predictable periods of relatively intense local background noise such as generated by noon and 5:00 pm traffic and children returning home from school. Moreover, such periods were not convenient times for respondents either. The background noise generated within the house was of very low level also. The experiment operated during warm weather so that the heating plant was off and the only activity within the house was that associated with conducting the interviews.

Figure 6 presents an octave-band spectrum of the background noise existing within the research house at 4:00 am. By ear, this noise seemed to consist of distant traffic noise and insect noise. At the times of the interview sessions, the minimum background noise levels tended to lie somewhat higher than indicated in Figure 6. Ordinarily broad band measurements with a sound level meter (C-network) fell in the low 40's (db) at times when no individual noise sources such as airplanes or nearby cars or trucks could be detected by ear. Because of the warm weather, often one or more windows somewhere in the house remained open during the interviews.



Average octave-band spectrum taken at 4:00 am.

Figure 6. Background noise spectrum in research house.

4. ANALYSIS AND DISCUSSION OF EXPERIMENTAL RESULTS

4.1. BOTHER AS FUNCTION OF INTENSITY AND STIMULUS CHARACTERISTICS

Two concepts are central to the present research; the degree of bother engendered by an appliance sound is a function of the intensity of that sound, and the degree of bother at a given intensity varies with the type of appliance producing the sound. Such concepts are often said to be self-evident, meaning perhaps, that almost everyone would arrive at them after some reflection on the appliance noise problem. Moreover, they seem to be reasonably consistent with our remembered experiences concerning appliance noises. Nevertheless, if we require direct experimental proof confirmatory documentation of these concepts is virtually non-existent.

As a consequence, the research house interview experiment was designed so that, if true, the two concepts would devolve as logical conclusions. If they should turn out to be not true, or only true in some very restricted sense, then the diffuseness of the reduced data should so indicate. Moreover, in addition to a demonstration of validity, we need some quantitative idea of the corresponding functional dependences, of the minimum perceptible differences, of the distributions of individual responses, etc.

Consider what may reasonably be expected to occur when we conduct an experiment on the degree of bother as a function of stimulus intensity. The respondents are presented a noise stimulus which reoccurs at various intensities and they are asked to make an independent judgment of the bothersomeness of each successive presentation. (The mechanics for maintaining a consistent frame of reference and independence of judgment are set aside for the present discussion.) If the precise question to be answered for each trial was, "Was that sound bothersome," requiring an unqualified yes or no answer, what might we realistically expect to find when the reduced data have been collated?

First, imagining the probable responses of an individual, we would anticipate that at very low intensity levels the sounds would be judged not bothersome. Possibly, a stimulus might have to be completely inaudible to elicit this response but certainly a not bothersome response could be achieved if the intensity level is carried to low enough values. In the opposite direction, it should be possible to raise the intensity level until even the most resistive respondent would agree that a stimulus was bothersome. Somewhere between these extremes of stimulus intensities must lie a range of intensities where a respondent switches his judgment from not bothersome to bothersome. It is largely immaterial for the purposes of the present discussion whether this change in judgment occurs gradually or as a threshold.

To this assumed response pattern of the individual respondent, add the assumption that the individual responses contributed by a group of respondents will tend toward consensus without expecting perfect agreement. Under these conditions and for a particular stimulus type, we should be able to construct a graph of the mean results resembling Figure 7.

The entire possible presentation space is divided into two regions representing the cumulative percentage of bothered and not bothered respectively at each intensity level. The boundary between the two regions may be clearly defined or vaguely defined depending upon the detailed characteristics of the individual responses and the statistics of the total ensemble of responses. In any event, this boundary should tend toward a single-valued S-shaped curve if bother is a meaningful concept to the respondents.

If the bothersomeness of different types of stimuli, i.e., different appliance—like sounds, is distinctly different then we may confidently expect to observe those differences as corresponding changes among the S-shaped boundary curves. Thus a boundary curve might shift right or left, become more or less precisely defined, or perhaps the slope of the central portion might change. Indeed, there seems to be no a priori reason to assign any particular shape to the central portion of the postulated S-shape curve as presented in Figure 7. Rather this curve must run from upper left to lower right and probably it tends toward a fairly smooth and continuous shape.

If instead of an elementary yes or no answer to questions of bothersomeness, the respondents are permitted to qualify a yes answer with an expression of the degree of bother, then the postulated graphical presentation in Figure 7 must be elaborated to accommodate the additional information. In our research house experiment using absolute stimuli, the respondents were asked to qualify a yes answer in terms of only a little bothersome, moderately bothersome, or extremely bothersome. This additional information concerning degrees of bother might be treated in various ways and we have elected to incorporate it as a parameter into the format of Figure 7 yielding Figure 8. The form of Figure 8 seems to be most appropriate to this appliance noise problem and it precisely represents the information we had proposed to provide to Whirlpool.

There are now three S-shaped curves which subdivide the format of Figure 8 into four regions. As in Figure 7, the area below and to the left of the lowest curve represents not bothered and the entire region above this curve represents bothered. In Figure 8, however, the entire area below and to the left of the middle curve represents the cumulation of not bothered and a little bothered and so it represents at most a little bothered. In a like manner, the entire area below and to the left of the highest curve represents at most moderately bothered.

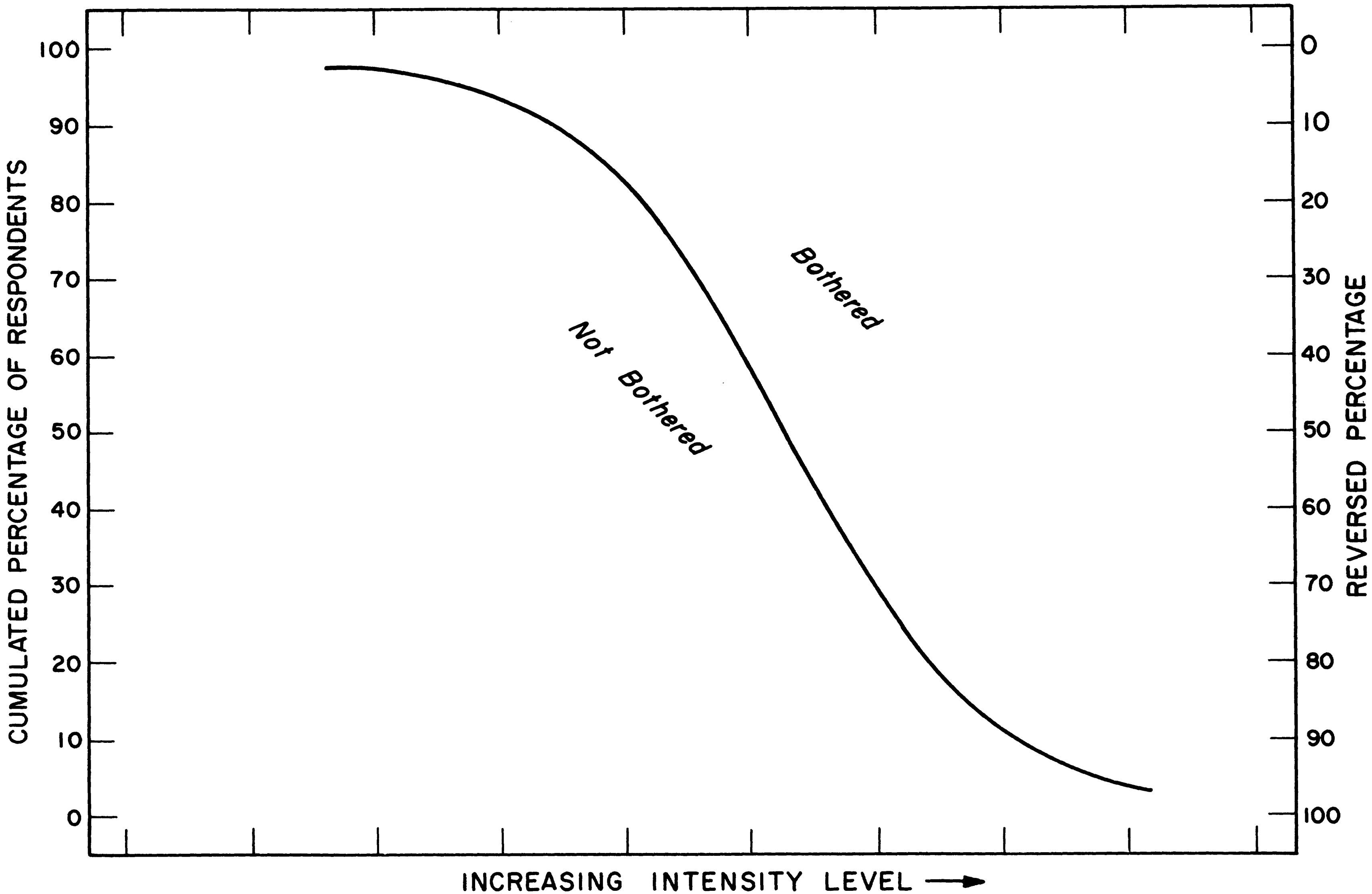


Figure 7. Postulated bother function.

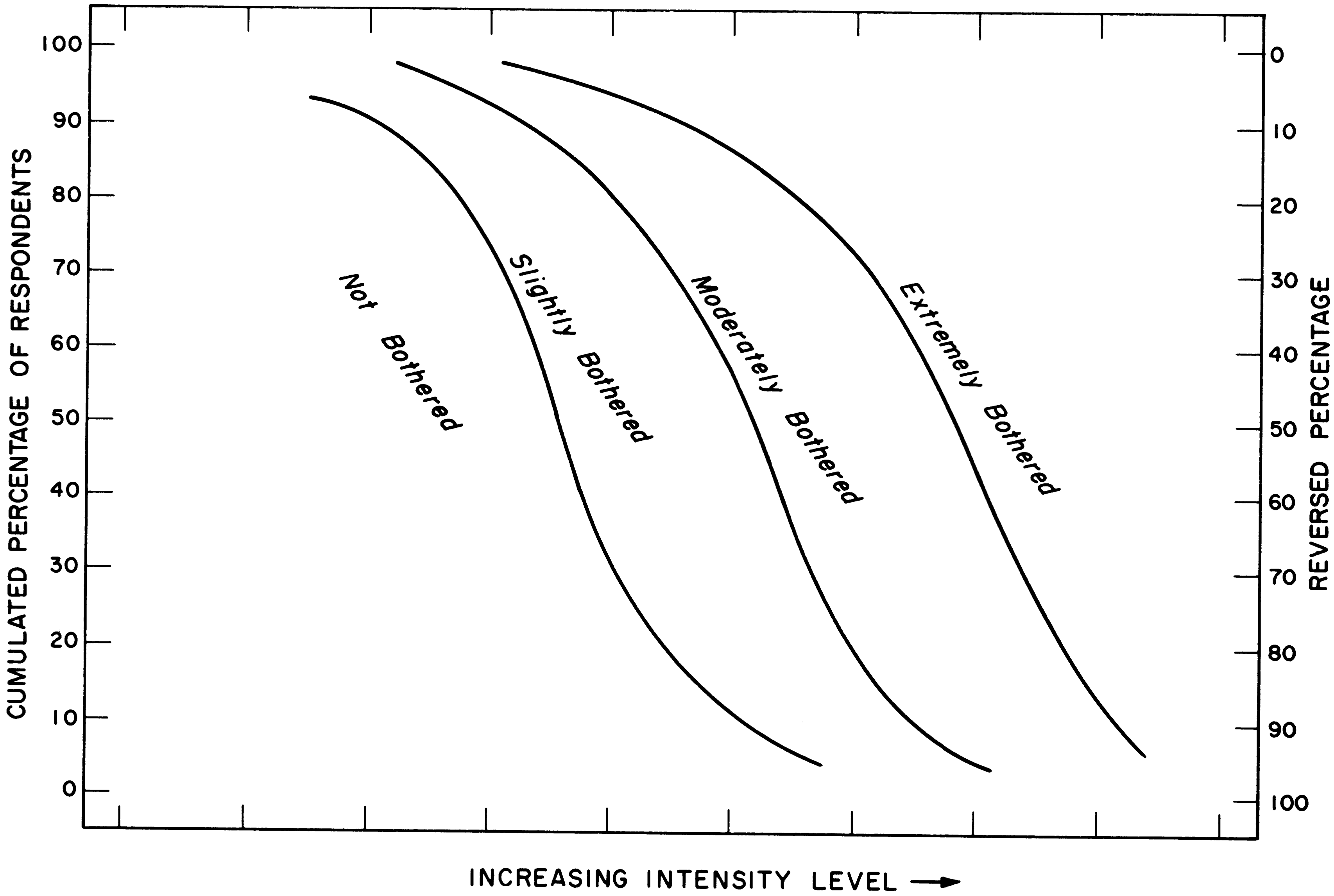


Figure 8. Postulated bother function with degree of bother as parameter.

The format used in Figure 8 was selected as the most convenient and practical display of this type of psychophysical data. It assists in working out answers to many types of applied questions as well as theoretical questions about responses. However, an exhaustive treatise on how to employ such graphs is clearly beyond the scope appropriate to this report.

Before presenting detailed experimental results in the format of Figure 8, there are several matters relating to the nature of the experiments performed and the generated data which need to be discussed. First, in order to contain the scope of the research house interview experiment, a limited range of stimulus intensity level was selected on the basis of preliminary tests; also, a limited number of intermediate levels. Consequently, the eight absolute stimuli were presented at seven levels spaced nominally 3 db apart. The ends of the S-shaped curves are less well defined than their central portions simply because no weaker or stronger stimuli were available to generate the required data.

A second pertinent feature of the reduced data arises from the natural spatial distribution of the sound signals around the living room of the rented house. The actual sound pressure at the respondents' ears differed depending upon where the respondent sat and these differences varied from one stimulus type to the next as well as with loudspeaker location. (In an earphone experiment, identical stimulus levels could be presented to all respondents by observing suitable precautions.) The observed differences from one seat location to another were not just small shifts in the nominal levels but often were larger than the nominal 3 db intervals between stimulus levels. As a consequence, the levels of all stimuli at all seat locations were measured and the raw data for each respondent adjusted accordingly before undertaking the various data reductions.

The non-uniform sound distribution resulted in wide fluctuations in the number of responses obtained at various intensity levels and introduced a need to average the data with respect to intensity level. Several averaging and smoothing techniques were tried and finally a type of sliding average was chosen as best suited to these data in contrast to some of the more rigorously defensible statistical procedures. Each plotted value is an average of the data at that intensity level and at the next two lower intensity levels. This procedure appears to have smoothed the curves sufficiently to reveal general trends in the data. Even so, considerable jaggedness remains in the middle portions of the curves and, of course, the ends of the curves have not been smoothed.

Some of the residual jaggedness is the natural consequence of a small sample size as can be verified by examining the computer print-out of the reduced and averaged data. Other irregularities can be traced to specific features of the experiment which influenced the respondents' judgments. For example, during the early part of an experimental session, a respondent will be building up a mental concept of a bothersomeness scale and many respondents appear to have reserved extremely bothered in anticipation of worse stimuli. Loudspeaker location also acts as a confounding variable to a certain degree in as much as the sample size is smaller for each loudspeaker location compared to the merging of data from all loudspeaker locations. In addition, the range of stimulus intensity varied with loudspeaker location due to the acoustic transmission characteristics of the research house.

An example of reduced experimental data presented in the format of Figure 8 is given in Figure 9. It represents the mean bothersomeness function, as obtained in our experiment, for the noise of a clothes dryer when the results for all loudspeaker locations have been merged. This graphical presentation is subject to all of the considerations discussed above. The curves are distinctly S-shaped and correctly ordered and yet they evidence considerable fluctuation, especially at the highest intensity levels. We have every confidence that a refined experiment of this same type with enlarged sample and extended intensity range could yield even more perfect function curves subject only to the variance appropriate to human beings operating within our experimental framework.

The full collection of these curves for all eight stimulus types both for individual loudspeaker locations and merged with respect to location are given in Appendix B. They represent the psychoacoustical functions relating to appliance noise which we set out to discover using the research house experiment. With the assistance of the explanations already given above, the collection of graphs in Appendix B stand as a completed statement of fact about the bothersomeness of appliance noise and closely related noise.

As mentioned above, most of the residual functions in these curves can be traced back to some particular aspect of the experiments. When such an identification has been accomplished, it would then be possible to either eliminate the offending position of the data or weight it appropriately and then recompute a more refined curve. However, the labor of identifying and testing all such fluctuations was prohibitive at this stage of the research, especially since the curves reproduced in Appendix B already involve three successive stages of computer-reduction of the raw data. The residual fluctuations, nevertheless, bear strong implications for future experiments of this same type. In the present report, they have already fulfilled a valuable role in alerting us to the nature and magnitude of some of the secondary effects which occur in experiments of this type.

The principal value of the graphs relating cumulative bother to stimulus intensity is that they are directly interpretable in terms appropriate to application engineering. Certainly the degree of bother is a function of stimulus intensity. Thus the graphs verify one of the central concepts discussed earlier and provide a quantitative form of the functional relationship. The second concept, that for a given intensity, the degree of bother is a function of the type of sound, is also positively affirmed by the graphs

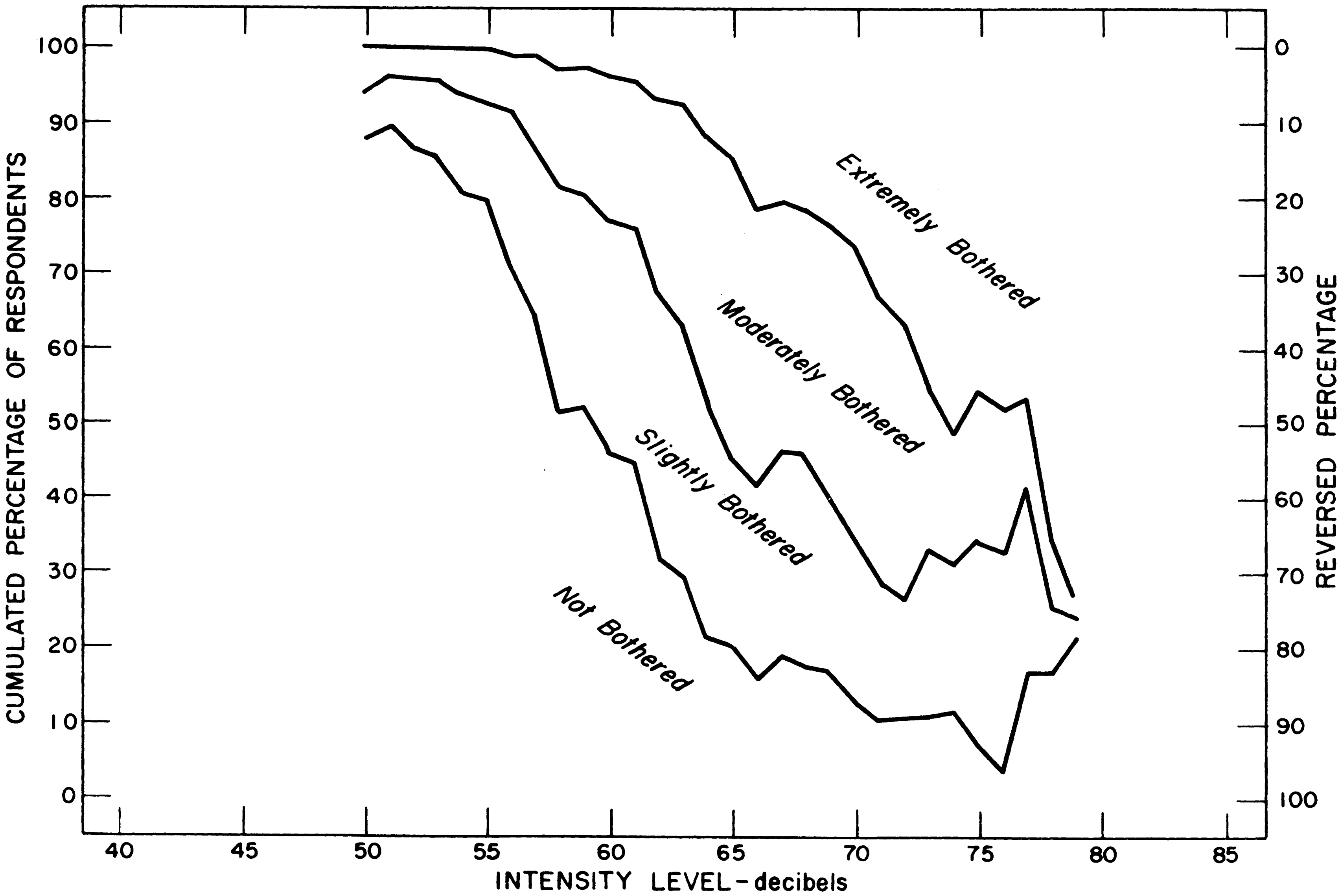


Figure 9. Example of bother function plotted from experiment data.

in Appendix B (see for example Figures 13 and 43, Dryer and White Noise respectively). The curves for White Noise are smoother, steeper, more compactly spaced and lie further to the left than the corresponding curves for Dryer. Clearly, Dryer type noise can be somewhat more intense, physically, than White Noise before it produces the same degree of bother response. This difference just cited is one of the more extreme which occurs among the graphs of Appendix B. A closer study of these graphs results in some confusion about the reality of subtler differences and the whole ensemble of 40 graphs is difficult to visualize for cross-comparison purposes.

These graphs are least effective in answering more basic questions which the authors, in their role of researchers rather than practical appliance engineers, must ask and answer. For example, "Are the eight physically different sound stimuli actually distinctly different stimuli with respect to the subjective judgment of bothersomeness?" In order to deal with questions of this type a correlational technique was employed (see Appendix C). Applying this statistical technique, the data, merged with respect to loudspeaker location, relating bothersomeness to intensity level can be represented by straight regression lines, as is done in Figure 10.

On this basis, the eight stimuli fall into two groups: WN and SBF forming one group and lying quite far to the left of the other six stimuli. Possibly an even finer subgrouping of the stimuli, perhaps with respect to slope, could be made but first the question of what constitutes a significant difference in Figure 10 must be answered. That question will be answered rigorously below but for the moment, merely assume that the spacing and ordering of the regression lines in Figure 10 does possess valid meaning. Several observations and conclusions can be made.

1. The purely laboratory sounds, not intended to resemble appliance noise are the most different from appliance-like sounds on the basis of bothersomeness.

2. Loudness calculations work best on broadly distributed sounds, similar to white noise, so we might anticipate that loudness of a sound may not be a particularly good predictor of bothersomeness for appliance sounds.

3. Much laboratory research in the area of psychoacoustics is conducted using white noise or bands of noise filtered from white noise perhaps because the steep slope makes for clear-cut results. Extrapolations to appliance noise problems from such research data may not be very successful.

4. Our laboratory sounds of O and O + PT, intended to partially simulate real appliance sounds, appear to have accomplished their purpose.

The question of what variations among the regression lines of Figure 10 constitute significant differences can be answered by considering the corresponding fiducial limits. Figure 11 presents the same regression lines for WN and

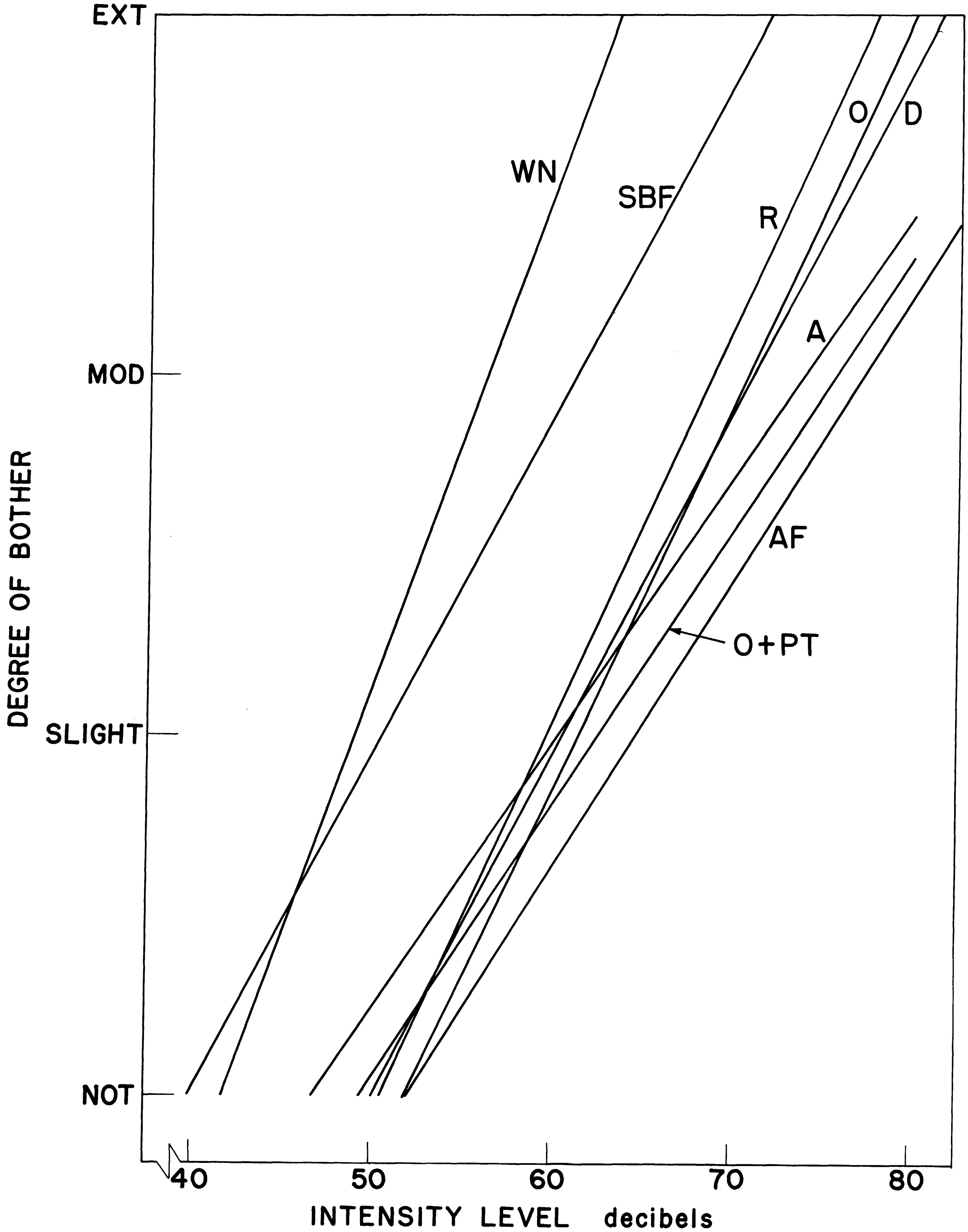
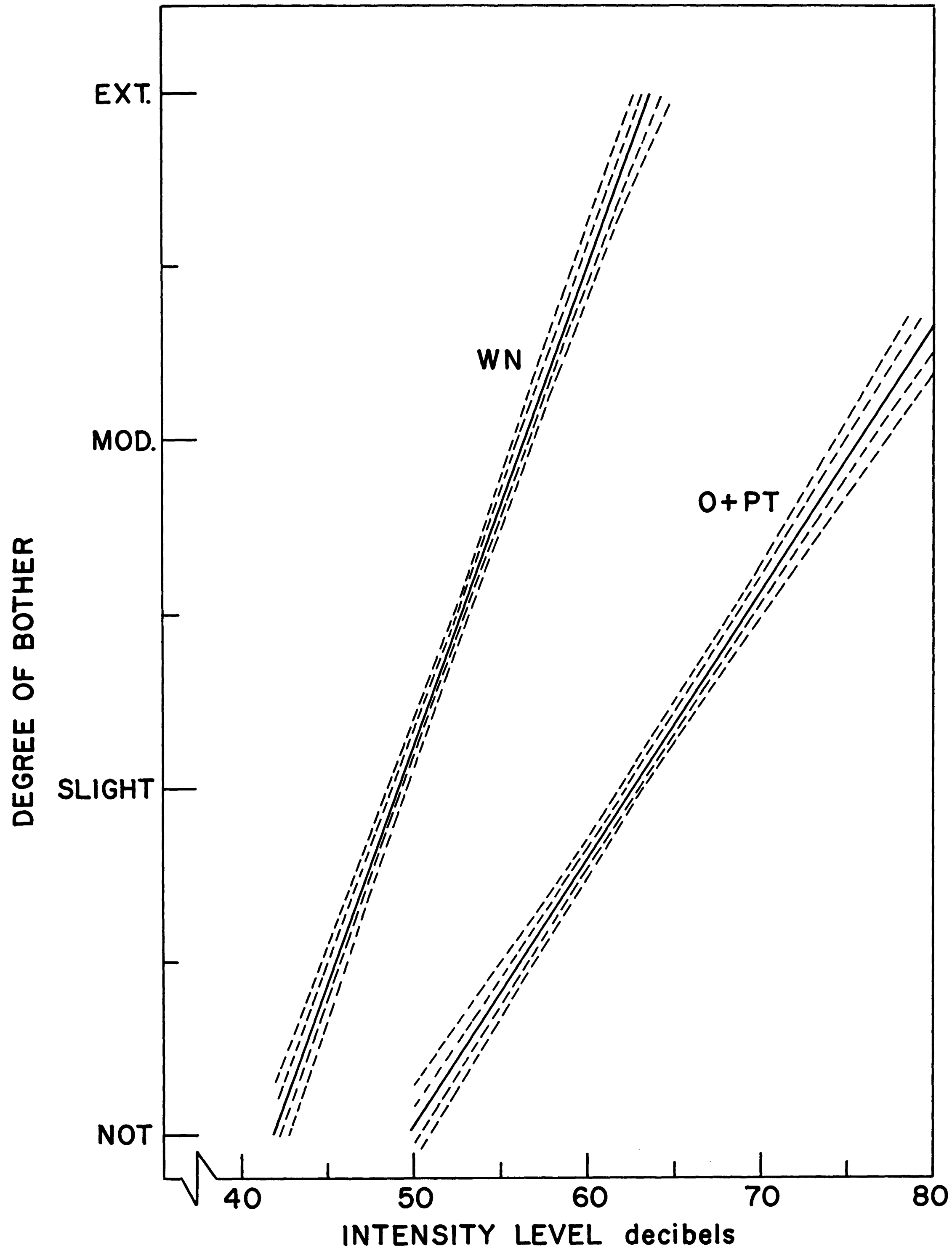


Figure 10. Regression lines for various stimuli.



Solid line - regression line

Inner pair of dashed lines - 10% confidence limits

Outer pair of dashed lines - 1% confidence limits

Figure 11. Regression lines for WN and O+PT with fiducial limits.

0 + PT (the most different of the purely laboratory sounds) as shown in Figure 10. In addition, these lines are surrounded by the fiducial limits at two levels of confidence. The closest set of limit lines correspond to a 10% confidence interval and the outer set to a 1% confidence interval. The correct interpretation is that, in the case of the outer set, if repeat experiments were performed, then 99 times out of 100 the new straight regression line would fall somewhere within these bounds.

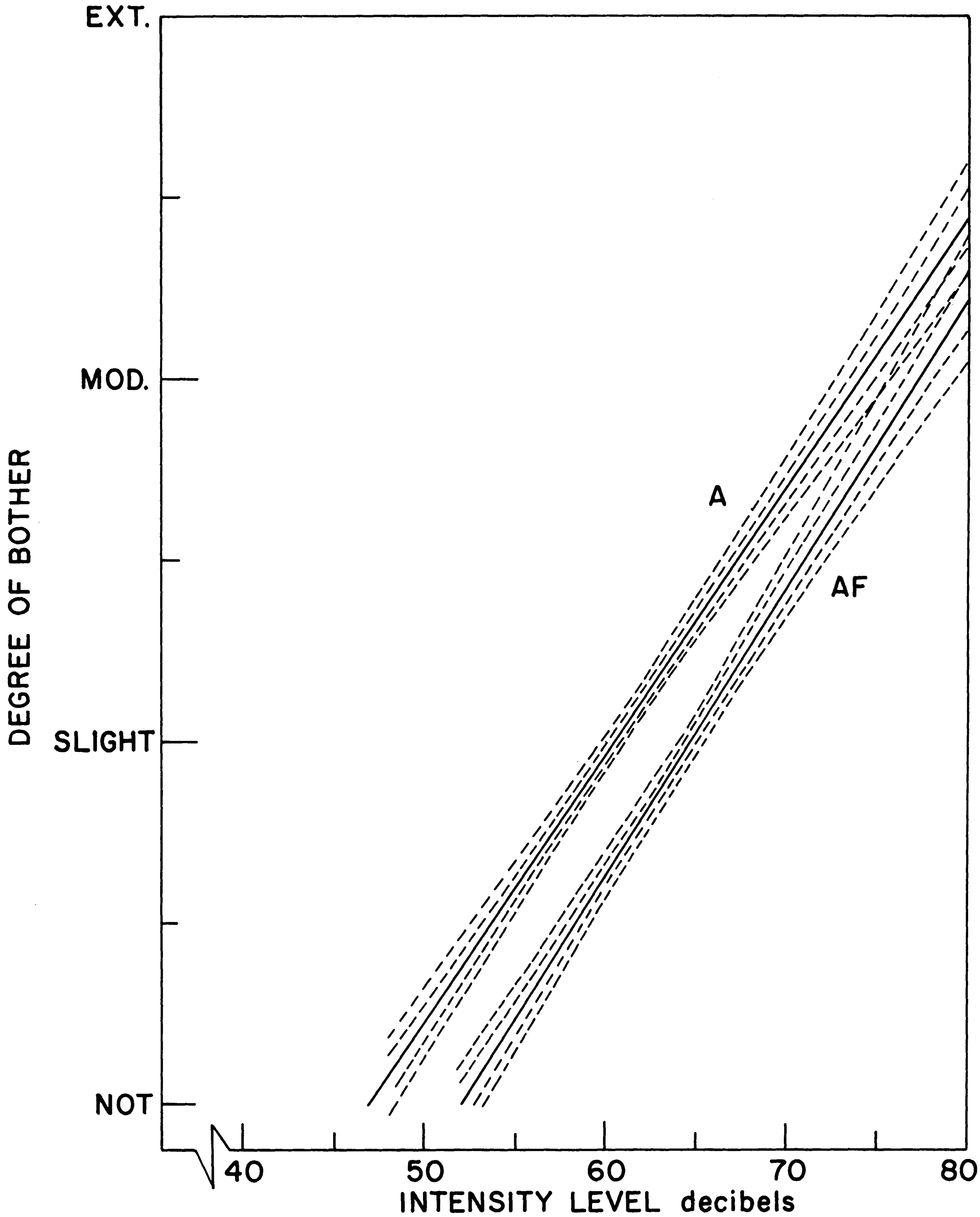
With the above interpretation, WN and 0 + PT given in Figure 11 represent distinctly different subjective functions over the range of intensity investigated. Indeed the difference is much greater and much more certain than one would estimate from the corresponding graphs in Appendix B.

Figure 12 displays the same sort of information with respect to A and AF. The test in this case is more revealing since the regression lines are parallel and quite closely spaced. In this case also, we conclude that subjectively A and AF are distinctly different, the fiducial confidence limits only merging slightly at the highest intensity levels. Indeed, the regression line for 0 + PT and its fiducial lines, as shown in Figure 11 neatly fit into the space between A and AF on Figure 12 with merging of the fiducial lines only at high and low intensity levels.

When the various regression lines and their fiducial confidence intervals are examined as a whole only D and O overlap sufficiently to question whether or not they should be considered separate acoustical entities in the context of our experiment. All other stimuli are distinctly different from one another. Viewed in this manner, the experiment has been extremely successful in separating sound types which, by design, were quite similar to one another. And this success was achieved with a small sample size and a large measure of experimental mud arising from the many confounding effects still incorporated in the experiment. The sounds tested here represent a very small sampling of a multidimension sound space but the clear results indicate that in future experiments one could attack the multidimensionality rather boldly with excellent chances of success.

The use of regression analysis and fiducial confidence limits, as represented by Figures 10-12 start to provide an answer to the question of what type of appliance sound will be least bothersome. It must be remembered that the present experiment has only barely touched upon the question of a multidimensional sound space (the usual frequency spectrum i.e., frequency vs. intensity representing only two of many possible dimensions). Nevertheless, in the context of the regression lines, ideal sounds from appliances should be represented by lines with shallow slope occurring far to the right (without political implications).

In the present experiment, it turns out that we did not explore very far in that direction. Indeed our laboratory sounds, WN and SBF, are very clearly



Solid line - regression line
Inner pair of dashed lines - 10% confidence limits
Outer pair of dashed lines - 1% confidence limits

Figure 12. Regression lines for A and AF with fiducial limits.

in the wrong direction. Of the other purely laboratory sounds, O was very like the realistic D sound while O + PT moved in the correct direction but not as far as a filtered realistic loud, namely AF. The trend would seem to be toward low frequency sounds principally of random characteristics but with some structuring or patterning discernable. However, more research is needed to find out how much improvement one can hope to achieve in this direction.

Summarizing, Appendix B contains the detailed functional data relating appliance noise intensity level to subjective bothersomeness. These are the relationships which were the objectives of this research program. Although the data often appear weak at first glance, application of regression analysis techniques show the results to be very clear. The various stimuli employed likewise have been shown to be subjectively distinct from one another. The results clearly demonstrate the validity of the two concepts discussed at the beginning of this section and go beyond the verbal concept to provide functional and quantitative detail.

After data reduction, the clarity and power of the results indicate that one may go back into the data and ask for at least probable trends with respect to more subtle psychophysical variables. Likewise with this experiment as a background, we are now in a position to undertake some critical check experiments in the laboratory and also to undertake more sophisticated field experiments.

4.2 BOTHERSOMENESS VS. LOUDNESS AS RELEVANT FUNCTION

In this research, the word bothersomeness has been used to connote that subjective psychophysical aspect of appliance sounds which we wished to investigate. This word was carefully chosen to be meaningful to the respondents and as neutral in implications as possible.

In the area of psychoacoustics, loudness is a psychophysical function which has been investigated extensively and technically defined. In fact, loudness is essentially the only subjective magnitude function which has been reduced to engineering practice and which is generally accepted by researchers in this area. Various attempts have been made to work out other functions on a quantitative basis, such as annoyance, noisiness, etc., but as yet such functions have not enjoyed universal acceptance. The loudness function has been developed both for pure tones and for bands of noise. The computational procedures yield good results for wide-band noises possessing fairly regular characteristics. Likewise, high relative accuracy is obtained for any noises of similar characteristics. However, the loudness of noises having widely differing frequency characteristics and irregular frequency distributions can not be predicted as accurately.

Because the loudness function and the associated computational procedure already exist and because loudness has been employed in appliance noise studies, we wanted to test whether this loudness function was an adequate psychophysical function for our purposes. In other words, was there any ultimate need to introduce the concept of bothersomeness as distinct from loudness?

There are several ways to investigate this question. The loudness and loudness level for all stimulus presentations were calculated from the corresponding octave band spectral data using Stevens's method (see Reference 11). Thus each response could be compared with either the sound intensity level or the loudness level of its generating stimulus in the course of data analysis. In one investigation, we plotted a set of graphs similar to those in Appendix B except that the cumulative percentage of bothersomeness was plotted against loudness level instead of intensity level. (This set of 40 graphs in terms of loudness level has not been reproduced in this report.)

If loudness were an adequate function for our investigation, then the cumulative percentage curves as a function of loudness for the eight absolute stimuli should have coalesced. They did not! The differences observed in Figures 13 and 43 for example (Dryer and White Noise respectively) in the forms of horizontal displacement, slope, spacing, etc., were decreased when loudness level was used instead of intensity level, however, the improvement was only about one third of the amount needed for coalescence.

From such comparisons, it can be deduced that the loudness function possesses a certain measure of appropriateness, as would be expected, but it is equally clear that loudness is not an adequate function to account for our data. The calculation of loudness level principally involves a non-linear weighting with respect to the frequency distribution of the sound. Thus, conversion to loudness level did reduce the differences between WN and SBF, which contain appreciable high-frequency noise and the other predominently low-frequency stimuli.

A more succinct way to examine the adequacy of loudness as the applicable psychophysical function for this experiment is to examine the proportion of variance (eta squared in the computer print-out) accounted for when applying the method of multiple correlation (see Appendix C). Table 9 below shows this variance for the eight absolute stimuli individually and for the aggregate of all eight stimuli when using either intensity level or loudness level as the independent variable.

First, it can be seen that considering the stimulus types individually there is no value to replacing intensity level with loudness level. This result is to be expected on the basis that conversion to loudness level is predominantly a frequency weighting. Likewise, this result emphasizes that within a particular sound type, prediction of response is not improved by substituting the subjective scale, loudness level, for the physical scale, intensity level.

TABLE 9

PROPORTION OF VARIANCE ACCOUNTED FOR BY DIFFERENT MAGNITUDE FUNCTIONS

Stimulus:	A	R	D	WN	SBF	O	O + PT	AF	Merged
Magnitude Function									
Intensity	.296	.486	.379	.412	.404	.413	.264	.241	.188
Loudness	.266	.473	.391	.482	.394	.422	.264	.250	.356

Note: Numerical values are eta squared.

Secondly, looking at the variance for the merged stimuli, the change from intensity level to loudness level produced an improvement by a factor of two (from 0.19 to 0.36). This improvement, of course, is largely the consequence of an appropriate weighting with respect to frequency. However, variance in the amount of 0.64 (i.e , 1 - 0.36) remains to be explained.

A complete accounting for this remaining variance can not be given at present. The largest amount of variance accounted for, taking the stimuli individually is about 0.48; a value which at least suggests the possibility of accounting for somewhat more variance if the most appropriate aggregating function is used. Moreover, this value of 0.48 was achieved with many confounding influences, e.g., loudspeaker location, still present in the data.

A certain amount of the residual variance must be due simply to the normal variability among human subjects performing within the framework of this kind of experiment. It is difficult to estimate how much variance should be ascribed to this cause but it certainly does not seem reasonable, all factors considered, that it approaches 0.64. Some indications bearing on this problem of inherent human variance are to be found in Section 4.5. In any event, the residual variance seems to be a composite of the variances due to inherent human variability, loudspeaker location, stimulus type, and other identifiable physical and demographic variables, as well as a liberal portion of muddiness associated with the exploratory nature of this experiment.

Another check on the appropriateness of the <u>loudness</u> function to the present appliance noise study arises from a consideration of the ABA stimuli. As it turned out, some of the ABA combinations ranged from say, A louder than B through an approximate loudness balance to B louder than A. The stimulus levels which occasioned approximately equal bothersomeness responses did not correspond to levels for a computed loudness balance. (The data from the ABA stimuli have not been as extensively analyzed as for the absolute stimuli except for consistency and hence a detailed presentation based on ABA data will not be given).

Another observation relating to the appropriateness of psychophysical functions deserves mention. A subsample of our respondents were queried with respect to annoyance rather than bothersomeness. Tests of the resulting data indicated no statistically significant differences. These results do not merit the interpretation that annoyance and bothersomeness are the same function but rather that within the scope of this one limited test, they were similar enough to escape differentiation. The reasons for preferring bothersomeness when discussing the appliance noise problem still prevail.

The above results are consistant, broadly speaking, with some results of Rule and Little (Reference 12) obtained in a distantly related psychophysical experiment conducted at higher intensity levels. They found, within a frame of reference related to jet aircraft sounds, that loudness and noisiness were essentially identical functions while annoyance was distinctly different.

Assembling the several items of evidence available from the research house experiments, loudness is not an adequate function to account for the bothersomeness of the several stimuli. Loudness can account for a portion of the variance but it does not go far enough. Thus the frequency weighting attending the loudness calculation partially narrowed the gap between the appliance-like sounds and the unappliance-like sounds, namely WN and SBF. From this view point, any frequency weighting function operating in the proper direction to adjust from the realm of physical frequency distribution to a corresponding subjective realm must, of necessity, account for some of the variance. The emphasis of interpretation must fall, therefore, on what loudness failed to accomplish. The residual variance after converting to loudness is sizeable, consequently loudness intrinsically cannot be the function sought.

Regardless of the name applied to an appropriate psychophysical function, the graphs of Appendix B must represent it within the appliance noise context. A variety of psychophysical functions, e.g., noisiness, annoyance, etc., have been proposed by several investigators from time to time and have been at least partially defined. We have not specifically tested these to determine if any of them are more satisfactory than loudness for present purposes. There are several reasons for such an omission which briefly are as follows:

1. These other functions differ from loudness principally in the details of frequency weighting whereas our experiments suggest the strong influence of parameters other than frequency.

2. Many of these other functions were proposed for and tested upon sound stimuli having quite different characteristics, intensity levels and frames of reference from appliance noise.

3. Computational procedures often have not been fully enough developed to permit cogent testing of them in a somewhat different experimental framework.

Thus bothersomeness as used in the research house experiments remains not loudness and beyond this, is defined by the frame of reference developed during the research house experiment.

4.3 EFFECT OF APPLIANCE (LOUDSPEAKER) LOCATION

In the course of the Phase One experimental studies, we began to suspect that our respondents were taking some acoustical aspects of their homes into account instead of judging the stimuli solely on the characteristics of the sounds at their ears. They seemed to be imagining an appliance located elsewhere in the house and estimating what its noise must have been like to have produced the stimulus heard in the living room. Thus they seemed to utilize direct knowledge of sound location, or sound properties like directional characteristics or frequency weighting (corresponding to the sound transmission characteristics of the house), or perhaps some combination of these in arriving at their judgments. However, which possibility was the correct explanation and the strength of its influence remained uncertain.

The observations described above motivated the introduction of several loudspeaker locations and an electrical filtering into the variable parameters for the research house experiment. It should be possible, upon analysis, to differentiate between a location effect and a frequency filtering effect. The answers with respect to these two possibilities are contained (perhaps concealed is a more accurate statement) in the detailed graphs of Appendix B.

A comparison of Kitchen and Basement results (also, Living room unfiltered and Basement) would involve both location and filtering effects simultaneously. A comparison of Living room-filtered and Basement (also, Kitchen and Living room-unfiltered) presumably involves only location effects since the electrical filtering has removed the frequency differences. (See Section 3.5 for the experimental architectural transmission functions). Finally, a comparison of Living room-unfiltered with Living room-filtered should involve only a pure filtering effect since location is constant.

A close study of the set of graphs in Appendix B leads to the conclusion that the location and filtering parameters seem to exert some influence on bothersomeness but the pattern of that influence remains very hazy. Upon turning to the application of multiple regression analysis procedures, a somewhat clearer picture begins to emerge. When taking location of the source into account, correlation was improved for five of the stimuli and decreased for three, namely, AF, O + PT and WN. That is, the decrease of correlation occurred for the more unrealistic stimuli and was most extreme for WN. In the case of the frequency filtering parameter, generally a small decrease in correlation occurred. When both parameters acted simultaneously, the results for the eight absolute stimuli scattered with perhaps a slight average improvement. The three realistic stimuli, A, R and D, as well as O exhibited increased correlation.

The results cited above are still rather weak but they point to location of the source as a valid parameter. At this juncture, the frequency filtering parameter appears to be very weakly, or even negatively, correlated; perhaps it is confusing to the respondents when not accompanied by a concurrent change in location.

The regression analysis used to obtain the results above probably did not reflect the full strength of the information contained in the data about location and filtering parameters. Time did not permit regrouping of the data to see if stronger results could be obtained.

The finding that source location probably is a significant parameter is certainly consistent with the differences among graphs in Appendix B. In terms of utilizing a functional formulation for the appliance noise problem, this result reinforces the need for a catalog of graphs, like those in Appendix B, which includes appliance location as an explicit parameter. Moreover, an important implication with respect to future research has been introduced by this location parameter. Unless all aspects of location can be simulated by means of stereophonic presentation, a prospect which seems doubtful of accomplishment, laboratory or earphone type experiments could not be used to investigate this location parameter; a realistic research house appears essential.

4.4 THE INFLUENCE OF ENVIRONMENTAL AND DEMOGRAPHIC CHARACTERISTICS OF RESPONDENTS

4.4.1. Introduction

Thus far this report has dealt exclusively with characteristics of the physical sound stimuli in determining the relative nuisance value they possess for the entire response group. The only parameters considered were the intensity, character, and source of the auditory stimulus. However, it is apparent that we are interested not only in the responses of the total group, but also in any clearly identifiable subgroups. These subgroups are of interest for two reasons: (1) For practical purposes, we may be able to treat some groups differently from others; and (2) the nature of the differences between subgroups is of theoretical interest because it indicates causes of complaints about noise.

Lacking any previous guidelines in this area it became necessary to deductively select all major characteristics that appeared to be relevant to the problem. Only those variables for which information could be gained with relative ease were employed. A population characteristic that cannot be easily identified in existing national census data can be of only limited practical use no matter how precisely correlated with information the researcher wishes to predict. For this reason, subtle psychological information is not considered in this section.

In order to study such characteristics intelligently it is necessary to remove the influence of acoustic differences that do not contribute to the relationship. In other words; age differences may cause different evaluations of the degree of bother associated with different characteristics of the sound. Education may reveal similar differences but with regard to intensity levels. For effective presentation and maximum stability of the sample in the case of age, all intensity levels could be combined, whereas, in the case of education all sound types could be combined. The problem of simplifying a fundamentally five dimensional array of data (degree of bother, source location, character and intensity of the stimulus, and the population characteristic under consideration) was handled with two different approaches. The first, a tabular method, revealed the relationships more clearly. A table was produced for each sound type at each level from every speaker location which revealed the distribution of bothersome responses within different characteristic groupings. For example, one table which has been prepared shows the number of respondents in each age bracket (of the code) who responded with each degree of bother at a fixed nominal level, speaker location and type of sound. In a similar manner there are tables for each of twenty different environmental and demographic characteristics of the sample for each possible combination of speaker location stimulus type, and stimulus intensity.

The second approach really is a complex of many methods and for various reasons was less successful than the tabular approach. They were all essentially multivariate analyses using the least-squares regression technique. Again, the methods utilized are explained in Appendix C.

Some of the tabular analyses utilize a metric designed primarily for the regression methodology. The metric is responsible for translating the verbal ordinal bothersome response categories into an interval scale (which, therefore, attaches a weight to each response). Again the design of this metric is discussed in the Appendix C.

The remainder of this chapter deals with the various characteristics which were investigated. The sub-headings of the chapter divide these variables up into logical topic areas, of which there are seven. For each subheading there is an inventory of the relevant variables, an explanation of why it was believed this type of information might contribute insight into the basic problem, and a discussion of the inferences gained from these variables. This discussion is divided into two section; first, a quick glance at the inferences made about those variables which were only partially explored, and second, a detailed statement about those variables which were completely investigated. The existance of the more complete investigation is identified in the inventory by the enclosure in brackets of variables handled in that manner. Variables discussed in this manner always are followed by some presentation of data (table or graph) as supporting evidence.

In the enumeration of the variables in the inventory, those items which are underlined were not analyzed in the tabular format (but were in the regression procedures) and, therefore, will have no attention addressed to them in this chapter. They are offered here only for the purpose of exhibiting all of the variables under consideration in this program in one place.

4.4.2. Variable Topic Area: Audiometry

i. High Frequency Hearing Loss: Summation in db of loss in both ears at 4000 cps.

ii. Mid Frequency Hearing Loss: Summation in db of loss in both ears at each of three frequencies; 500, 1000, and 2000 cps.

iii. Low Frequency Hearing Loss: Summation in db of loss in both ears at 125 and 250 cps.

iv. American Medical Association Binaural Hearing Loss: Standard clinical examination using 125, 250, 500, 1000, 2000, and 4000 cps tones. Designed to identify individuals with hearing difficulty sufficient to cause distress in social discourse. (Reference 13).

v. Audiometry Summation Code: Special code created from i, ii, and iii above to scale overall hearing loss of individuals. It placed much heavier emphasis on slight hearing loss than the AMA Binaural Hearing Loss scale.

Some relationship between an individual's ability to perceive sound and his nuisance evaluation of noise is intuitively obvious. One might expect that the nuisance value would vary directly with the keeness of hearing as measured by the various scales listed above if nuisance value is a function of the absolute level of perceived sound. However it is also obvious that overall hearing loss does not change the relative levels of particular sounds. It was further supposed that the nuisance value of a given sound type would be dependent upon the frequency range of that sound relative to the frequency range where hearing loss prevailed. The three loss-variables (see inventory above) which subdivided the frequency range are potentially useful where hearing loss occurred in only one part of the audible spectrum or was predominant in one part.

The audiometry examination was discussed in section 1.2 and the examination data form is displayed in Appendix A.

From an examination of the tabulated data, the same general inference appears to hold for all five audiometry variables. The observed nuisance value could not be related to the frequency of the sound stimulus or the frequency range of hearing loss. These observations may have been hampered by the limited range of the sound types and the relative rarity of occurrence of more than slight hearing loss in the sample.

An inference can be drawn relating extreme loss with a greater amount of bother. This applied to all audiometry variables utilized, at all intensity levels, from all sound sources, and for all sound types but was least apparent in the case of White Noise. (Of the eight sound types investigated White Noise contained a broader frequency spectrum and a larger proportion of high-frequency energy.)

These inferences are based on weak and not always consistent trends in the tabulated data and are subject to much individual variation from respondent to respondent.

4.4.3. Variable Topic Area: Architecture

i.	House Construction Material:	Classified into the following categories; Solid Brick, Brick Face Masonry, Brick Face Frame, Wood Frace Frame, Shingle Face Frame and Stucco Face Frame.
ii.	House Value:	Estimated market value of house including property.
iii.	Basement:	Classified into the following: Full, Partial, Crawl Space, None.
iv.	Roof Construction:	Classified into the following: wood-shingle, asphalt shingle, built-up.
v.	Number of Dwelling Units:	Number of completely independent family units within a structure.
vi.	Number of Levels:	Number of levels exclusive of the basement with a split-level home separately categorized.

The characteristics of the structure in which an individual dwells may effect noise nuisance judgments substantially in two independent manners. From the social and psychological point of view, the general atmosphere of the home environment may be a contributing factor. More directly, however, the structure of the house will have specific acoustical properties. It was hypothesized that a house's quietness or noisyness could effect the respondents general frame of reference in our experiment. For example, a person accustomed to an acoustically well-designed house would possibly be less inclined to accept a sound at a given level than respondents from other types of houses.

Detailed information on the collection of data dealing with the respondents own home (which was facilitated by the architectural field survey) is to be found in section 3.5. Within that section there is also a discussion of the representativeness of the homes of our respondents of those in both the north central U.S. and the entire nation.

Structures made of heavier and usually more expensive construction materials generally housed respondents who were less bothered by our sounds. This effect was more especially true of those persons dwelling in solid brick or brick face masonry structures and directly contrary to the hypothesis suggested above. The differences between brick face frame and other types of frame were not as marked. This was probably due to the fact that brick facing is not acoustically much of an improvement over other types of frame. Obviously this is an area requiring much more intensive study of the acoustic properties of structural materials and the relative care with which the houses were actually built. We can only use this variable as a very rough gauge of the overall effect we are searching for.

Despite the fact that more expensive building materials seemed to be related to our bother index, overall house-value was not. This is completely consistent with the income effect to be noted in the demographic section (as the value of the residence an individual lives in is highly correlated with income).

4.4.4. Variable Topic Area: Sample Methodology

The Quality and Non-Quality Research House Samples:

For methodological purposes it is important to ascertain if the various deviations of our sample from a random sample had a distorting effect on our data. A comparison is, therefore, made of the quality and non-quality samples. It is sufficient to indicate at this point that no strong bias was observed which tended to differentiate the two sample strata in question. This statement is with regard to the two groups evaluation of the nuisance associated with auditory stimuli only and is subject to the reservations about valid sampling procedures discussed in section 2.

4.4.5. Variable Topic Area: Psychoacoustical

i.	(Psychoacoustic Index):	Discussed in the following text.
ii.	Spend Additional for Quiet:	Willingness to spend additional money, when purchasing a new appliance, for quieter operation.
iii.	Overall Appliance Complaint:	Complaint regarding appliance noise divided into those which were spontaneous and those which followed a specific probe.

In this section we are testing the relationship between an individual's verbal judgment about himself (in an interview situation) and his psychophysical response to sounds in the research house. To do this we first constructed an index, called the Psychoacoustic Index, to combine the answers from several interview questions to form a single 4 point scale.

The questions used in the index construction are as follows:

1. Generally speaking are you happy with the appliances you use..., or do you have any complaints about the way appliances work?

   (If a complaint) What is your complaint

2. (If sound not mentioned in question above) Do you have any complaints about the sounds any appliances make?

3. Do you think a manufacturer should spend his time...making his appliances quieter or improving his appliances in some other way?

4. Suppose you were buying a new home appliance. You find a model that is as good as the others on the market, and also is substantially quieter... Would you be willing to spend any additional money for the quieter model?

Using the four questions above a four point psychoacoustic index was created as follows:

1. Noise mentioned as a complaint in Q 1(spontaneous complaint; regarded as very highly noise sensitive)

2. Q 2 answered affirmatively. (complaint after probe; regarded as having some sensitivity to noise above average)

3. Either Q 3 or Q 4 answered affirmatively (little sensitivity to appliance noise; but still some degree of concern).

4. Conditions 1 - 3 do not apply (no noise sensitivity).

There are a large number of indications that spontaneous complainers (i.e., scale position 1) are more bothered by the sounds in our study than the other 3 categories, but differences among the three remaining categories are very slight. The details of the greater bother for spontaneous complainers differs to a minor degree; e.g., for Dryer sounds the difference is greatest at the highest sound levels, while for white noise the difference is more marked at low sound levels.

This finding has several levels of significance for our study. First, it is certain (see Table 6, page 26 of this report) that our method of obtaining our sample over-represents those respondents who have spontaneous complaints regarding the noise. Thus our overall results overstate the degree of bother as a multiplicative function of these combined biases. Future studies or further analysis of this data might control this effect, either statistically by a weighting function, or experimentally by enticing respondents to the interview without indicating noise as the central interest. (Table 6 indicates that the research house group is not biased in terms of other complaints.)

Second, this finding is a kind of "validation" of the interview technique, since it indicates a connection between the interview response and an objective psychophysical response.

Third, it indicates that at least a portion of the spontaneous complaints regarding noise (presumably generalizing to complaints received by dealers and service facilities) are to be explained in terms of individual differences among consumers. Appliances producing identical sounds, installed in identical situations, will give rise to varying complaint levels depending upon a psychological characteristic of the owner.

The willingness to spend additional for quieter appliances appeared as a very important influence in the AID analysis, where it was second in importance only to a separation of sounds into intensity levels above 55db and below 55db in predicting degree of bother. (see Appendix C for explanation of AID analysis)

4.4.6. Variable Topic Area: Research House Interview Conditions

i. Number of respondents at session: Varied between one and eleven. Usually did not exceed seven. Mode four.

ii. Frame of Reference:	To determine if the wording of the nuisance evaluation question plays any role. Some respondents were asked our key questions in terms of "annoyance" while the majority were asked in terms of "bother" (i.e., "Did that sound 'annoy' you?" Did that sound 'bother' you?"
iii. Time of Day:	Sessions were held morning,afternoon and evening at respectively 10 AM, 1 PM and 7:30 (or 8:00) PM.
iv. Seat Position:	Discussion of this variable, concerned with variation of stimuli about the living room of the research house, found in Chapter 1.

Except for seat position these variables are fundamentally related to psychology. Does the environment of the respondent during the interview or the phrasing of similar questions have any impact on their evaluation of the nuisance value of a noise? If so, any future experiments along this avenue of research would have to take steps towards their more effective control.

The number of respondents at a session did not seem to substantially effect responses. This is as expected inasmuch as the interviews were written and interviewers acted to direct conversation in a manner that would disallow group evaluation of sounds. However, where only a single respondent was present responses did tend more toward extremes. It would appear, then, that the mere presence of other respondents tended to produce more central judgments.

Those respondents who were asked if they were "annoyed" tended to express judgments indicating more irritation than those who were asked if they were "bothered."

The time of day of the experiment did not seem to exert any impact with the exception of the fact that there was a very slight indication that respondents interviewed in the evening were slightly more bothered by the same sounds which were offered to different respondents during the daylight period. This may be a function of greater irritability of people as the day grows later and they become more tired, or it may be related to the generally lower levels of prevailing background noise in the evening.

Seat position is a highly complex variable which could not be studied in detail at this point. Where actual sound intensity levels were used in this study, they counted to reflect the actual levels at each seat. An extension of the Whirlpool study might allow for the lengthy analysis required.

4.4.7. Variable Topic Area: Demographic Characteristics of the Sample

These variables are mostly self-explanatory.

i.	age.	
ii.	(education)	
iii.	(income)	There were two estimated sources. One derived from the relationship with the house value of the respondent, and the second was estimated by the field interviewer. (Income is a sensitive subject with most persons and we did not wish to endanger the success of the entire program in attempting to elicit this single piece of information by direct questioning.)
iv.	sex	Insufficient number of males to warrant closer consideration of this variable.
v.	race	Insufficient number of non-whites to warrant closer consideration of this variable.

Where acceptability is a function of demographic features it may be possible to vary the noise in question to best suit the normative judgments of that section of the population in which a particular product will be marketed. In any event, if different sections of the population do vary in their judgment it is well to at least expect where greater or lessor displeasure with regard to appliance sounds will be prevalent.

There appears to be a strong trend toward bothersomeness increasing with educational levels. There can be little doubt that college educated individuals are more bothered with the sound stimuli we presented than the remainder of the population. There was some indication, however, that the degree of completion of High School education was not a factor. Those who completed High School, in fact, seemed slightly less bothered than those who attended High School but received no diploma. Since our sample was so highly educated (relative to the national population) it was difficult to analyze in detail the trends of the less educated.

It is noteworthy that these differences between educational levels with regard to bother judgments of sounds was clearest in the case of the White Noise stimuli. However, the same general trend appears to prevail for all stimulus types at all levels and from all speaker locations. Different types of sounds, therefore, do not seem to be differentially more or less bothersome to respondents with different educational backgrounds.

Because of the limited sample size, it cannot be determined if these increases in bother associated with education are completely uniform or can be defined precisely. Hence, it cannot be known at this point if each additional year of education contributes to the relationship being considered here. However, from Table 10 it is quite clear that at least three major educational groups do vary significantly.

TABLE 10

MEAN DEGREE OF BOTHER BY EDUCATION*
(Expressed as a Per Cent of Total Bother)

Nominal Level[a]	Final School Attended in Education of Respondent: Elementary	Secondary	College
0	6.4%	10.7%	13.0%
3	15.0	14.7	19.6
6	too few cases	26.3	33.6
9	30.3	34.0	41.3
12	too few cases	43.0	50.6
15	53.6	55.0	61.9
18	45.6	57.3	69.9
All Levels	25.3	33.6	53.3
Number of Cases	567	4814	6147

*This table is dependent upon the metric specifically designed to scale the verbal bother responses. A discussion of this metric is found in Appendix C. It suffices to indicate that "not bothered" has a value of 1.0 and "extremely bothered" has a value of 4.0. In this table if all respondents indicated "not bothered," 1.0 expressed as a percent of total bother would be 0.0% whereas if the converse were true and all respondents were "extremely bothered" the table entry would be 100.0%. Since each entry in the table is expressed as a percent of total bother they cannot be summed in any direction. All stimulus types are merged.

[a]The concept of nominal level is discussed in the introduction; in most cases they tend to increase (at least electrically) 3 db with each step.

After intensive investigation it can safely be concluded that within this sample there was no systematic relationship between the broad income brackets of respondents and the degree of bother registered to auditory stimuli. The apparently random differences of income groups were true for all of the eight sound types. It was also true of any formulation for speaker location. Furthermore, both of the preceding observations were true at all intensity levels.

Viewed simultaneously the income and education variables have an unusual impact. Initially, it might be pointed out that it is only with the rather special sample we have that such a circumstance might occur; education being related to auditory bother and income not so related. In most samples income and education is highly correlated; in this sample large numbers of graduate students (married and therefore not a part of the institutional population which had been removed from the sample) at high educational levels have small incomes. Usually it is rather difficult to divorce the contribution of education from income and vice versa when they appear to contribute to the relationship being studied.

In the present case it is possible to state with a rare degree of confidence that it is education and not income which is contributing importantly to the relationship.

4.4.8. Variable Topic Area: Home Environment

i.	Neighborhood Noise Level:	Estimated without aid of sound level meter during the architectural field survey. All estimates made by Bruce Erickson. Categorized with verbal scale varying from "very noisy" to "very quiet."
ii.	Appliances Located in Kitchen:	Enumeration of all noise producing major appliances located in kitchen of home.
iii.	(Washing Machine Location):	basement, kitchen, utility room
iv.	Appliance Ownership:	Number of appliances owned by respondent (or present in household in the case of rented residence). "Appliance" is defined by those items listed on page 2 of field interview questionnaire;
v.	Washing Machine Ownership:	Special variable to separate owner from non-owner.
vi.	Washing Machine Type Owned:	Special variable to separate owner of automatic washer from wringer type washer.

This topic area may exert impact on respondent judgment for a variety of reasons; fundamentally, however, they are all psychoacoustic in nature. The attempt is made to determine if the specific noises and overall noise levels to which respondents are accustomed seriously affects their evaluation of our stimuli. The neighborhood noise level variable actually is an attempt to establish a relationship between the background level to which the respondent is adapted and the overall sensitivity of the respondent.

Persons with a greater number of appliances located in their kitchen are less bothered with all of our stimulus types at lower intensity levels. At higher levels, however, there is a tendency for this same group of people to be more bothered. This observation is more pronounced in those sound types which had more of the high frequency spectrum represented in them. Speaker location appeared to play a role too; however, the number of cases at the various speaker locations are not sufficient to justify the statement of any inferences.

The location of the washing machine within the home of the respondent is a factor which deductively appeared to be extremely potent in influencing noise nuisance judgments. From intuition it might be inferred that the more centrally located within the living quarters of the household is this appliance, the more impact it might have. Whether this closeness would result in more or less bothered reactions was not, however, clearly deducible. The initial formulation of the data revealed the following information:

CUMMULATIVE DEGREE OF BOTHER BY WASHING MACHINE LOCATION
(all sound types, speaker locations and intensity levels merged)

Washing Machine Location	Cummulative Degree of Bother			
	not bothered	at most little bothered	at most moderately bothered	at most extremely bothered
Utility Room	48%	64%	86%	100%
Basement	38	61	82	100
Kitchen	37	60	81	100

At this level the differences are not very striking. This variable proved fascinating in that it revealed just how subtle and elusive some of the implications of this investigation can be. Further investigation of this variable by intensity levels revealed differences which were not highly significant. However, when speaker location and stimulus type were intensely investigated in the proper focus some highly interesting revelations are made.

The fact that the application of stronger filters which principally reduce the high frequency components has the effect of decreasing bothersome levels of response has been discussed previously and is of no great import here. The important implication to be noted here is that respondents having washing machines located in the kitchen are far more bothered by non-filtered sounds than those having basement located washing machines. Not presented here is the group with washing machines in their utility room. As might be expected their position is roughly intermediate between kitchen and basement with regard to the effect noted.

CUMULATIVE DEGREE OF BOTHER FOR SELECTED SPEAKER LOCATIONS AND SOUND TYPES BY WASHING MACHINE LOCATION
(all intensity levels merged)

Washing Machine Location: Kitchen

Speaker Location and Sound Type	Cumulative Degree of Bother: not	at most little	at most moderate	at most extreme
non-filtered locations; agitate[1]	26%	38%	65%	100%
Filtered locations; agitate[2]	40	65	81	100
all locations; agitate filtered	44	77	92	100

Washing Machine Location: Basement

Speaker Location and Sound Type	not	at most little	at most moderate	at most extreme
non-filtered locations; agitate	35%	59%	86%	100%
filtered locations; agitate	43	67	87	100
all locations; agitate filtered	49	72	90	100

[1] Living room unfiltered and kitchen

[2] Living room filtered and basement

The application of high frequency filters, then, would considerably reduce the degree of bother for those groups of respondents who have appliances located in household areas where those frequencies are likely to be audible. In fact, one might almost conclude that those who are adapted to hearing the high frequency sounds note a marked improvement when that portion of the spectrum is removed. Those who are not accustomed to the high frequencies do not clearly note the difference from their home appliances in the unfiltered situations nor do they note any considerable improvement after the filter is applied.

One final trial was made to see if the effect of intensity could also be studied in this light. The number of cases with this number of control categories was very small and the inferences are, therefore, somewhat dubious. In this light, however, it is still interesting to note the following points. At low intensity levels the difference between agitate and agitate filter bother levels were nearly non-existent regardless of washing machine location in the respondents home. At higher levels, however, the differences between agitate and agitate filter became apparent and coupled with this were still larger differences (in the same direction as noted before) for those with kitchen located washing machines.

4.5 CONSISTENCY IN RESPONDENT JUDGMENT OF THE DEGREE OF BOTHER ASSOCIATED WITH AUDITORY STIMULI

As utilized in this chapter the term consistency refers primarily to the reproducibility of individual judgment. It is important to recognize that consistency after all responses have been aggregated is very different from the individual consistency under investigation here. The vital difference is depicted in the following example: (although the data presented here are hypothetical, in parts of this chapter the identical format is employed in the analysis of actual data).

Let us assume a sample of 100 respondents. They are asked on two separate occasions to record which of a pair of two sounds, to be referred to here as sound "A" and sound "B", is more bothersome. Let us assume further that in each instance 50 respondents indicated "A" and 50 "B." Checking for consistency after these replies have been aggregated would result in the conclusion that the group was completely consistent. Indeed, it is clearly true that the group, viewed as a whole, was consistent. However, if individual consistency is sought the conclusion may be far different. It might well turn out that of the 50 who selected "A" as more bothersome in the first exposure, 10 respondents changed their preference to "B" when reexposed to the pair of sounds; likewise 10 of the original "B" group may have switched to "A." In a more critical light the overall consistency would be stated as only 80%. Note that the aggregated analysis which appears to demonstrate 100% consistency, when reevaluated employing the second method, may result in anything from 100% consistency to 100% inconsistency.

Individual consistency is of crucial importance to this study and will, therefore, be discussed at considerable length in this chapter. If a reasonable degree of individual consistency is not demonstrable this would tend to cast doubt on the general validity of the current program and additional programs proposed in the area of psychoacoustics.

It is to be recalled that one of the premises of this program was that realistic consumer responses to noise could be expected only within an environment that approximated real life situations. As this approach is unique it is necessary to show not only that the respondent tends to make accurate judgments in this non-laboratory setting but, moreover, that such judgments or preferences are not whimsical but rather reflect a constant set of values that the respondent may be called upon to repeat.

Since there was full awareness of this problem from the outset of this study the entire research program was constructed so that a full investigation was possible at the analysis stage. In the discussion which follows the results of a number of methods of exploring consistency are reported. The procedure will be to begin with methods which make the least demands upon the respondent (in order to maintain a consistent reply) and proceed to those methods making greater demands on the respondent.

4.5.1. Analyzing Consistency With Paired Stimuli

Perhaps the simplest consistency related to judgments of auditory stimuli is the ability to chose repeatedly between a pair of sounds as to which is more bothersome. This approach required a judgment of relative bother, not absolute. The sole requirement is that one is able to decide clearly which of a pair of sounds is more disliked.

The data of the research house experiment testing expressly for this form of consistency reveal that respondents are capable in the majority of instances (in fact nearly 80% of the time) of giving reproducible judgments. This is despite the fact that every effort was made to insure that consistency could not be based on irrelevant factors. Repetitions of paired sounds were randomly distributed throughout sections two and three of the interview (Appendix A) and separated by up to one hour in time. In addition, the order in which the pair of stimuli was offered was reversed for the second presentation.

Table 11 on the following page reports the proportion of consistent responses within each of the 29 pairs of sounds offered individually for all sound sources at which each of the pairs was presented.

It can be observed from Table 11 that the proportion of consistent responses ranges from 61 to 87%. This range, however, does not permit the assumption that the pairing of certain stimuli result in a greater consistency. Table 12, below, represents data excerpted from Table 11 which has been selected and grouped in a manner that optimally demonstrates the lack of uniform proportions of consistency associated with respondent exposures to identical pairings of sounds at different nominal levels. It would be more accurate to state that the differences noted in Table 11 is more likely to be the result of error due to sample size.

It should be noted, however, that it is possible that differences in the nominal levels of the pairs, the degree to which "A" and "B" are matched in db and the location of the sound source may be mitigating against greater uniformity. The examination of the overall impact of speaker location is afforded by the data presented at the bottom of Table 11. Here again, with the possible exception of the Kitchen, differences are clearly within sampling error. The type of approach used in Table 12 is occasionally possible even within a fixed speaker location. The stimulus pairs of 120 cps. and 4000 cps at the living room unfiltered location may be observed to vary in consistency from 63% to 86%. It may safely be concluded that the absence of uniformity is not due to changes in the speaker location.

TABLE 11

RESPONDENT CONSISTENCY IN SELECTING THE MORE DISLIKED
OF A PAIR OF SOUNDS BY SOUND TYPE AND SOURCE

Description of Paired Stimuli[1]			Proportion of consistent responses by sound source				
A[2]	B	Nominal level[3]	All[4]	Living room un-filtered	Living room filtered	Kitchen	Base-ment
R	SBF	15	81%	78%	76%	86%	5
A	WN	12	82	86	76	83	
D	O	15	79	80	70	84	
R	AF	15	77	78	74	78	
O+PT	O	18	75	69	62	80	
120	4000	15	85	86	72	96	
WN	R	15	83	88	75	86	
O	R	12	77	73	70	88	
AF	WN	9	89	83	95	89	
SBF	O+PT	12	85	83	83	87	
AF	A	9	80	77	76	86	
A	R	15	64	65	55	71	
120	4000[6]	9	80	63	88	87	
A	AF	15	80	81	71	87	
AF	O	18	65	64	65	65	
4000	120	9	75	71	74	79	
120	256	15	87		86		87%
A	WN	15	70		70		70
O	SBF	18	80		80		79
O+PT	SBF	18	80		75		81
O	WN	12	78		78		78
D	AF	18	80		71		82
D	O	18	74		72		75
R	O+PT	15	71		60		75
D	WN	9	73		80		71
A	SBF	15	72		60		76
D	SBF	12	71		74		70
D	A	12	67		73		68
All paired sound types aggregated			77	76	73	83	75

[1]Key to the abbreviations used in describing the stimuli:

R: refrigerator

SBF: standard band of frequencies

TABLE 11 (Concluded)

A: washing machine, agitate cycle
WN: white noise
D: dryer
O: band of noise
AF: "A" filtered
O+PT: "O" plus a pure tone superimposed
120; 256; 4000: 120 cycles per second, etc.

For a detailed description of these sounds refer to section 3.4 and Appendix A.

[2] "A" is always the sound heard first at the initial time the respondent was exposed to the particular pair of stimuli at that particular nominal level.

[3] Nominal level refers to the intensity of the sound. For a full discussion of this concept see section 3.4. It is to be noted that a given nominal level for a stimulus does not necessarily equate it in db with other stimuli at the same nominal level. In addition, the nominal level is designated in terms of the electrical level at which this sound appears on the recording tape. Each nominal level, by itself will vary in acoustical energy depending upon the speaker location and the respondent seat position. With such latitude for variation it is not surprising that the paired sounds do not often agree in db. This is useful, however, in that it allows for attempts at the determination of equi-bother points of sound pairs (this aspect of the study has not been investigated at the time of the writing of this report). The effect of the difference in db level between the pair of stimuli upon consistency is discussed later in this chapter.

[4] Each respondent heard the two stimuli comprising the pair from the same speaker source; nor did the repetition of the pair vary the source.

[5] Due to time limitations of the interview not all pairs were asked at all speaker locations; empty columns represent those situations which were not presented. The "All" column is composed in each case of all responses made to a given pair; therefore, the figures in this column could be accused of not being strictly comparable as they do reflect the sum of different items in all cases where pairs being compared do not cover the same speaker positions.

[6] This pair of stimuli, at level 9 was given four times. This case is composed of the pairing of the first and second exposures presented. The data are displayed separately in order to reveal the differences in consistency which may occur when stimuli pairs are repeated with greater frequency.

TABLE 12

SELECTIVE PRESENTATION OF DATA FROM TABLE 11

Description of Paired Stimuli			Proportion of Consistent Response	Consistency Differences Between Identical Pairs
A	B	Nominal level		
A	W	12	82	12%
A	W	15	70	
120 cps.	4000 cps.	9	80	
4000 cps.	120 cps.	9	75	10
120 cps.	4000 cps.	15	85	
O+PT	O	18	75	14
O	O+PT	9	61	
SBF	O+PT	12	85	5
O+PT	SBF	18	80	

Furthermore, no particular stimulus type has a constantly higher degree of consistency associated with it regardless of the other stimulus type it is paired with. See Table 13.

TABLE 13

AVERAGE CONSISTENCY FOR A STIMULUS TYPE INCLUDING ALL PAIRINGS IN WHICH THIS STIMULUS WAS UTILIZED

Stimulus Type[1]		Average Consistency
A	WN (2), SBF, D, AF (2), R	74%
R	O+PT, SBF, AF, WN, O, A	76
WN	A (2), O, D, R, AF	79
SBF	O, O+PT (2), A, D, R	78
O+PT	O (2), SBF (2), R	74
D	AF, O (2), WN, SBF, R	74
AF	D, WN, A (2), O	79
O	SBF, O+PT (2), WN, D, (2), R, AF	74

[1] Abbreviations used for stimulus type are identical to those used in Table 11 (a key for them is found at footnote 1 of that table).

[2] Numbers in brackets following the abbreviated stimulus type indicate multiple pairing of the two stimuli types involved.

The nominal intensity levels at which paired sounds are presented also do not effect individual consistency. All sound pairs offered at the nominal level of 9 had an average consistency of 76%, at 12 it was 77%, at 15 it was 77% and at 18 it was 76%.

Attention was also directed toward determining if consistency differed substantially with special sections of the population. No clear difference in the proportion of consistency was observed within different age, education or income groups of the sample. In addition, the psychoacoustic index did not act to separate groups having more or less consistency of response.

It is of interest to note that those respondents with poorer hearing were on the whole no less consistent than the rest of the population. The only instance where hearing deficiencies appeared to have a marked impact upon consistency is the case where high frequency stimuli are presented as part of the pair and even here the impact is not uniform in nature. See Table 14.

TABLE 14

RESPONDENT CONSISTENCY IN JUDGING THE PAIRING OF STIMULI
120 CPS AND 4000 CPS BY HIGH FREQUENCY LOSS[1]

Respondent Hearing Ability in the High Frequency Range	Proportion of Consistent Responses		
	Nominal Level: 15	9; first exposure	9; second exposure
Excellent	84%	86%	80%
Good to Average	87	79	78
Sub-average and poor	85	63	63

[1]The method of calculating loss is discussed in section 4.4. All sounds presented in this instance included all sound source locations excepting the kitchen. High frequency loss is employed because the relationship is demonstrated most strongly.

To this juncture the question of the relative intensity of the sounds paired for consistency comparisons has been avoided. Due to the different actual decibel levels of different sound types at the same nominal levels, the factor of speaker location variation and the effect of the different seat positions the respondents chose to sit in, the first sound with respect to the second in any given pair fluctuated anywhere between ± 20 decibels. In Table 15 the variation for a typical pair may be observed.

TABLE 15

DECIBEL LEVELS FOR PAIRED SOUND STIMULI
120 and 256 CPS AT NOMINAL LEVEL 15 BY
SOUND SOURCE LOCATION

Stimuli DB Levels		Intensity Level Relationship		Location of the source[1]
120 cps	256 cps	Stimulus with higher level	Number of DB more intense	
67	54	120 cps	13	basement
67	55	120 cps	12	basement
67	64	120 cps	3	basement
66	65	120 cps	1	basement
65	65	neither	0	basement
65	66	256 cps	1	basement
67	69	256 cps	2	basement
64	67	256 cps	3	basement
61	66	256 cps	5	living room unfiltered
58	66	256 cps	8	living room unfiltered
57	70	256 cps	13	living room unfiltered

[1]The number of sources utilized in any given pair is guided by the track position of the stimuli on the tape recorder. At one time or another all of the five mode-type sources are employed in these paired stimuli questions; however, at any given point only two to three locations are used in a single pairing.

It is intuitively obvious that for each individual there must occur a point at which sound "A" seems so much louder than sound "B" that it is judged as more bothersome regardless of the characteristics of the paired sounds in question. Likewise, one would expect to find a range of intensity level differences for the paired sounds where the bothersome judgment becomes difficult for an individual. Within such a region, consistency in terms of repetitive judgments would no doubt be inclined to decrease for a given individual. If all respondents acted uniformly in this respect, the reported consistency should logically be expected to decline even though the population taken as a whole might still report 50% more bothered by stimulus "A" and 50% by stimulus "B". If the respondents do not act uniformly, that is, if this "equi-bother" range is differently placed for each individual then the aggregated consistency of individuals need not necessarily be effected by the distribution of consistent responses regarding whether "A" or "B" was most bothersome. Table 16 is a hypothetical flow diagram which depicts the two possible extreme situations previously discussed.

TABLE 16

PERCENTAGE OF RESPONSES IN TWO HYPOTHETICAL SITUATIONS WHICH DEMONSTRATE THE EXTREMES OF THE POSSIBLE RELATIONSHIPS BETWEEN THE DISTRIBUTION OF CONSISTENT BOTHERSOME RESPONSES IN A PAIR OF STIMULI AND OVERALL RESPONSE CONSISTENCY

Case a: Wide Difference of Individual Judgment of the "Equi-Bother" Point: Consistency Level is Relatively Static.

Level of Stimulus "A" in Decibels	Level of Stimulus "B" is fixed at 55 decibels		
	"B" consistently judged more bothersome	response is inconsistent	"A" consistently judged more bothersome
30	100	0	0
40	75	20	5
	15	10	10
45	60	25	15
	15	10	15
50	45	25	30
	10	15	10
55	35	25	40
	10	15	10
60	25	25	50
	20		25
65	5	20	75
75	0	0	100

Case b: Individuals Demonstrate Overall Agreement As To The Equi-Bother Range: Consistency Fluctuates Considerably.

Level of Stimulus "A" in Decibels	"B" consistently judged more bothersome	response is inconsistent	"A" consistently judged more bothersome
30	100	0	0
40	90	5	5
	25		
45	40	25	35
	25		
50	10	5	85
60	0	0	100

It can be seen from Case "a" that here individuals do not concur in their opinion as to the decibel difference required between the intensity levels of stimulus "A" and stimulus "B" to cause them to approach the equi-bother point. As different persons approach their individual equi-bother point they have a high probability of contributing an inconsistent response to the aggregates. As the level of one stimulus continues to increase relative to the second some individuals become consistent while other individuals move into their own equi-bother range and fill the space just vacated. For this reason the proportion of inconsistent responses remains fairly constant over a broad range of decibel differences between stimuli.

Case "b" is the alternative extreme; first the overwhelming majority of respondents agree that stimulus "B" is more bothersome. After a slight rise in the level of stimulus "A" a very great proportion of the respondents exhibit inconsistent responses and the remainder are equally distributed between stimuli "A" and "B." With the continued rise in the level of "A" all (or nearly all) shift their judgment such that now "A" is judged as more bothersome. This would be the best demonstration conceivable of all individuals agreeing as to where the equi-bother point is and the proportion of inconsistent responses, therefore, varies accordingly.

In actual situations utilizing our twenty-nine experimental pairings it happens that case "a" is far more common than case "b". In addition, when case "b" type situations do occur the condition never existed where the differences in decibels range for that pairing was broad enough to demonstrate the entire pattern constituting the hypothetical situation. Table 17, however, does demonstrate the general principle fairly well.

TABLE 17

DEMONSTRATION OF AGREEMENT OR DISAGREEMENT IN REAL DATA

Case "a": Percentage Disagreement; Paired Stimuli are Washing Machine Agitate and White Noise

Intensity Levels		Consistency		
Stimulus at Higher Level	Range of DB Higher[1]	Over-all	Proportion Consistent More Bothered by	
			Agitate	White Noise
White Noise	11	84%	6%	94%
Agitate	1-7	80	76	74
Agitate	8-18	78	59	41

[1]Narrower groupings of DB ranges did not reveal information substantially different from that above; therefore, data has been compressed into the three groups presented.

TABLE 17 (Concluded)

Case "b": Percentage Agreement; Paired Stimuli are 120 cps and 4000 cps

Stimulus at Higher Level	Range of Decibels Higher	Loudness Levels[1]	Overall	Proportion of Consistent Respondents More Bothered by	
				4000 cps	120 cps
120 cps	-10 to + 10	4000 cps is louder	77%	90%	10%
120 cps	11 to 17	Equal loudness (± 3 db)	82	88	12
120 cps	22 to 35	120 cps is louder	52	46	54

[1]Equal loudness occurs at about the point where 120 cps is 14 db more intense than 4000 cps. This varies somewhat with the intensity level of the sound. The 14 is true at a db range of 55-65 and the plus or minus figure allows for equal loudness to be reasonably true across the entire db range which these two paired sounds were given at in the above analysis (being roughly 35 to 80 db).

In case "b" of Table 17 the marked plunge in consistency can be noted when stimuli involved are equal in bother for the portion of the population which remains consistent. Most respondents must be near their equi-bother point here and apparently some persons just have a keener ability to make distinctions between the pair. Note that it cannot be asserted if the trend here will continue as it did in the analogous hypothetical case; that is, if the difference between the levels of stimulus A and stimulus B is expanded even further that the 4000 cps sound will uniformly be chosen as more bothersome than the 120 cps sound and that this will occur during a narrow range of decibel changes. In this latter suggested situation, consistent with the proposed theory, the overall consistency of the respondent population would also be expected to rise sharply. Therefore, it might be said that this equi-bother concept is double faceted and is not a point at all but a range defined by two points. The first at which the equi-bother point is arrived at (this being demonstrated on Table 17, case "b" where the 120 cps sound is found equally as bothersome as the 4000 cps sound and the intensity level of the 120 cps sound is 22 to 35 decibels higher) and the second at which the equi-bother point is passed. This last point is not demonstrated on Table 17 case "b" but it is observed that this equi-bother range is at least 13 decibels wide.

A portion of this table has been presented in terms of loudness also to demonstrate that the equal loudness point is not the same as the equi-bother point (or range). This again reinforces the conclusion that the loudness function may be quite different from the bother function for a particular frame of reference when the latter is clearly defined.

It is sufficient to merely note that case "a" of Table 17 is equivalent to case "a" in the hypothetical situation mapped on Table 16. Like that section of Table 16, consistency does not vary to any great degree although the choice by respondents as to which sound is more bothersome shows considerable variations.

Relating the preceding analysis to Table 11 reveals that the consistency in that table does not appear to vary with the sound types which were paired, probably because this consistency is dependent upon how closely the given levels of the two stimuli approached the equi-bother range; (and where they actually did, for what percentage of the overall response to this pair this condition was true). Since it can be demonstrated that the equi-bother point in several instances was not approached (for the majority of the sample) it is understandable that in these cases the overall consistency is higher than in those cases where the equi-bother point was approximated. Table 18 shows above average consistency where the equi-bother point was clearly never reached.

TABLE 18

DEMONSTRATION OF A CASE IN WHICH THE INTENSITY DIFFERENCES IN A PAIR OF STIMULI PRESENTED WAS NOT BROAD ENOUGH TO ENCOMPASS THE EQUI-BOTHER RANGE FOR THE MAJORITY OF THE POPULATION
(The pair of stimuli are Washing Machine Agitate Cycle and the same stimulus when filtered)

Intensity Levels		Consistency		
Stimulus at Higher Levels	Range of Decibels Higher	Over-all	Proportion of Consistent Respondents More Bothered by	
			Agitate	Agitate Filtered
Agitate	0-6	82%	25%	75%
Agitate-Filtered	1-3	78	28	72
Agitate-Filtered	4-6	80	28	72

There is a great deal of significance clearly associated with the extent of the range in which the majority of bother judgments are static. It serves to reinforce the analysis which was undertaken in the earlier sections of this report in their significance. In terms of the actual methods of reducing complaint levels of an appliance noise it would be better to attempt to transform any 4000 cps sound so that it more closely approaches the 120 cps sound rather than attempting any reduction of less than 27 decibels (at least in the case offered on Table 17) insofar as the majority of the population would register no significant difference in bother within that range. However, in terms of the actual transformation of the sounds, to what extent a sound would keep the properties of the 120 cps sound if the noise produced could not be exactly made to represent that sound is not known from this study. This entire subject is out of the range of this section which intends to deal with consistency.

Beyond this, however, the degree of consistency for each individual as they approach the equi-bother range is an intriguing subject for a laboratory experiment but out of the scope of the present study too. Stricter tests for consistency in the current study have been reserved for the judgments of respondents to a single sound when repeated and allied areas to be discussed in the immediately succeeding sections.

4.5.2. Analysis of Consistent Change Associated with Changes in Stimulus Intensity

The analysis of consistency in this respect is concerned with whether respondents can evaluate successfully the increased intensity of a stimulus whose audio character is the same in all other respects. For example, if a sound is presented at 50 decibels and the respondent indicates it causes a "little bother;" then if the same sound is presented at 53 decibels, in order to remain consistent this respondent must judge this second sound to be <u>at least</u> a "little bothersome."

The method employed in this analysis was to check the response for each individual for the lowest intensity level at which the sound was presented against each successively higher level. There were seven nominal levels utilized and each nominal level was three decibels separated from the one just below it.

Nominal Levels Compared		Difference in Decibels
1	2	
0	3	3
0	6	6
0	9	9
0	12	12
0	15	15
0	18	18

It is to be understood that the exact level of intensity in decibels for each nominal level will vary with the type of sound in question for all respondents and will further vary for each respondent depending upon the speaker location and seat position in the house. However, for each respondent within any given sound type each successively higher nominal level is exactly three decibels higher.

Since in all cases comparisons are being made against the lowest level, a large percentage of the respondent population is disqualified from the remainder of the analysis. This is true because those respondents who indicated they were "not bothered" at the lowest level were not capable of being inconsistent at each higher level compared with it. (Other procedures might have been employed but the one described was practical for our data.)

The level of respondent consistency did not vary substantially with the location of the sound source except for the basement. Responses made by individuals to sounds originating in the basement, for any given sound type, were from 5% to 10% more inconsistent than those from other locations. This might be because the sound levels for the lowest nominal level were lower in the basement than the other locations.

It also must be noted that the individuals being checked for consistency here are those that differ from the majority opinion in that they originally are bothered at low levels where other respondents are not; no doubt some of these respondent replies merely represent errors in the recording of their judgment on the interview form. It is likely that these recording errors account for a large proportion of what might be called residual error; errors which remain despite very large differences in intensity of the two stimuli being compared. This can be observed in Table 19.

Table 20 presents the inconsistency associated with changes in level for the first three nominal level steps; that is, those who are inconsistent in going from levels 0 to 3, 0-6, and 0-9. Because the basement responses were so different they have been eliminated.

Before leaving the subject area of respondent consistency within change of intensity level for a given stimulus, it is of interest to note the extent to which inconsistencies registered with each change of the decibel difference are made by the same individuals. Table 21 presents this type of concept in the form of a flow diagram. It utilizes the data on exposure of respondents to the Washing Machine Agitate sound at all of the available nominal levels being 0,3,9,12,15, and 18. All the itemized inconsistencies displayed here are a result of comparison with the lowest nominal level. Inconsistencies of the form of the example listed below are not accounted for.

	level 0	level 3	level 9
bother level	little	moderate	little

The only categories which are accounted for in this inconsistency formulation are those where consistency at any of the four higher nominal levels fall below those judgments made at the lowest nominal level (i.e., "0").

TABLE 19

RESIDUAL INCONSISTENCY OF RESPONDENTS ASSOCIATED WITH CHANGES OF INTENSITY LEVEL FOR A GIVEN STIMULUS BY STIMULUS TYPE[1]

Stimulus Type	Proportion Inconsistent[2]
Washing Machine, Agitate (A)	14%
Refrigerator (R)	0
Dryer (D)	4
White Noise[3] (WN)	0
Standard Band of Frequencies (SBF)	4
Octave Band (O)	18
Octave Band with Pure Tone Superimposed (O+PT)	0
Washing Machine, Agitate Filtered[3]	16
Average for All Sound Types	7

[1]For all speaker locations at which the appropriate stimuli were presented. The range being compared here was 18 decibels wide.

[2]Proportion of those capable of being inconsistent who were so; this means that those who indicated not bothered at the lowest nominal level were excluded.

[3]Because either the highest or lowest nominal level was not included in the tape program the stimuli being compared are only 15 decibels apart.

TABLE 20

INCONSISTENCY ASSOCIATED WITH CHANGES IN INTENSITY OF 3, 6 AND 9 DECIBELS BY STIMULUS TYPE (BASEMENT RESPONSES HAVE BEEN EXCLUDED)[1]

Stimulus Type	Proportion Inconsistent[2] Intensity Increase: 3 db	6 db	9 db
Washing Machine, Agitate	25%	3/	10%
Refrigerator	36	17	14
Dryer	24	9	7
White Noise[4]	16	6	3
Standard Band of Frequencies	15	8	7
Octave Band	42	36	3/
Octave Band plus Pure Tone Superimposed	26	15	6
Washing Machine, Agitate Filtered	27	15	10
Total for All Sound Types Averaged[5]	25%	13%	10%

[1]Each comparison is made against the lowest nominal level. Depending upon various factors discussed in the text the lowest nominal level falls between 35 and 50 decibels.

[2]Proportion inconsistent of those respondents capable of being so.

[3]Data not available because these stimuli were not included in the actual tape program.

[4]White noise was not asked at the second lowest nominal level (3); therefore, for comparison purposes the step between nominal level 6 and 9 was used to represent the data in the 3 db column.

[5]In calculating the totals, per cents for those columns with missing data were estimated on the basis of data actually reported at other decibel difference levels.

The total averages for all sound types with the basement included is: 3 db - 28%; 6 db - 19%, 9 db - 12%.

TABLE 21

FLOW DIAGRAM TRACING THE REPETITION OF RESPONDENT BOTHERSOME JUDGMENT INCONSISTENCIES WITH CHANGES IN THE INTENSITY LEVEL OF THE WASHING MACHINE AGITATE CYCLE SOUND[1]
(numbers entered in tables are "individuals" and not percentages)

Bothersome Response At Level 0	Nominal Levels Under Comparison				
	0 - 3	0 - 9	0 - 12	0 - 15	0 - 18
Little	11	3	3	3	2*
Moderate	9	1	1		3*
Extreme	1	1	1	1	
Total	21	5	5	4	
		new inconsistents			
Little		5	2	-	
Moderate		4	2	-	
Extreme		-	-	-	
Sub-total		9	4	0	
Grand Total		14			
			new inconsistents		
Little			1	-	
Moderate			-	-	
Extreme			-	-	
Sub-Total			1	0	
Grand Total			10	no new inconsistent cases	
Grand Total				4	new inconsistents
Little					3
Moderate					2
Extreme					-
Sub-Total					5
Grand Total					10

*These cases were inconsistent at compared levels 0-3 but were consistent at 0-9, 0-12, and 0-15. This situation occurs nowhere else on this table.

[1]This table is based on two responses per individual where present since it incorporates the data for each individual at each of the two speaker locations at which the respondent heard the stimulus (this includes two responses per person in cases where respondents replied to both questions they were asked). There is only one instance of an individual represented twice on the above table (responded inconsistently at both speaker locations) and this is included in the figure of 11 under 0-3 for those with a judgment of little bothered at level 0. This reinforces the notion that it is generally not the same individuals contributing the inconsistencies in the majority of cases.

Table 21 indicates that the errors which occur at different levels are not made by the same individuals in the majority of instances. Those cases present which do remain inconsistent over a large portion of the intensity difference range do indicate that there are some individuals who tend toward overall inconsistency. However, and more importantly, the larger number of inconsistencies which occur as single entities by individuals tend to lend support to the notion that there is a residual inconsistency level with numerous and different respondents entering and leaving. This is probably due either to momentary lapses in the respondents ability to recall the general frame of reference in which the questions are being asked or some mechanical error incurred by the respondent in the process of recording the judgment made.

4.5.3. Respondent Consistency in Judging the Degree of Bother of Identically Repeated Stimuli

This test for consistency is perhaps as critical as any that could conceivably be devised. It demands of the individual the ability to exactly repeat the bothersome judgment of a stimulus whenever it is presented. Seven such stimuli were employed in the tape program and the repetitions were randomly distributed throughout the interview with time separations of as little as two minutes and as long as an hour. Table 22 enumerates the stimuli repeated, the time differences (in terms of the number of responses made during the interval between the original and repetition), and sets forth important information about the stimulus in each case which preceded the sound used for the repetition analysis. The preceding sound is important because it may serve to set a frame of reference, in particular to establish a "contrast" effect. Table 22 also makes judgments regarding the character of the preceding stimuli to determine if the character of one (in terms of degree of bother and the possible frame of reference set) differs substantially from the other. This last point is an estimation based upon the differences between the regression curves of the two "preceding sounds involved (as presented in section 4.1).

The method to be employed here is to first analyze individually the consistency associated with each of the seven repeated stimuli and then merge all seven for a study of overall consistency and consistency by seat location and speaker location.

Table 23 reports the consistency of respondents for each of the seven sounds in three strata. The first stratum consists of those totally consistent; that is they repeated on the second exposure exactly what they stated in the first. The second stratum might be said to be semi-consistent; the variation between the first and second exposure was only one step on the bother scale. The third stratum are highly inconsistent in that they varied between exposures by two or three steps. In addition, strata two and three are broken down into the direction of change from the first to the second exposure; that is, more or less bother on the second exposure than the first.

TABLE 22

INFORMATION RELATING TO THE SEVEN REPEATED ABSOLUTE STIMULI USED FOR THE PURPOSE OF DETERMINING A CRITICAL LEVEL OF RESPONDENT CONSISTENCY

Repeated Stimuli Identification			No. of Questions Separating Original and the Repetition	Identification of Sounds Preceding the Stimuli Being Judged for Consistency[2]	Difference in Intensity Levels of the Preceding Stimuli	Estimated bias introduced by differences in the character of preceding stimuli[3]
Name of Stimulus	Nominal Intensity Level	Locations of Sources Used[1]				
A	3	I	26	a) O+PT, 15, I b) D, 3, I	12	small
R	18	I	6	a) AF, 6, I b) O+PT, 18, II	12	very small
R	15	II	48	a) O+PT, 9, II b) SBF, 6, II	3	large
WN	12	I	7	a) O, O, II b) O+PT, 15, I	15	intermediate
SBF	3	II	80	a) SBF, 18, I b) SBF, 3, II	15	none
O	12	I	11	a) AF, 15, II b) SBF, 9, II	6	very large
O+PT	15	I	19	a) SBF, 3, II b) O+PT, 6, II	3	large

[1] I; includes Living Room Filter, Living Room Unfiltered and Kitchen II; includes Basement and Living Room Filtered.

[2] The key to the abbreviated presentation on the table is as follows: Lines demarked by an "a)" refer to the identification of the stimulus which preceded the first exposure of the sound being investigated for consistency while lines demarked by a "b)" refer to the identification of the stimulus which preceded the repetition of the sound being investigated for consistency.

The initial letters identify the stimulus type.

The arabic number following the abbreviated name identifies the nominal intensity level.

The roman number which occurs last on each line is the location of the speaker source used and the key is the same as under footnote 1.

[3] The general notion of the estimating method is offered in the text.

TABLE 23

RESPONDENT CONSISTENCY IN JUDGING THE DEGREE OF BOTHER IN IDENTICALLY REPEATED STIMULI FOR EACH OF SEVEN SOUNDS INDIVIDUALLY

a) Refrigerator, Nominal Level of 15

Bother Response of First Exposure[1]	Bother Response of Second Exposure[2]			
	Not	Little	Moderate	Extreme
Not (37.5%)	6.4%			6.9%[4]
Little (40.4%)		10.1%	20.2%[3]	
Moderate (56.3%)		19.2%[5]	21.3%	
Extreme (60.5%)	3.7%[6]			12.2%

Total consistency:	50%[7]
Consistency ± one degree of bother:	89%[8]
Gross inconsistency:	11%[9]

[1]The percentages in parentheses represents the proportion of the total responses for each bothersome level at the first exposure who remained exactly consistent after the repetition.

[2]Percentages represent proportions of the entire sample reporting.

[3]Percentage represents proportion of the entire sample which was semi-consistent (a change in one degree of bother) and became more bothered in second exposure.

[4]Percentage represents proportion of the entire sample which was highly inconsistent and became two or more degrees more bothered in the second exposure.

[5]Identical to footnote 3 but respondents became less bothered.

[6]Identical to footnote 4 but respondents became less bothered.

[7]Figure obtained by addition of all percentages on the diagonal.

[8]Figure obtained by addition of all percentages on the diagonal plus those percentages one step above or below the diagonal.

[9]Figure obtained by addition of all percentages above or below the diagonal by more than one step.

Note: The above footnotes will not be listed for the remaining six repeated stimuli but do apply in every instance.

For the purpose of initial clarity shading is provided; it is true for the other six tables but omitted.

TABLE 23 (Continued)

b) White Noise, Nominal Level of 12

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (71.9%)	12.4%			3.2%
Little (52.4%)		5.9%	10.8%	
Moderate (60.9%)		73.0%	15.1%	
Extreme (42.7%)	2.1%			37.6%

Total Consistency: 71%
Consistency ± one degree of bother: 95%
Gross Inconsistency: 5%

c) Washing Machine Agitate Cycle, Nominal Level of 3

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (81.0%)	67.7%			5.2%
Little (40.0%)		4.2%	12.1%	
Moderate (40.0%)		7.9%	2.1%	
Extreme (0.0%)	0.5%			0.0%

Total Consistency: 74%
Consistency ± one degree of bother: 94%
Gross Inconsistency: 6%

TABLE 23 (Continued)

d) Refrigerator, Nominal Level 18

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (33.3%)	1.6%			3.7%
Little (33.3%)		2.6%	8.9%	
Moderate (37.0%)		24.1%	8.9%	
Extreme (71.1%)	5.2%			45.0%

Total Consistency: 58%
Consistency ± one degree of bother: 91%
Gross Inconsistency: 9%

e) Standard Band of Frequencies; Nominal Level 3

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (53.5%)	35.7%			8.1%
Little (50.0%)		8.8%	29.9%	
Moderately (45.8%)		6.4%	6.4%	
Extreme (66.7%)	1.8%			1.2%

Total Consistency: 53%
Consistency ± one degree of bother: 90%
Gross Inconsistency: 10%

TABLE 23 (concluded)

f) O, Nominal Level of 12

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (68.8%)	11.8%			2.1%
Little (55.3%)		13.9%	12.4%	
Moderate (50.0%)		24.6%	17.7%	
Extreme (50.0%)	6.4%			11.2%

Total Consistency: 55%
Consistency ± one degree of bother: 92%
Gross Inconsistency: 8%

g) O+PT, Nominal Level of 15

Bother Response of First Exposure	Bother Response of Second Exposure			
	Not	Little	Moderate	Extreme
Not (62.5%)	8.0%			3.7%
Little (50.0%)		10.1%	17.0%	
Moderate (31.6%)		14.9%	19.7%	
Extreme (21.5%)	3.7%			23.4%

Total Consistency: 61%
Consistency ± one degree of bother: 93%
Gross Inconsistency: 7%

Before beginning the discussion of Table 23 it might be well to summarize the high points and rank order the seven repeated stimuli by proportion of consistency for convenience in referring to the data; this is performed in Table 24.

TABLE 24

SUMMARY OF CONSISTENCY FOR ABSOLUTE STIMULI BY STIMULUS TYPE

Rank Order; Proportion of Fully Consistent Responses	Identification of Stimulus		Proportion of Consistent Responses	
	Sound Type	Nominal Level	Totally	± one degree of bother
1	A	3	74%	94%
2	WN	12	71	95
3	O+PT	15	61	93
4	R	18	58	91
5	O	12	55	92
6	SBF	3	53	90
7	R	15	50	79

First there must be a general point made about the overall levels of consistency in this respect. This data fits well with regression correlations and curves drawn from other sections where data were reported. Consistency is fairly high, Table 25 shows an overall level of total consistency at about 61%. However, the 39% inconsistency rate is sufficient to cause serious impact to correlation measures. It is also sufficient to cause data distributions which scatter a bit.

At this point the focus of attention is returned to the effect of the sounds which precede the stimuli upon which the consistency analysis in the repetition format is based. Despite what might appear to be a somewhat elaborate display of data this is again one of those aspects of the research program which is so intricate in its nature that our information for basing conclusive judgments is thin. So many parameters are running uncontrolled that it is difficult to state conclusively that one preceding sound is more bothersome than another. This may not be true for the entire sample even if it could be demonstrated that one preceding sound is more bothersome than another for the majority of the population. In addition this last point is not clearly demonstrable because the preceding stimulus was not one which an attempt was even made to control. Therefore, in most instances the two preceding sounds of the stimuli being investigated for consistency differ with respect to the type of sound they are, nominal level and speaker location. Even if all of these were controlled, the seat position of the respondent could substantially effect the actual level at which different respondents received the same preceding stimuli.

Despite all these shortcomings and the reservations that must, therefore, be made in any interpretation of this subject from our data, there is a surprising consistency in the results suggesting the operation of a contrast effect.

TABLE 25

RESPONDENT CONSISTENCY IN JUDGING THE DEGREE OF BOTHER IN IDENTICALLY REPEATED STIMULI FOR ALL SEVEN SOUNDS COMBINED
(Numbers in parentheses show the proportion of total response for each bother level that is consistent for each exposure individually)

Bother Response of First Exposure	Bother Response of Second Exposure				
	Not (79.7%)	Little (34.3%)	Moderate (48.5%)	Extreme (76.1%)	Total
Not (67.8%)	21.9%	7.5%	2.5%	0.4%	32.3%
Little (47.2%)	3.6	7.8	4.0	1.2	16.6
Moderate (50.4%)	1.4	6.2	12.1	4.3	24.0
Extreme (69.3%)	.6	1.4	6.4	18.7	27.0
Total	27.5%	22.9%	25.0%	24.6%	100.0%

Total Consistency:	60.5%
Consistency ± one Degree of Bother:	92.0%
Proportion Expressing More Bother with the Second Exposure:	19.9%
Proportion Expressing Less Bother with the Second Exposure:	19.6%

TABLE 26

PROPORTION OF CONSISTENT RESPONSES (TOTAL) BY LOCATION OF THE SOUND SOURCE

Location of Sound Source	Proportion of Consistent Responses
Living Room Unfiltered	65.6%
Living Room Filtered	54.9%
Kitchen	62.9%
Basement	60.2%

The first factor worthy of notice is that of the rank ordering of the seven consistency stimuli (Table 24); numbers 4, 5 and 6 show considerable differences in terms of respondents shifting more noticeably towards increased bother or decreased bother while those sounds which are higher up on the rank order tend to show a more equal distribution. Only refrigerator at level 15, the least consistent, tends to contravert this relationship. Table 27 demonstrates this fact more clearly.

TABLE 27

TRENDS IN THE DIRECTION OF INCONSISTENT BOTHERSOME RESPONSES BY SOUND TYPE

Rank Order of Stimulus	Identification of Stimulus		Change in the degree of bother that occurs when the consistency stimulus is repeated		
	Name	Level	more bother	less bother	difference
1	A	3	17.3%	8.4%	8.9%
2	WN	12	14.0	15.1	1.1
3	O+PT	15	20.7	18.6	2.1
4	R	18	12.7	29.3	16.6
5	O	12	15.0	31.0	16.0
6	SBF	3	38.0	9.2	28.8
7	R	15	27.1	22.9	4.2

Proceeding under the assumption that the more definite directions the inconsistency movements take in those sounds generally lower in the rank orderings are the results of bias caused by the impact of the preceding sounds it remains to be shown that such preceding sounds logically would be expected to push the bother responses in the direction in which they went for each instance individually.

The next logical step is to examine the theoretical proposition and the actual fact for each of the three situations; R at Level 18, O at Level 12 and SBF at Level 3 which tend to lend some support to the inferences being driven at.

Case I: R at Level 18.

From Table 27 it can be observed that the majority of inconsistent respondents have become less bothered at the repetition of the sound. The preceding stimuli for the repeated consistency sound must have been generally more bothersome in nature. Table 22 reveals the preceding sounds to be AF at Level 6 coming from the Living Room and Kitchen and O+PT at Level 18 in the Basement and

Living Room Filtered. The regression analysis of the earlier section revealed that the basic character of these two preceding sounds was similar. However, despite the fact that AF comes from speaker locations with generally higher intensity the nominal level difference is so substantial (12 decibels) that we would expect in most instances that respondents would be experiencing O+PT levels considerably higher (both in decibels and phons). Hence, the facts are consistent with the preconceived theory in this case.

Case II: O at Level 12

The reader ought to be able to follow the approach and, therefore, less detail will be required in this case and the next. Like the previous case respondents are less bothered with the repeated sound than the original. In this case SBF is at a lower intensity level (6 decibels) and at the same speaker locations as AF. Nevertheless, the key here is that AF is generally a far less bothersome sound than SBF.

Case III: SBF at Level 3

This case makes a fine demonstration quite by chance as it was not planned. Both preceding sounds are the same as the stimulus being used for consistency and differences between them do not enter the situation. The first preceding sound is at a much higher level than the second; not only is it higher nominally but the presence of it at upstairs speaker locations will make the differences between the actual levels with the second preceding sound even larger. Needless to say the repetition here is found more bothersome.

The data presented is far from conclusive to be sure. R at level 15, the least consistent sound, does not contribute to the hypothesis. The three sounds which are most consistent do not have what would appear to be equally balanced preceding sounds that would contribute to their consistency being higher either.

The one condition which has been neglected in this analysis is that certain sounds might inherently have more consistency associated with them than others. In previous experiments by others (and other sections of our own experiment) it has clearly been substantiated that consistent judgments by respondents are more often forthcoming for white noise (WN) than most other types of sounds. To what extent this is generally true of the other sounds being considered here is not known but undoubtedly it must play some role.

It is sufficient to point out, however, that the preceding sound clearly offers some degree of bias to judgments. The three cases pointed to (out of a possible total of seven) that demonstrate this point vividly are just too improbable to fairly attribute to chance. Future experiments along these general lines will certainly have to exert caution in programming of sounds to minimize the bias caused by the preceding sounds before fair judgments can be made. Certainly it appears to be mostly true that those sounds which demonstrate wider deviations from the equal distribution of inconsistencies (between more and less bothered) tend to be responded to with from 10 to 20% less consistency than those cases where the aforementioned equal distribution exists. For whatever reason it is clear that some force is accounting for the movement of the inconsistent primarily in one direction and it is this force that is lowering the overall consistency rate for cases where it occurs. Since little else seems to be related to consistency it appears to this research staff at least that this matter of the preceding sounds plays a role of some sort in the determination of overall respondent consistency in cases where individuals are asked to judge the degree of bother from a single sound.

One final methodological step was pursued to attempt to put the type of consistency analysis under discussion here into the more rigorous format of the standard least squares regression. Table 28 reports the results:

TABLE 28

RESPONDENT CONSISTENCY IN JUDGING THE DEGREE OF BOTHER IN IDENTICALLY REPEATED STIMULI FOR EACH OF THE SEVEN SOUND TYPES PRESENTED AS MEASURED BY LINEAR CORRELATION METHODS.[1]

Stimulus	Sound Level	R
A	3	.45
R	18	.47
R	15	.52
WN	12	.79
SBF	3	.52
O	12	.65
O+PT	15	.70

[1]The introduction of speaker sound source as a third variable did not explain any additional variance.

The rather large difference encountered between the distribution of inconsistencies in Table 23 and that presented above is a perfect demonstration of how these methodological tools operate differently. The regression format is far more critical. For example; A at level 3 appears to be 74% consistent in Tabular format; however, this is a result of a clumping of all (or almost all) respondents at the not bother level. Since the regression operates over the entire range, and the remainder of the range is not as consistent, the percent of variance explained in the regression format is relatively low.

In general it may be inferred that those sounds which do well in both formats of analysis are of the nature such that respondents are more capable of judging them consistently. This is of course an inference only at this point and considerably more experimentation is required to strongly support this case.

The two sound types which fulfill the aforementioned requirements most adequately, it can be noted, are White Noise and O+PT.

4.5.4 Concluding Remarks

This chapter has demonstrated the variety of types of consistency which may be expected of a respondent from tests which are relatively simple to those which are quite difficult. It has been shown that a relatively high degree of consistency is possible if the tests are simple and conversely that where a very difficult judgment is required consistency will decrease. However, it can be clearly borne out that the pessimists who deny the possibility of psychoacoustical experiments with randomly sampled populations are incorrect in their judgment. It is further borne out that such bothersome judgments can be made reasonably well in non-laboratory situations that approximate the typical household environment. Throughout this section mention was made of numerous laboratory pre-experimentation programs which would aid markedly the final outcome of any future experiments conducted in a true-life setting in terms of respondent consistency. This is both in terms of improvements in the methodology of conducting the interview and in actually improving the stimuli selected for use. At this point, however, it is sufficient to say that to the extent that data in previous sections reveal trends they are no doubt accurate and not to be ignored because of the notion that our respondents were inconsistent. In almost all cases relationships were not strongly expressed and here this is in agreement with what would be expected with the level of consistency we are generally working with. Once again, trends that appear, such as they are, probably can be confirmed insofar as it is necessary to demonstrate that the respondent can consistently repeat his judgment.

APPENDIX A

INTERVIEW QUESTIONNAIRES AND DESCRIPTION OF STIMULI

Institute of Science and Technology

Appl. Annoyance Study - Int. II

Yes

Date ____________________ Maybe

Yes - qualified

Address __________________________ No

Date House Interview ______________

INTRODUCTION:

I'm an interviewer for the University of Michigan Institute of Science and Technology. A few days ago we sent you a letter asking if you would cooperate in a survey concerning home appliances manufactured in Michigan.

Helping the University of Michigan with this study is also helping to create a brighter economic future for your home state. Will you allow me to come into your home and ask you a few questions?

1. Generally speaking, are you happy with the appliances you use, that is, things like refrigerators, washing machines, dryers, vacuum cleaners, garbage disposals and so forth, or do you have any complaints about the way appliances work?

☐ COMPLAINT ☐ NO COMPLAINT

(IF A COMPLAINT) 1a. Which appliance? ____________________

1b. What is your complaint? ____________________

____

1c. Any others? ____________________

(IF YES) Which appliance? What is your complaii

____

____

(ASK IF SOUND WAS NOT MENTIONED IN Q1, OTHERWISE SKIP TO Q3) Do you have any complaints about the sounds any appliances make?

☐ YES ☐ NO

2a. (IF YES) Which appliances? ____________________

____

3. One of the things we are after in this study is a better idea of the kinds of appliance problems that consumers would like to have studied.

3a. Do you think a manufacturer should spend his time and money making his appliances quieter or improving his appliances in some other way?

☐ QUIETER ☐ OTHER

3b. (ASK IF OTHER) Are there any special problems with appliances that you would like to see worked on? ____________________

____

3c. Why do you feel this way? ____________________

____

4. Now we want to ask especially about sounds of appliances. Suppose you were buying a new home appliance. You find a model that is as good as the others on the market, and also is substantially quieter.

4a. Now we want to know, would you be willing to spend any additional money for this quieter model? ☐ YES ☐ NO

4b. (ASK IF YES) Were you thinking about any special appliances when you said yes? Which ones? ____________________

4c. (ASK IF YES) As you probably know, new appliances such as washing machines, dryers and so on frequently cost upwards of 300 dollars - some more, of course, and some less.
But thinking about $300 as the price for an appliance, we would like to know how much more you would pay for this quieter model:

☐ up to $10 ☐ $10-$25 ☐ $26-$50 ☐ over $50

4d. (ASK IF YES AND 3c NOT ASKED) Why would you be willing to pay more for quietness in an (appliance)? ____________________

____

4e. (ASK IF NO) If the quieter model cost about the same as the others, would you prefer this quieter machine? ☐ YES ☐ NO

(ASK IF YES AND 3c NOT ASKED) Why would you prefer a quieter model?

____

5. Do you know of anything special about your hearing; that is, is it better than that of most people, is it about average, or do you have any hearing defect?

☐ SPECIAL ☐ NOT SPECIAL (SKIP TO Q6)

5a. (IF SPECIAL) What is special about it? ____________________

____

6. Will you tell me which of the following appliances you have in your home:

Air Conditioner	(Yes)	(No)	Garbage Disposal	(Yes)	(No)
Automatic Washer	(Yes)	(No)	Refrigerator	(Yes)	(No)
Wringer Washer	(Yes)	(No)	Deep Freeze	(Yes)	(No)
Clothes Dryer	(Yes)	(No)	Vacuum Cleaner	(Yes)	(No)
Dishwasher	(Yes)	(No)			

7. (ASK IF OWNS WASHER) Where in your home is your washing machine located?
   (Basement) (Kitchen) (Utility room) (Other __________)

PERSONAL DATA

P1. Are you married? ☐ Married ☐ Single ☐ Divorced ☐ Widowed

P2. Do you have any children? (IF YES) How many live here with you?
   (0) (1) (2) (3 or more)

P2a. (IF CHILDREN) How old is your youngest child? __________

P3. How many grades of school have you finished? __________

P3a. (IF MORE THAN 8) Have you had any other schooling?
   WHAT? __________

INTERVIEWER: ESTIMATE (BY OBSERVATION WHERE REQUIRED)

1. Age of respondent (18-24) (25-34) (35-44) (45-54) (55-64) (65 up)
2. Race of respondent (white) (non-white)
3. Judging by home, income level:
   (Under $5000) ($5000-7499) ($7500-9,999) ($10,000 or over)
4. Type of house (brick) (wood-frame) (other frame)
5. General location (urban) (suburban) (rural)
6. Living room: (live room) (dead room)
   (Judge "deadness" by: drapes vs curtains, overstuffed furniture, rugs with pads vs no pads or bare floors, wall board vs plaster walls)

---

1. Thank R's for their participation.
2. Explain purpose of house sound tests.
3. Explain procedure at house.
4. Invite R's to attend
   —transportation, baby sitters, times.

1. Would you be willing to help us by being our guest in our research home?

(YES) (NO) end here

2. (If Yes) When would it be most convenient for you? (If possible, set up an exact appt. date and time.)

Time of Day__________

DAY OF WEEK__________ (Be Sure to Check Date With Pocket Calendar)

DATE__________

2a. If appt. is set up: May I have your name and telephone number for our records?

Name______________________ Telephone: NO__________

2b. If no appt. is set up: May I take your name and telephone number so that we can call you at a later time to set up an appointment? (end here)

Name______________________ Telephone: NO__________

3. Will you require a baby sitter or transportation?

CHECK IF NEEDED

(baby sitter)

(transportation)

GIVE R AN APPT CARD, EXPLAIN HOW TO REACH THE HOUSE, Fill in date and time.

Appointment Card:

The University of Michigan
Institute of Science and Technology

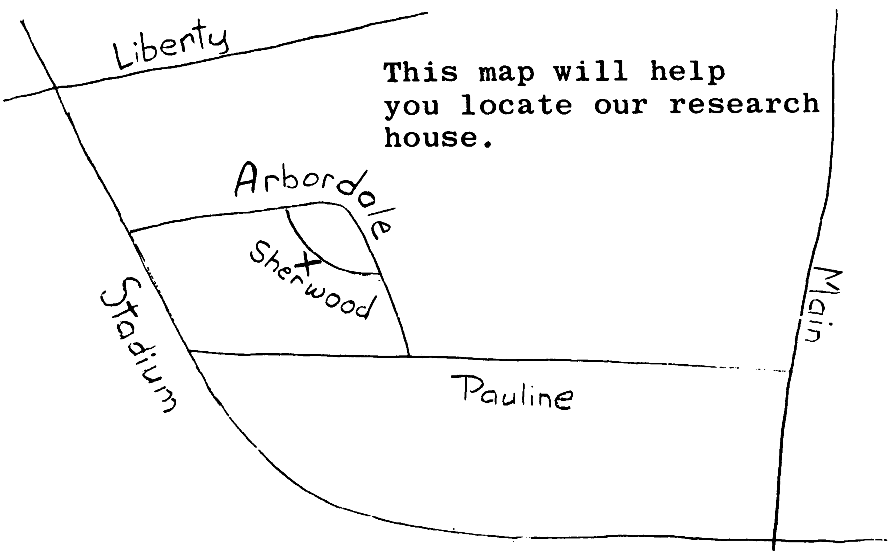
(Mr.)(Miss)(Mrs.) ______________________________

Thank you for agreeing to be our guest at our research home located at 926 Sherwood, Ann Arbor.

On (day)__________ (date)__________
At (time)__________________________

We are looking forward to seeing you. If for any reason you need to change your appointment, please let us know by calling us at NO 3-1511, Extension 315 North Campus, or NO 3-7116.

FRONT



BACK

IST APPLIANCE SOUND STUDY

1. Do you have a washing machine in your home? ____________

2. Does your washing machine load from the top or from the front?

   ____

3. Does your washing machine have a wringer or does it spin dry?

   ____

4. Does your home have a basement? __________

5. Is your basement finished? That is, is it dry walled, paneled, or plastered? Do you have acoustical tile on the ceiling or anything of this sort? ________________________________

   ____

6. Do you feel that sound research will be helpful to you as a consumer? How could this be? __________________________

   ____

   ____

7. In order to better understand your answers, we would like to have a general idea where the major appliances are located in your home. Therefore, would you please draw us a rough sketch of the floor plan of your house, indicating where the washing machine, the dryer, refrigerator, stove, and dishwasher (if you have them) are located? Use the next page for your sketch.

FLOOR PLAN

Respondent provided sketch of own home floor plan as requested in item 7 on previous page.

Audiometry form:

NAME ______________________ CODE NO. ____

ADDRESS ______________________

DATE ____ AGE ____ SEX ____

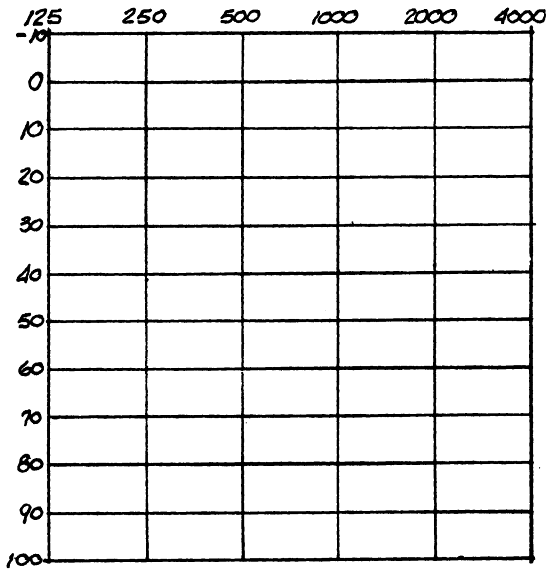
OCCUPATION ______________________

PREVIOUS NOISE EXPOSURE FACTORY GUNFIRE ETC.

____

EAR PROTECTION YES ☐ NO ☐

LEFT X RIGHT O



IST APPLIANCE SOUND STUDY: PART I

In this test, we will be asking you whether or not you find our sounds bothersome. There are many meanings which may be attached to this word, bothersome. Words such as disagreeable, annoying, disturbing, and irritating are similar in meaning. A sound may be bothersome to a person because of its racket, pitch, tone, loudness, softness, sharpness, etc. It might be too flat, too dull, too piercing, too sharp, or perhaps you may not like the sound pattern because of the beat, thump, hiss, click, whir, blast, slosh, etc. The character of the sound may also make you feel on edge, nervous, jumpy, or perhaps it would bother other members of your family.

To illustrate the method of scoring we will work a few examples before starting the actual test.

Practice Questions:

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
A	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
B	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
C	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
1	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
2	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
3	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
4	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ____________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
5	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
6	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ____________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
7	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
8	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
9	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
10	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
11	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
12	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
13	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
14	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
15	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
16	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
17	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
18	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
19	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
20	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
21	☐ No	☐ No	☐ Only a little bothersome
	☐ Yes	☐ Yes	☐ Moderately bothersome
			☐ Extremely bothersome (makes me uncomfortable)

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
22	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
23	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
24	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
25	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
26	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
27	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
28	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
29	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
30	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
31	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
32	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
33	☐ No	☐ No	☐ Only a little bothersome
	☐ Yes	☐ Yes	☐ Moderately bothersome
			☐ Extremely bothersome (makes me uncomfortable)

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

34	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
35	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
36	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
37	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
38	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
39	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
40	☐ No	☐ No	☐ Only a little bothersome
	☐ Yes	☐ Yes	☐ Moderately bothersome
			☐ Extremely bothersome (makes me uncomfortable)
41	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
42	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
43	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
44	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
45	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
46	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
47	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
48	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

a) If you are bothered by this sound, would you please tell me why you are bothered? ______________________________

____

b) If this sound were made by one of your appliances—an appliance that did something useful for you—would you find this: ☐ acceptable ☐ not acceptable

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
49	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
50	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
51	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
52	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
53	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
54	☐ No	☐ No	☐ Only a little bothersome
	☐ Yes	☐ Yes	☐ Moderately bothersome
			☐ Extremely bothersome (makes me uncomfortable)
55	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
56	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
57	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
58	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
59	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely
60	☐ No	☐ No	☐ A little
	☐ Yes	☐ Yes	☐ Moderately
			☐ Extremely

IST APPLIANCE SOUND STUDY: PART II

Practice Questions:

X.	Which sound bothers you more?	X.	A ☐	B ☐
		Y.	A ☐	B ☐

In Number Y, would you please tell us why one sound bothers you more than the other ____________________

____

Z.	Which sound bothers you more?	Z.	A ☐	B ☐

= =

1.	Which sound bothers you more?	1.	A ☐	B ☐
		2.	A ☐	B ☐
		3.	A ☐	B ☐
		4.	A ☐	B ☐
		5.	A ☐	B ☐
		6.	A ☐	B ☐
		7.	A ☐	B ☐
		8.	A ☐	B ☐
		9.	A ☐	B ☐
		10.	A ☐	B ☐

In Number 10, would you please tell us why one sound bothers you more than the other ____________________

____

11. Which sound bothers you more?

11.	A ☐	B ☐
12.	A ☐	B ☐
13.	A ☐	B ☐
14.	A ☐	B ☐
15.	A ☐	B ☐
16.	A ☐	B ☐
17.	A ☐	B ☐
18.	A ☐	B ☐
19.	A ☐	B ☐
20.	A ☐	B ☐

In Number 20, would you please tell us why one sound bothers you more than the other ______________________________

____

21. Which sound bothers you more?

21.	A ☐	B ☐
22.	A ☐	B ☐
23.	A ☐	B ☐
24.	A ☐	B ☐
25.	A ☐	B ☐
26.	A ☐	B ☐
27.	A ☐	B ☐
28.	A ☐	B ☐
29.	A ☐	B ☐
30.	A ☐	B ☐
31.	A ☐	B ☐
32.	A ☐	B ☐
33.	A ☐	B ☐
34.	A ☐	B ☐
35.	A ☐	B ☐

36. Which sound bothers you more?

36.	A ☐	B ☐
37.	A ☐	B ☐

In Number 37, would you please tell us why one sound bothers you more than the other ______________________

____

38. Which sound bothers you more?

38.	A ☐	B ☐
39.	A ☐	B ☐
40.	A ☐	B ☐

IST APPLIANCE SOUND STUDY: PART III

1. Which sound bothers you more?

1.	A ☐	B ☐
2.	A ☐	B ☐
3.	A ☐	B ☐
4.	A ☐	B ☐
5.	A ☐	B ☐
6.	A ☐	B ☐
7.	A ☐	B ☐
8.	A ☐	B ☐

In Number 8, would you please tell us why one sound bothers you more than the other ______________________

____

9. Which sound bothers you more?

9.	A ☐	B ☐
10.	A ☐	B ☐
11.	A ☐	B ☐
12.	A ☐	B ☐
13.	A ☐	B ☐
14.	A ☐	B ☐
15.	A ☐	B ☐
16.	A ☐	B ☐
17.	A ☐	B ☐
18.	A ☐	B ☐
19.	A ☐	B ☐
20.	A ☐	B ☐

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
21	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
22	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
23	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
24	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
25	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
26	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
27	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
28	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
29	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
30	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
31	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
32	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
33	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
34	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
35	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
36	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
37	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
38	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
39	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
40	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
41	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
42	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
43	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
44	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

Sound Number	Did you hear the sound?	Was the sound bothersome?	If so, how bothersome?
45	☐ No ☐ Yes	☐ No ☐ Yes	☐ Only a little bothersome ☐ Moderately bothersome ☐ Extremely bothersome (makes me uncomfortable)
46	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
47	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
48	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
49	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely
50	☐ No ☐ Yes	☐ No ☐ Yes	☐ A little ☐ Moderately ☐ Extremely

TABLE 29

ABBREVIATIONS DESIGNATING STIMULUS TYPE

ABSOLUTE STIMULI

A	Automatic washing machine in agitation cycle
D	Electric clothes dryer
R	Refrigerator
AF	Automatic washing machine, same as A above except filtered through 1/3 octave band filter centered at 125 cps.
O	Band of random noise centered at 125 cps; same filter as for AF above.
O + PT	Band of random noise; same as O above with superimposed pure tone at 120 cps.
WN	Broadband white noise.
SBF	Standard band of frequencies; a band of noise from 670 cps to 1000 cps.

ADDITIONAL ABA STIMULI

125	Pure tone at 125 cps.
256	Pure tone at 256 cps.
4000	Pure tone at 4000 cps.

TABLE 30

IDENTIFICATION OF STIMULI

IST APPLIANCE SOUND STUDY: PART I

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
A *B C	Three sounds not included in Table 29 which were added to tape to train respondents. Responses to these stimuli were not analyzed.		
1	1	0	A
2	1	-18	R
3	1	-18	O
*4	2	0	D
5	1	-15	R
*6	2	- 6	A
7	2	- 3	AF
8	2	-18	O+PT
9	1	- 6	O
*10	1	0	SBF
11	2	-15	SBF
12	1	- 3	O+PT
13	2	-15	A
14	1	-12	D
15	2	-18	A
16	2	- 9	O+PT
*17	2	- 3	R
18	1	0	D
19	2	- 9	SBF
20	1	- 6	O
*21	1	- 9	A
22	1	-18	A
23	2	0	A
24	2	-18	O
25	1	- 6	WN
*26	1	- 6	D
27	2	- 9	R
28	1	-12	AF
29	1	0	R
30	2	-12	O+PT
31	1	- 3	O+PT
*32	1	- 6	WN
*33	1	- 3	AF

TABLE 30 (continued)

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
34	2	0	O+PT
35	1	0	R
36	2	-18	D
37	2	-15	R
38	1	-15	D
39	2	-15	A
40	1	- 9	SBF
41	2	- 3	O
42	2	-12	WN
*43	1	-12	O+PT
44	2	-12	SBF
45	2	- 3	R
46	2	-18	R
47	1	- 9	WN
*48	2	- 9	A
49	1	- 9	AF
50	2	0	WN
51	1	-15	O+PT
52	2	- 6	O+PT
53	1	- 3	A
54	1	-15	AF
55	2	-12	O
56	1	-18	D
57	2	- 6	WN
58	1	-15	O
59	1	-15	SBF
60	1	- 9	O+PT

IST APPLIANCE SOUND STUDY: PART II

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
X Training	1	- 6	WN - A - WN
* Y Stimuli Data	1	- 3	AF - A - AF
Z Not Analyzed	2	- 9	120 - 4000 - 120
1	2	- 3	120 - 256 - 120
2	1	- 6	R - SBF - R
3	2	- 3	A - WN - A
4	1	- 9	WN - SBF - WN

TABLE 30 (Continued)

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
5	1	- 6	A - WN - A
6	1	- 9	SBF - WN - SBF
7	2	0	O - SBF - O
8	2	- 3	WN - A - WN
9	2	- 9	O - O+PT - O
*10	2	0	O+PT - SBF - O+PT
11	1	- 3	D - O - D
12	1	- 3	R - AF - R
13	2	- 6	O - WN - O
14	2	0	D - AF - D
15	1	0	O+PT - O - O+PT
16	1	- 3	120 - 4000 - 120
17	1	- 3	WN - R - WN
18	1	- 6	SBF - R - SBF
19	1	- 6	O - R - O
*20	2	0	D - O - D
21	2	0	SBF - O+PT - SBF
22	2	- 3	R - O+PT - R
23	1	- 9	AF - WN - AF
24	1	- 6	SBF - O+PT - SBF
25	2	- 9	D - WN - D
26	2	- 3	A - SBF - A
27	1	- 9	AF - A - AF
28	1	- 3	A - R - A
29	2	0	SBF - O - SBF
30	2	- 6	D - SBF - D
31	1	- 9	A - AF - A
32	1	- 3	O - D - O
33	1	- 9	120 - 4000 - 120
34	1	- 3	A - AF - A
35	2	0	O - D - O
36	1	- 3	R - WN - R
*37	1	- 3	AF - R - AF
38	2	- 6	WN - O - WN
39	1	- 3	R - A - R
40	1	0	AF - O - AF

IST APPLIANCE SOUND STUDY: PART III

1	2	- 9	O+PT - O - O+PT
2	2	- 3	SBF - A - SBF
3	2	- 6	D - A - D
4	1	0	O - O+PT - O

TABLE 30 (Continued)

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
5	2	- 9	4000 - 120 - 4000
6	1	0	0 - AF - 0
7	2	- 6	A - D - A
*8	1	- 9	WN - AF - WN
9	2	- 9	WN - D - WN
10	2	- 3	0+PT - R - 0+PT
11	1	- 3	4000 - 120 - 4000
12	2	- 3	256 - 120 - 256
13	1	- 6	0+PT - SBF - 0+PT
14	2	- 6	SBF - D - SBF
15	1	- 6	R - 0 - R
16	2	0	AF - D - AF
17	1	- 9	4000 - 120 - 4000
18	1	- 6	WN - A - WN
19	1	- 3	AF - A - AF
20	2	- 9	120 - 4000 - 120
21	1	-12	0
22	2	-15	D
23	2	-15	SBF
24	1	- 3	0
25	1	-18	AF
26	1	-15	A
27	2	- 6	D
28	2	0	SBF
29	2	-15	AF
30	2	- 3	A
31	1	- 6	SBF
32	1	-18	SBF
33	2	- 3	SBF
34	2	0	AF
35	2	- 9	WN
36	1	-12	WN
37	2	- 3	0+PT
38	1	-12	R
39	2	-15	0
40	2	0	R
41	2	-12	AF
42	1	-18	0+PT
43	1	-18	WN
44	1	-12	SBF
45	2	0	0

TABLE 30 (concluded)

Sound Number (4)	Track (1)	Relative Level db (2)	Stimulus (3)
46	2	- 6	AF
47	1	- 3	WN
48	1	- 9	D
49	1	- 6	A
50	1	- 3	R

(1) Two channel tape recorder utilized. Basement-located loudspeaker always operated from track 2. Kitchen located loudspeaker always operated from track 1. Living room located loudspeaker operated from track 1 or 2 depending on particular test mode.

(2) Relative rms electrical level, called nominal level in body of report, expressed in decibels. Zero represents most intense level.

(3) See Table 29.

(4) Corresponds to numbering on preceding questionnaires.

* Additional comments were solicited from respondents on these stimuli.

APPENDIX B

BOTHER AS A FUNCTION OF INTENSITY

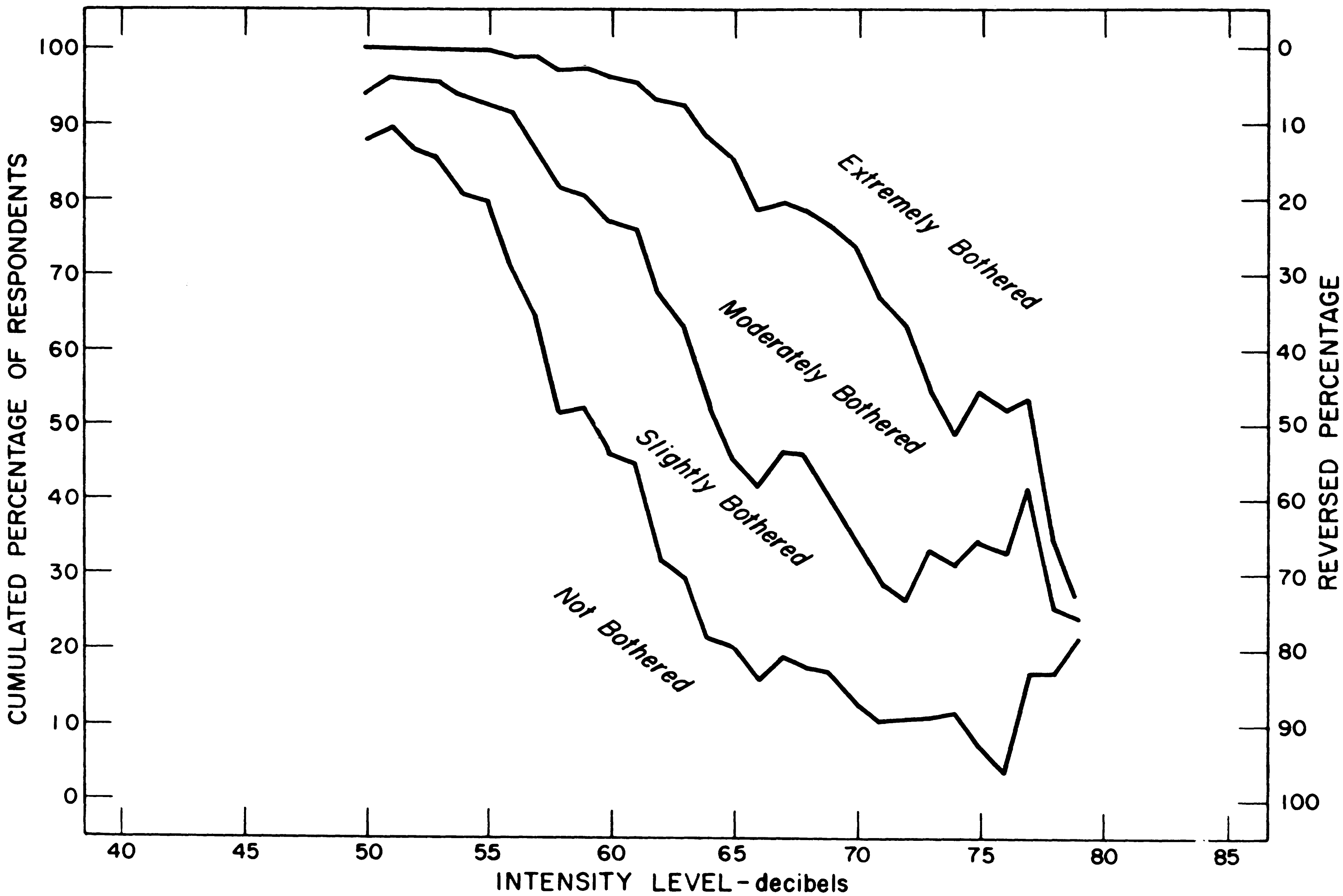


Figure 13. Dryer; all source locations merged.

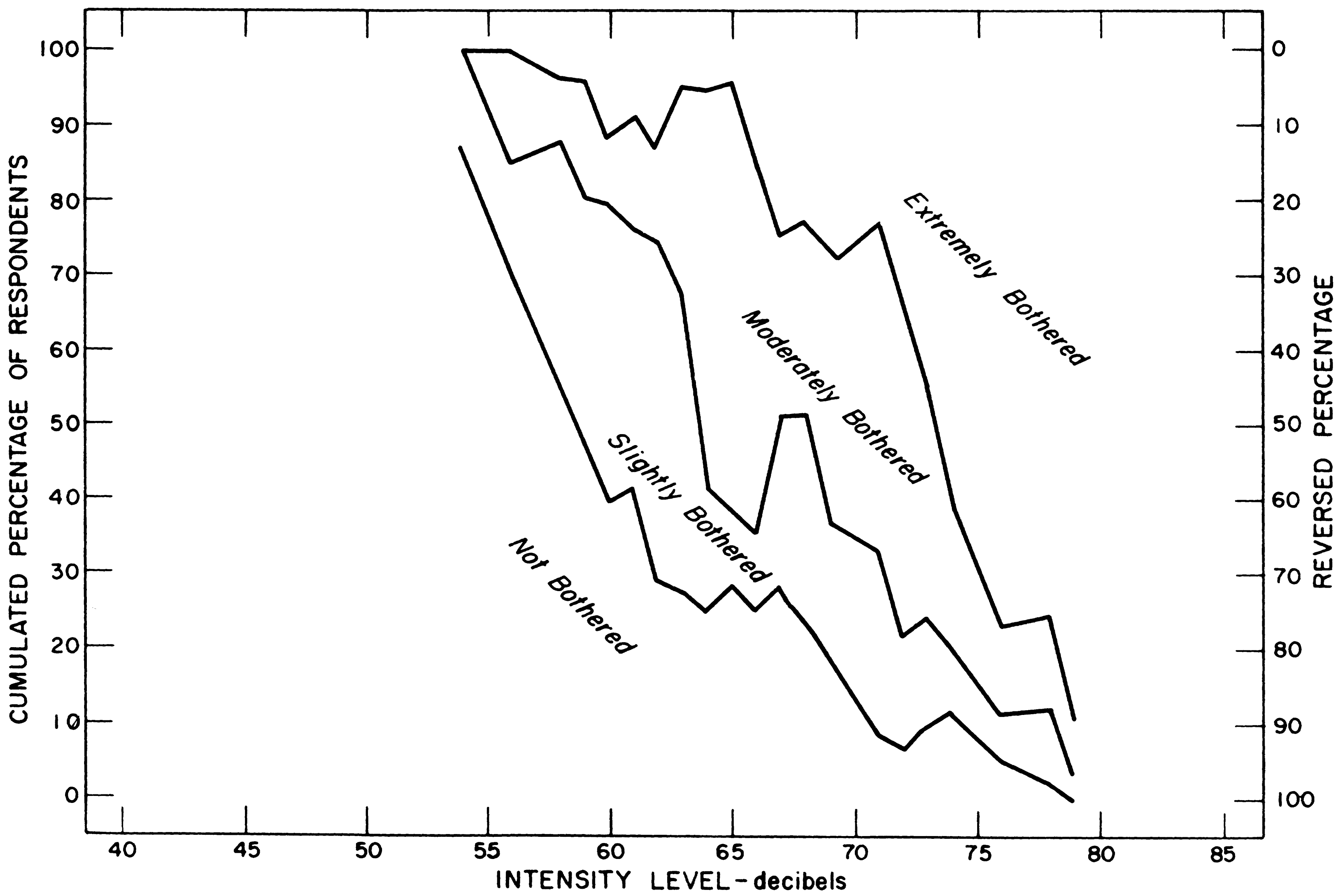


Figure 14. Dryer; source in living room, unfiltered.

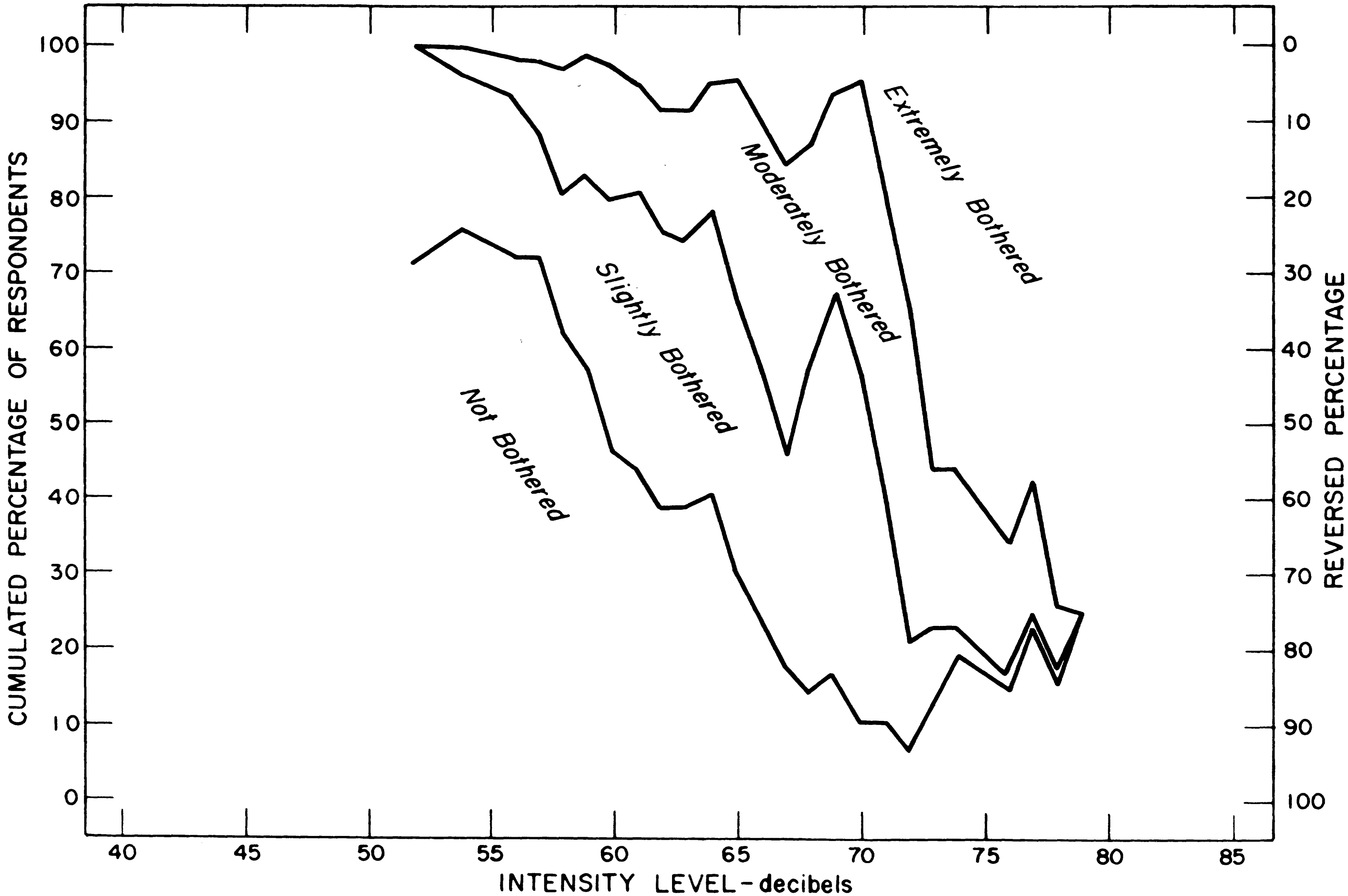


Figure 15. Dryer; source in kitchen.

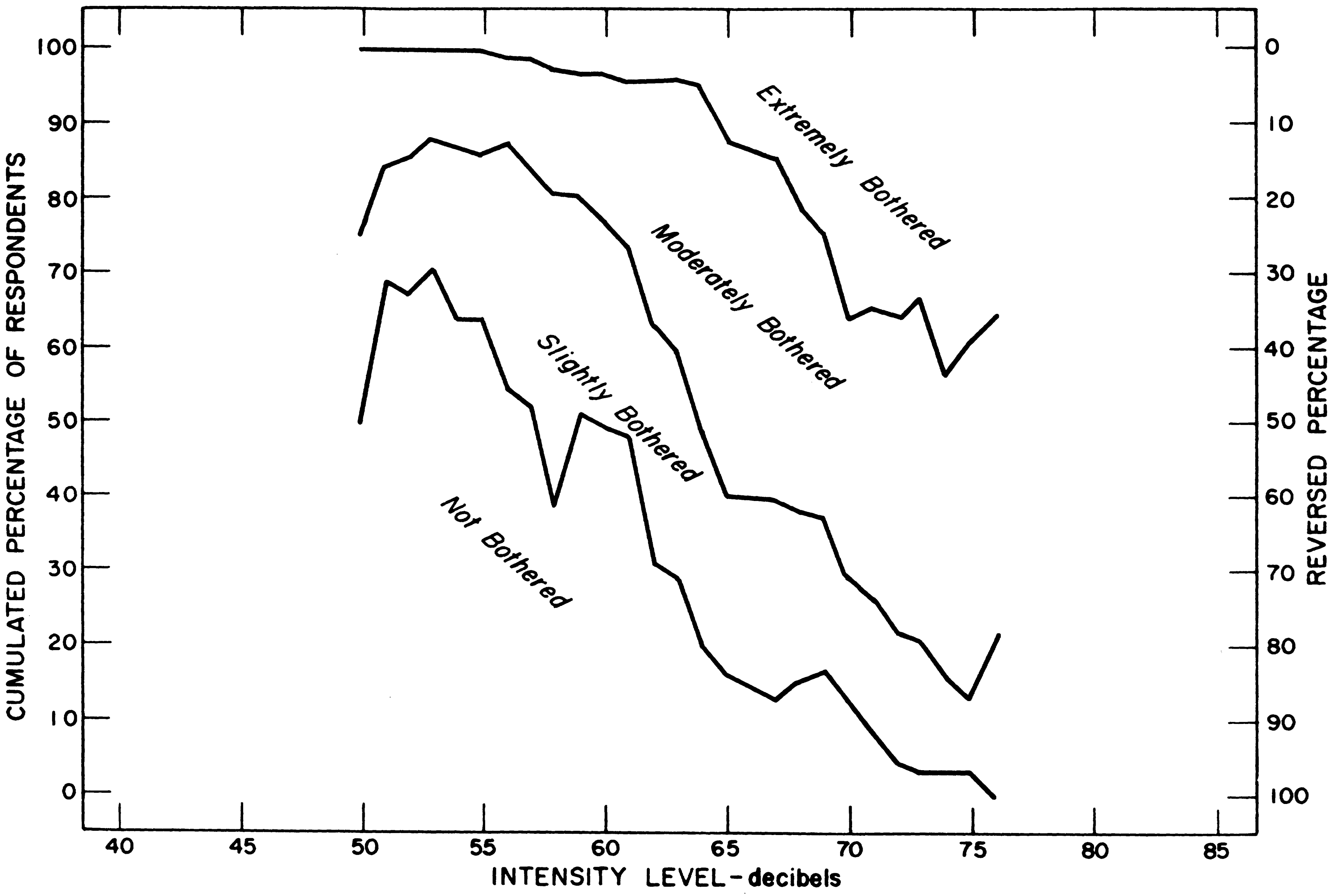


Figure 16. Dryer; source in living room, filtered.

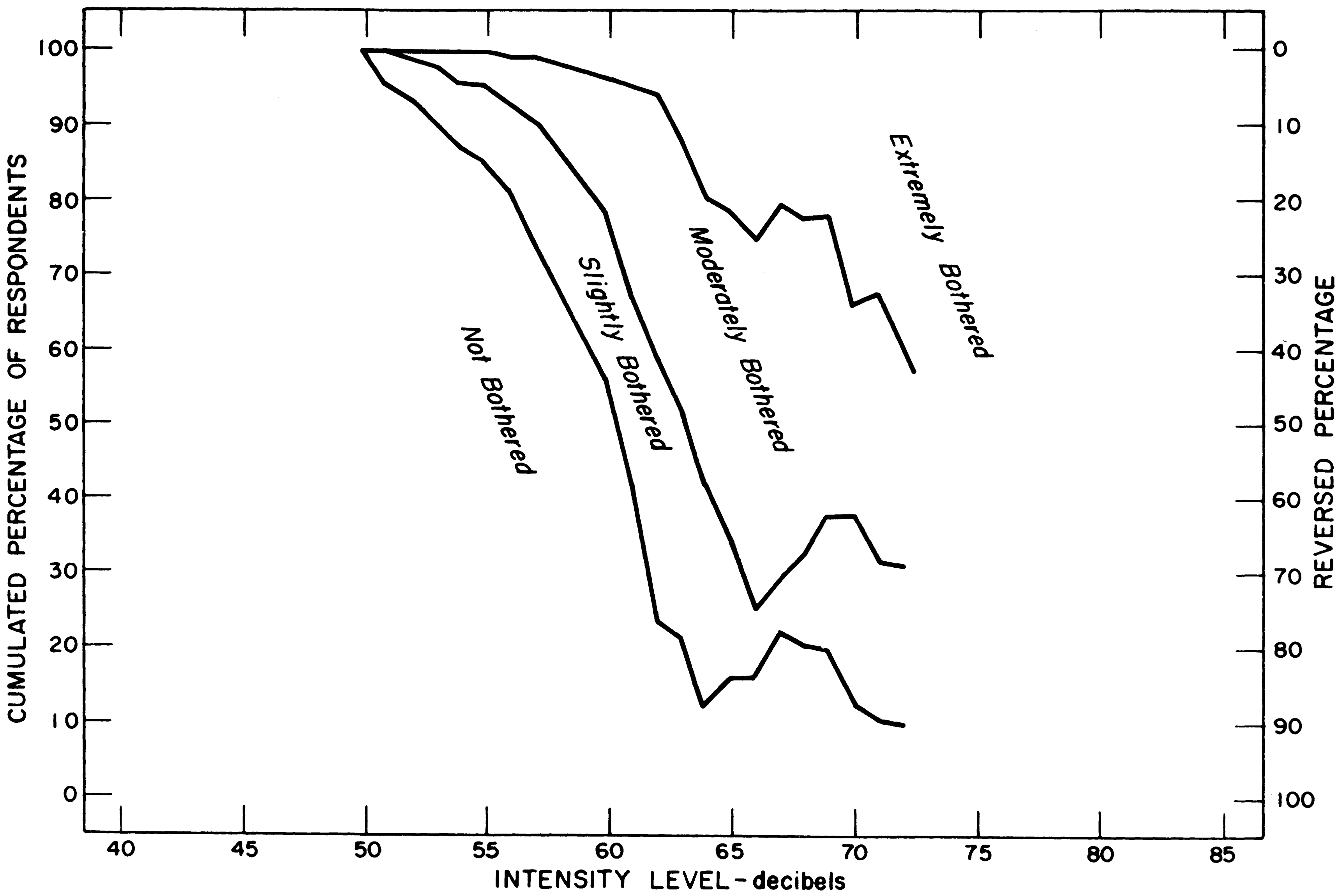


Figure 17. Dryer; source in basement.

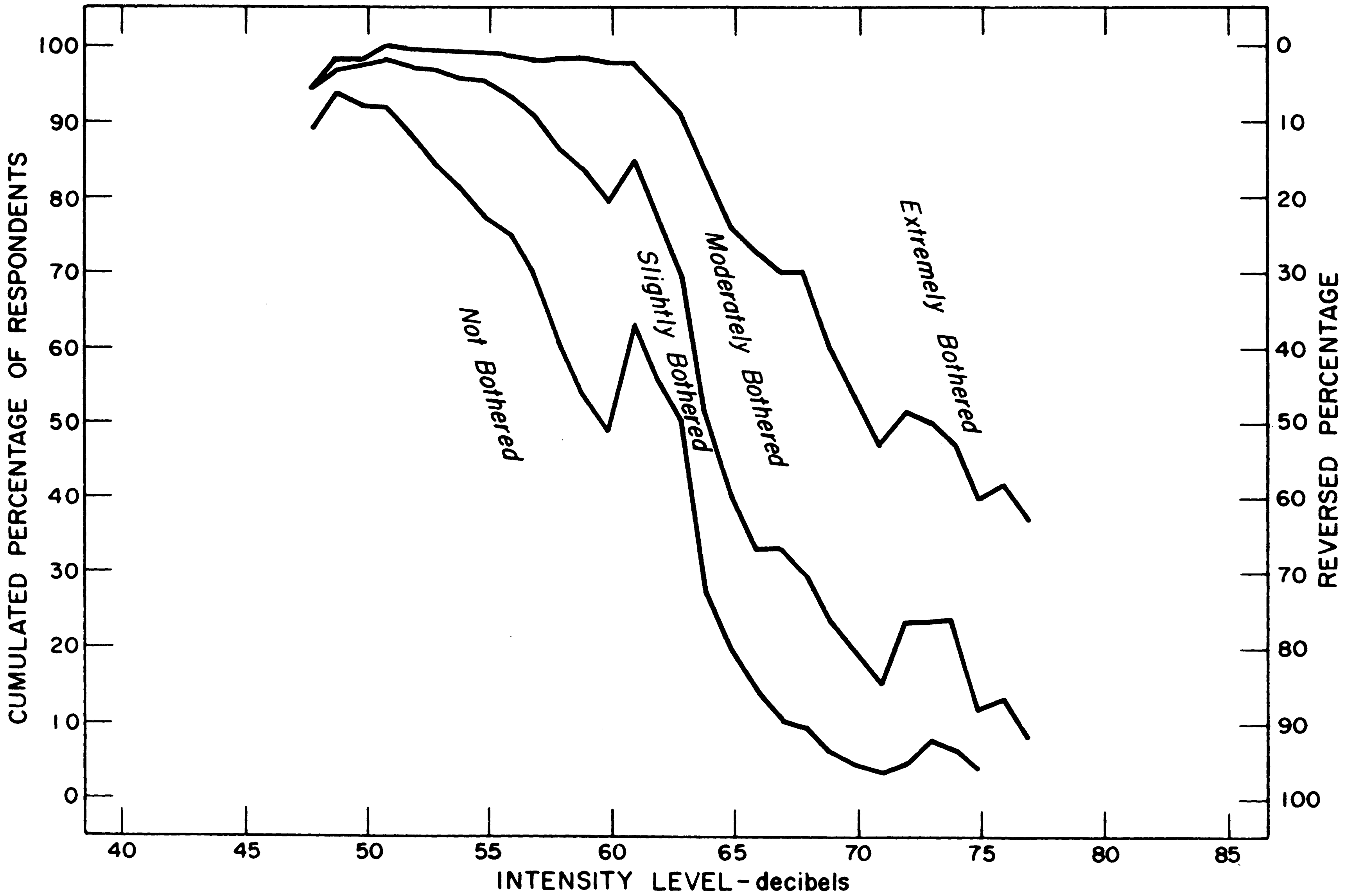


Figure 18. Refrigerator; all source locations merged.

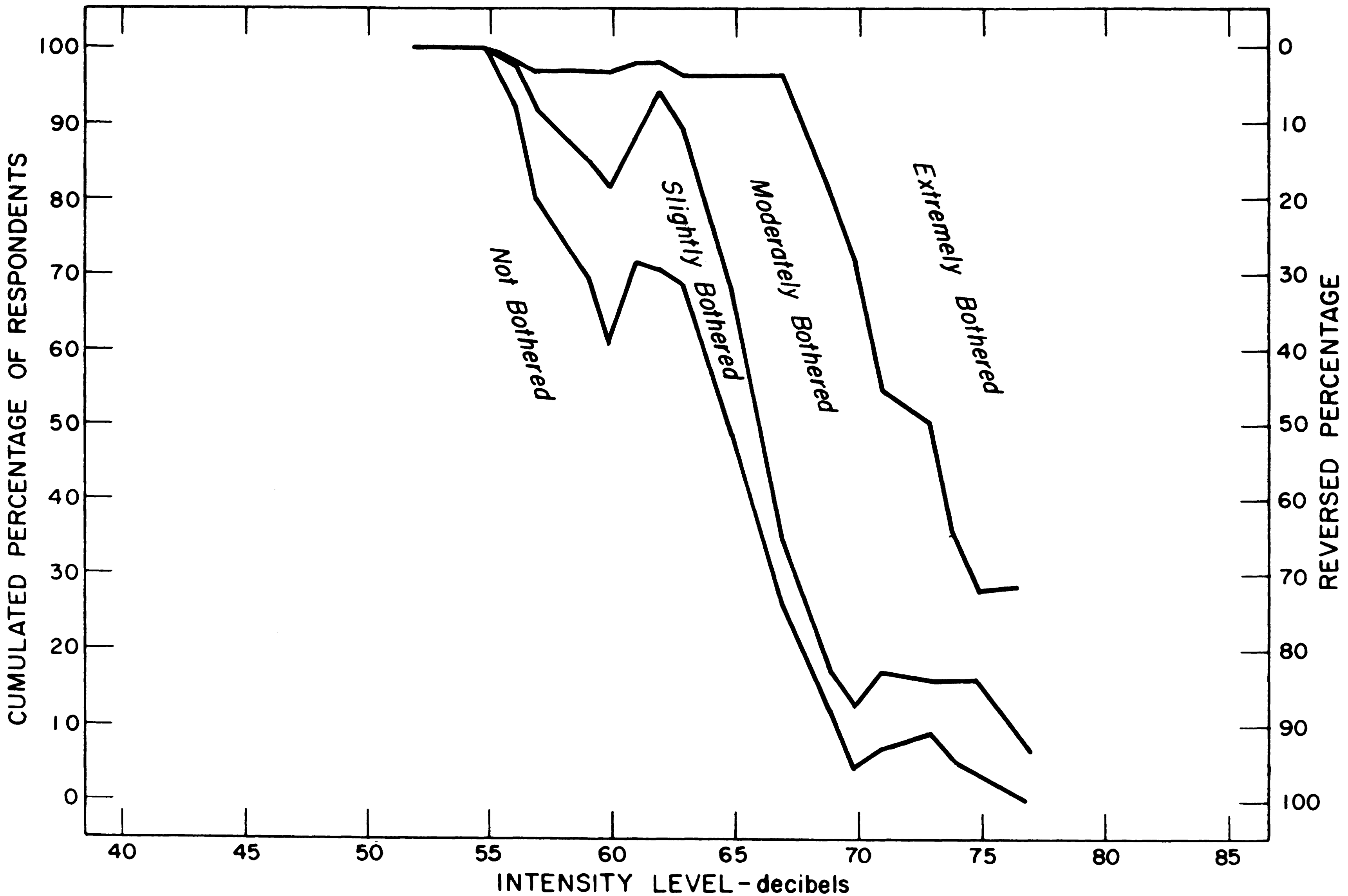


Figure 19. Refrigerator; source in living room, unfiltered.

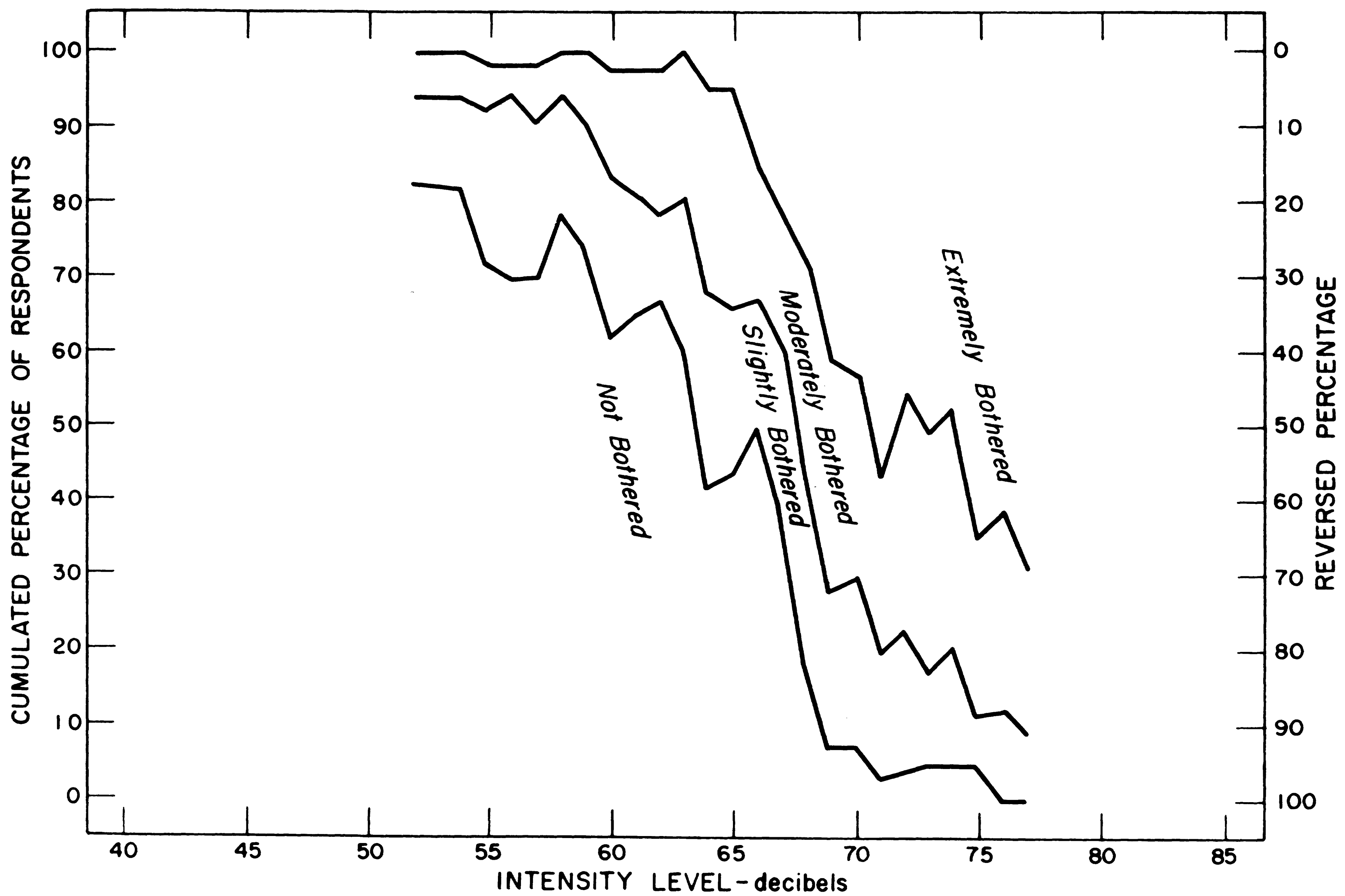


Figure 20. Refrigerator; source in kitchen.

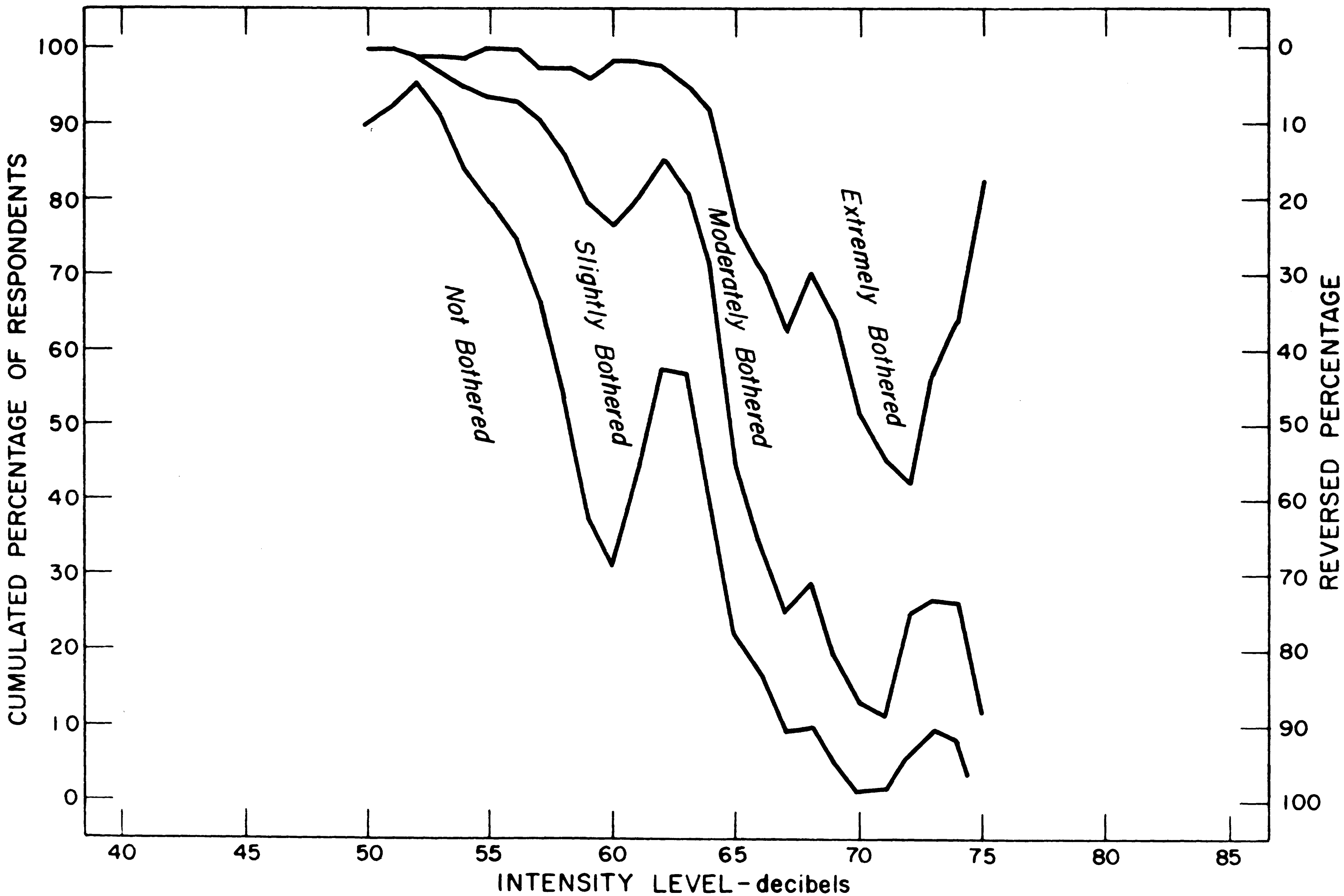


Figure 21. Refrigerator; source in living room, filtered.

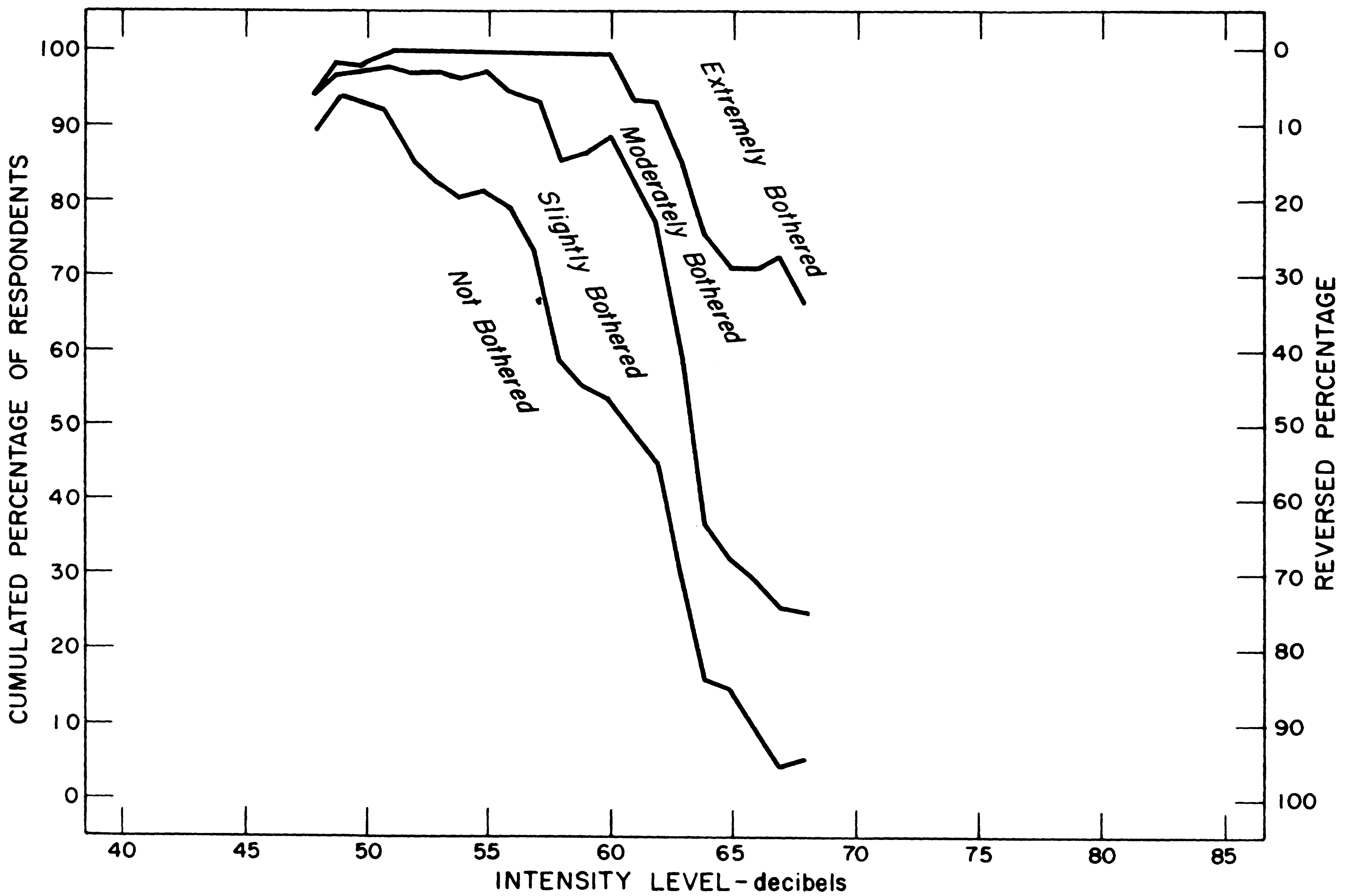


Figure 22. Refrigerator; source in basement.

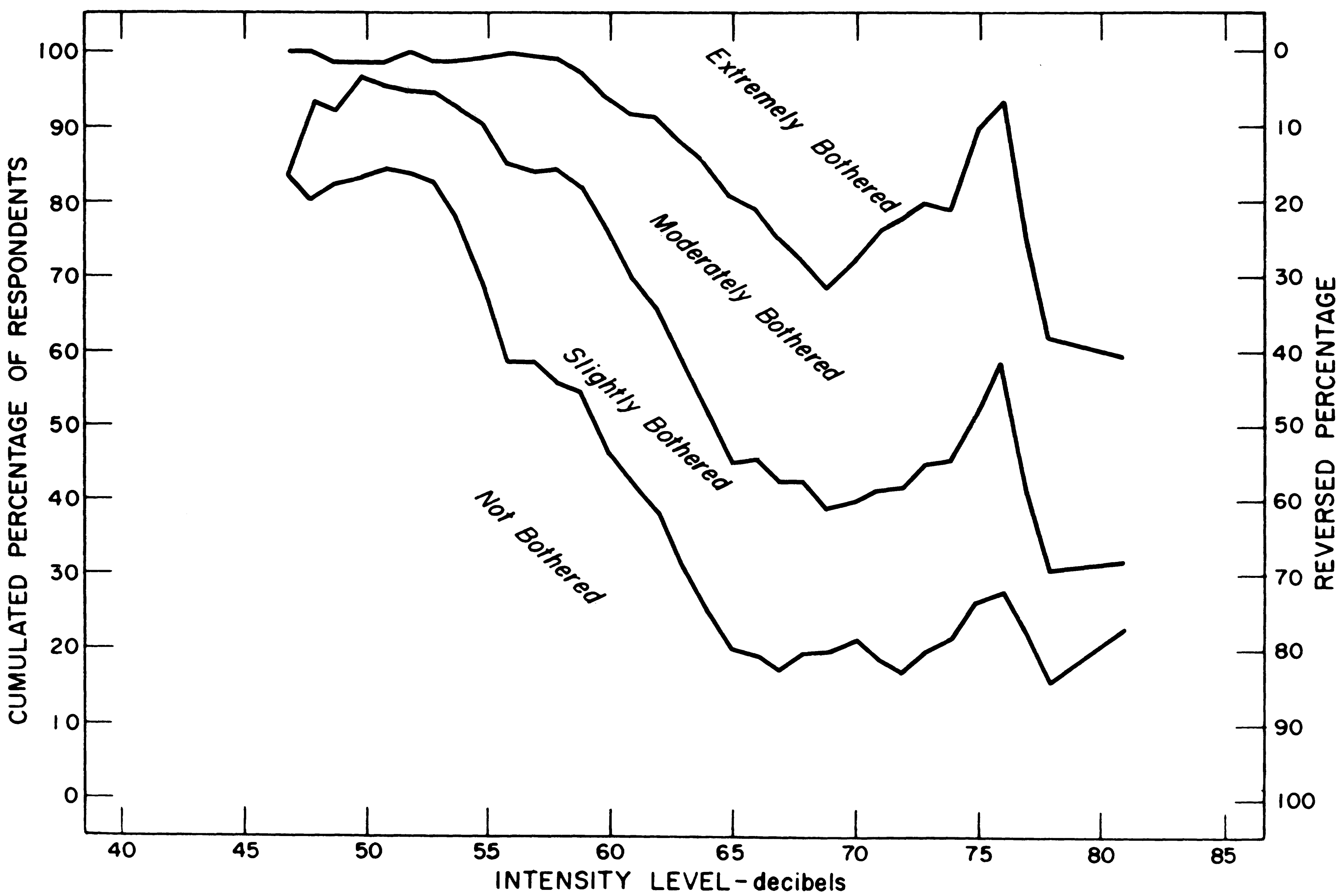


Figure 23. Washer, agitate cycle; all source locations merged.

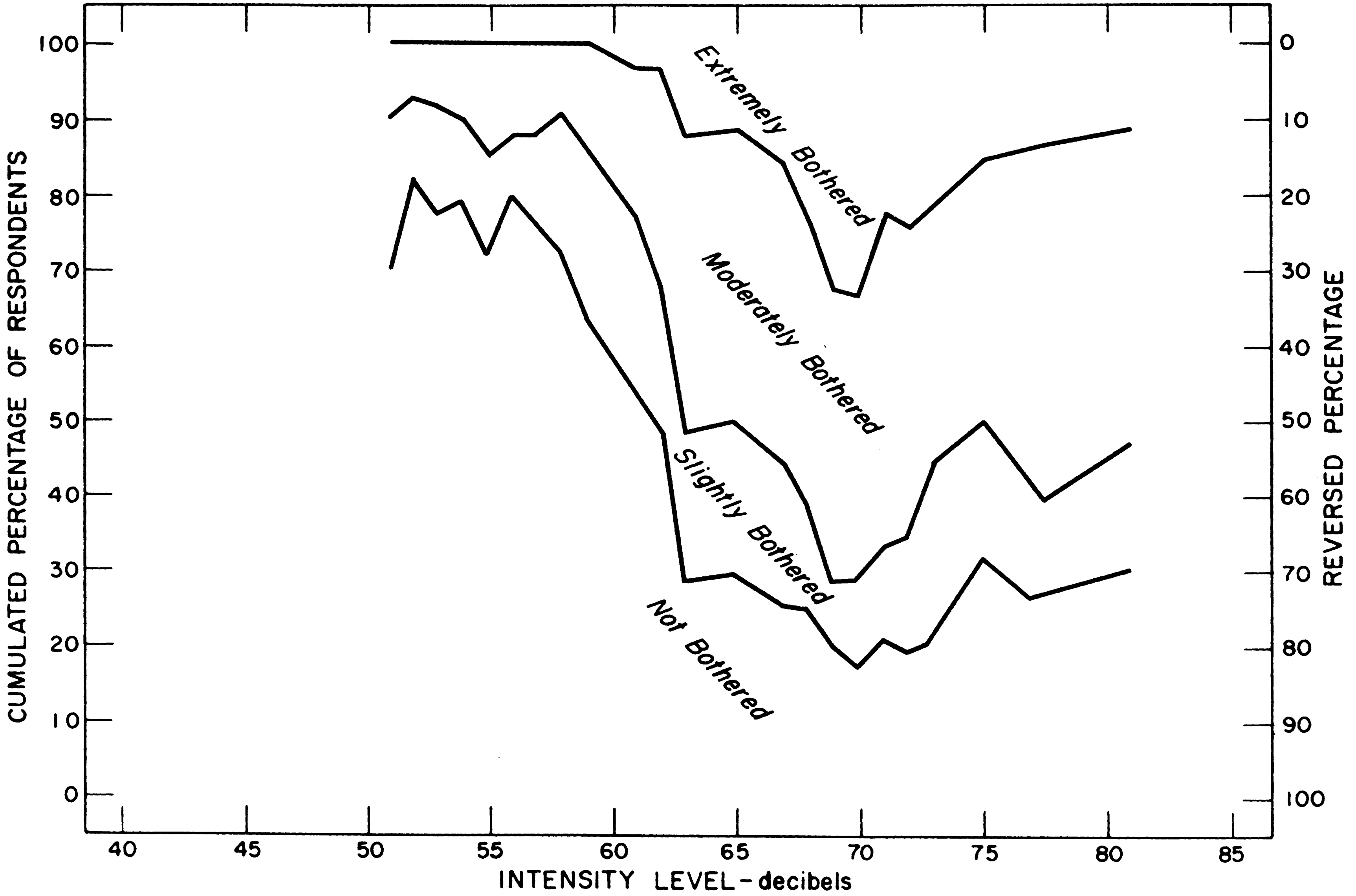


Figure 24. Washer, agitate cycle; source in living room, unfiltered.

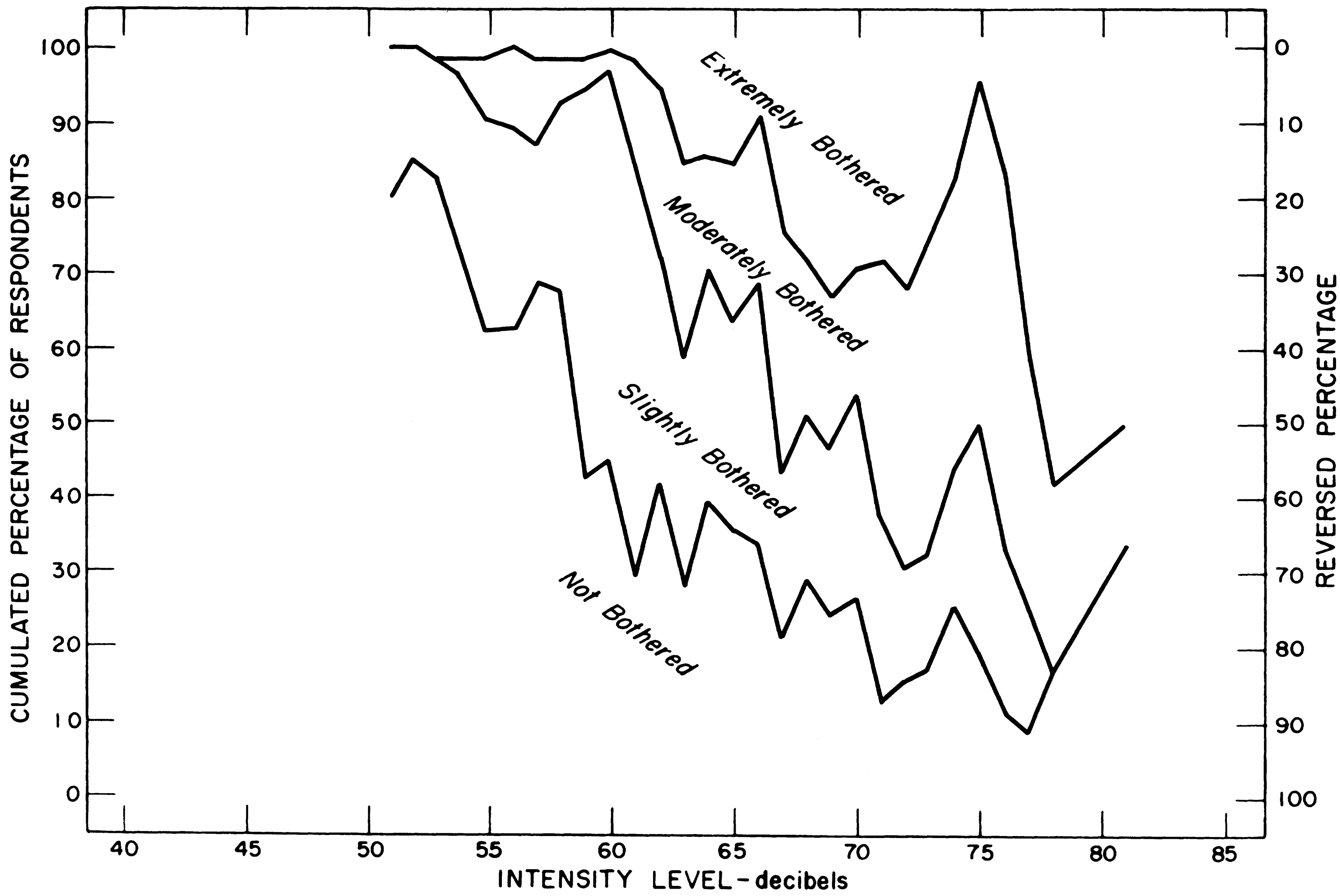


Figure 25. Washer, agitate cycle; source in kitchen.

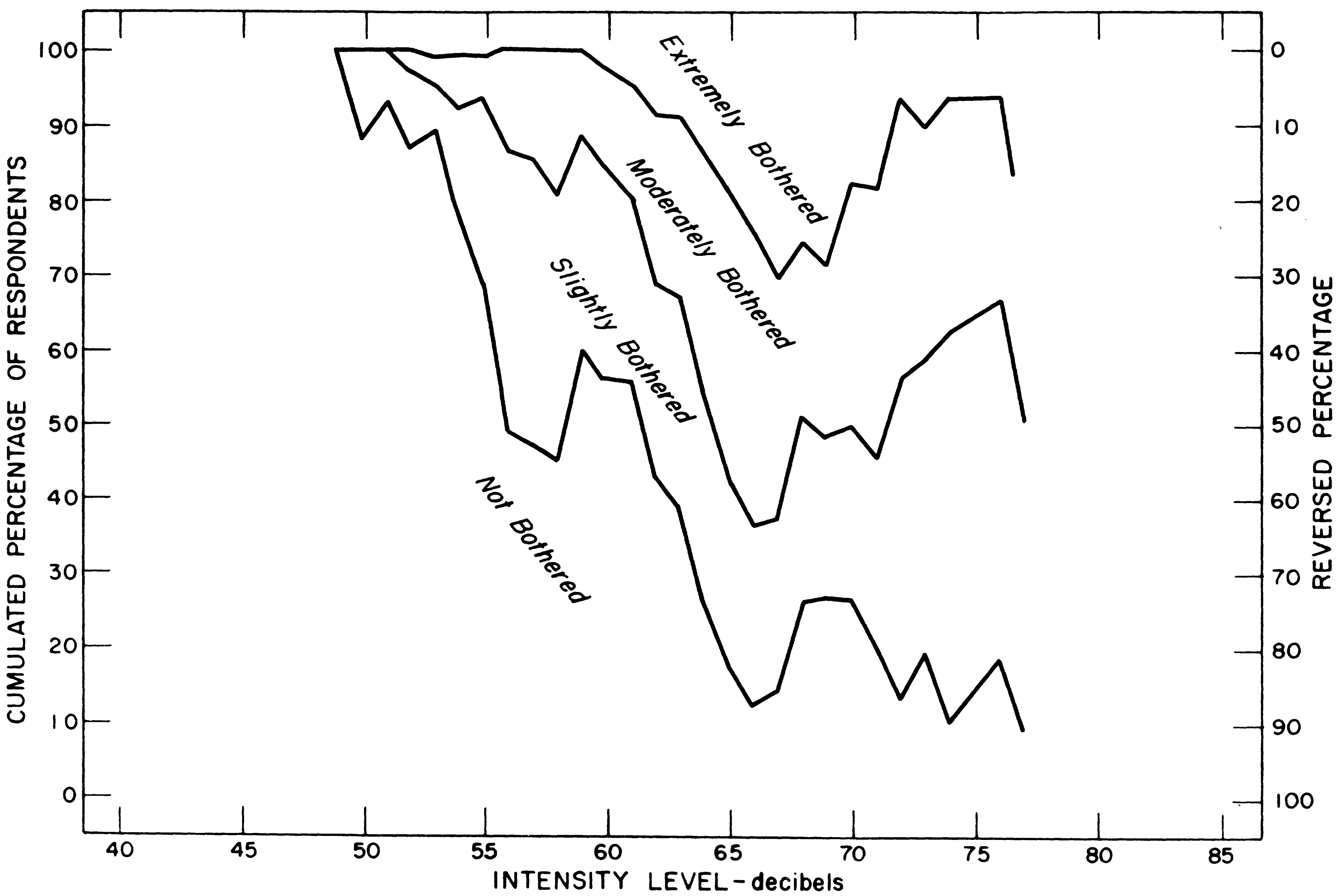


Figure 26. Washer, agitate cycle; source in living room, filtered.

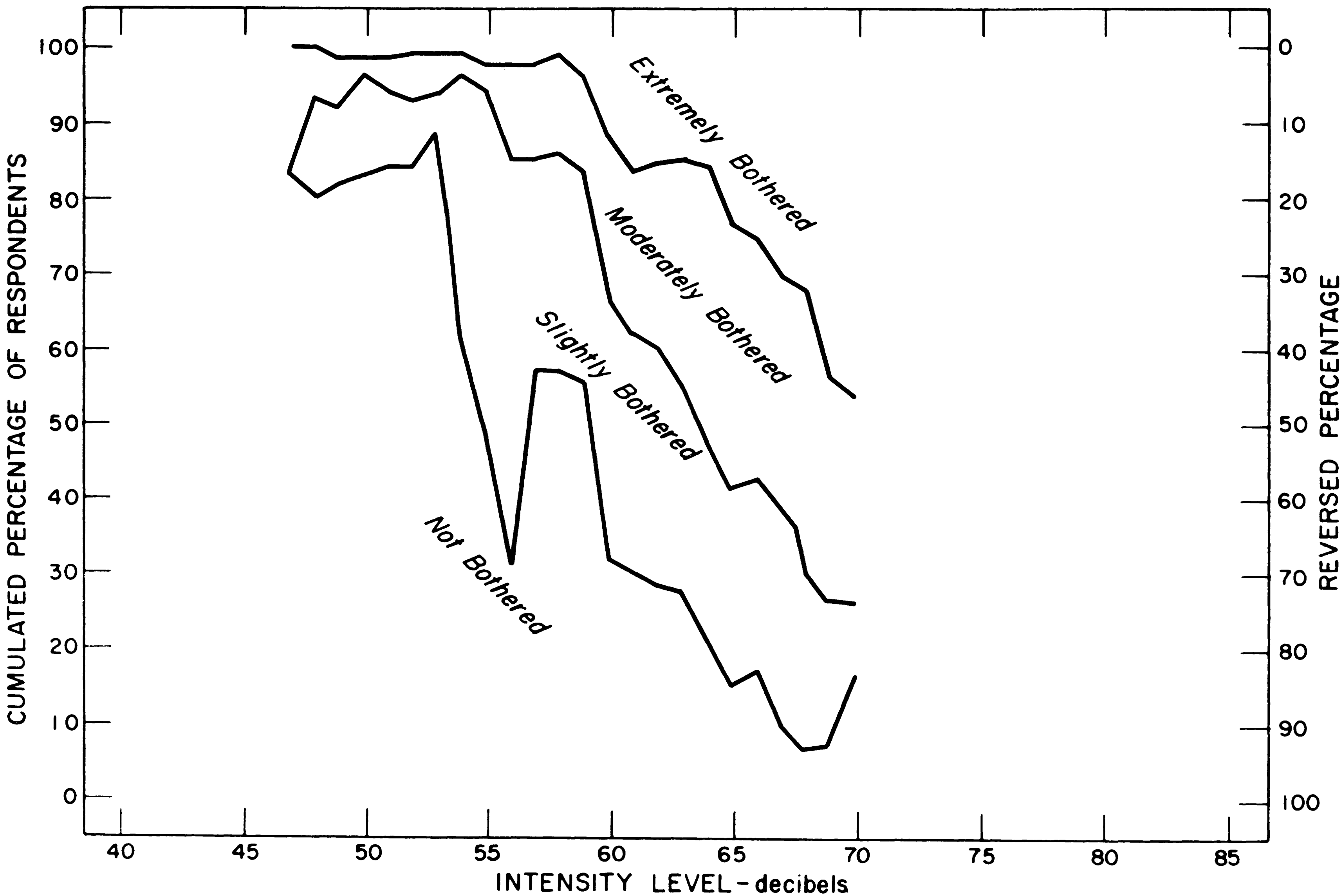


Figure 27. Washer, agitate cycle; source in basement.

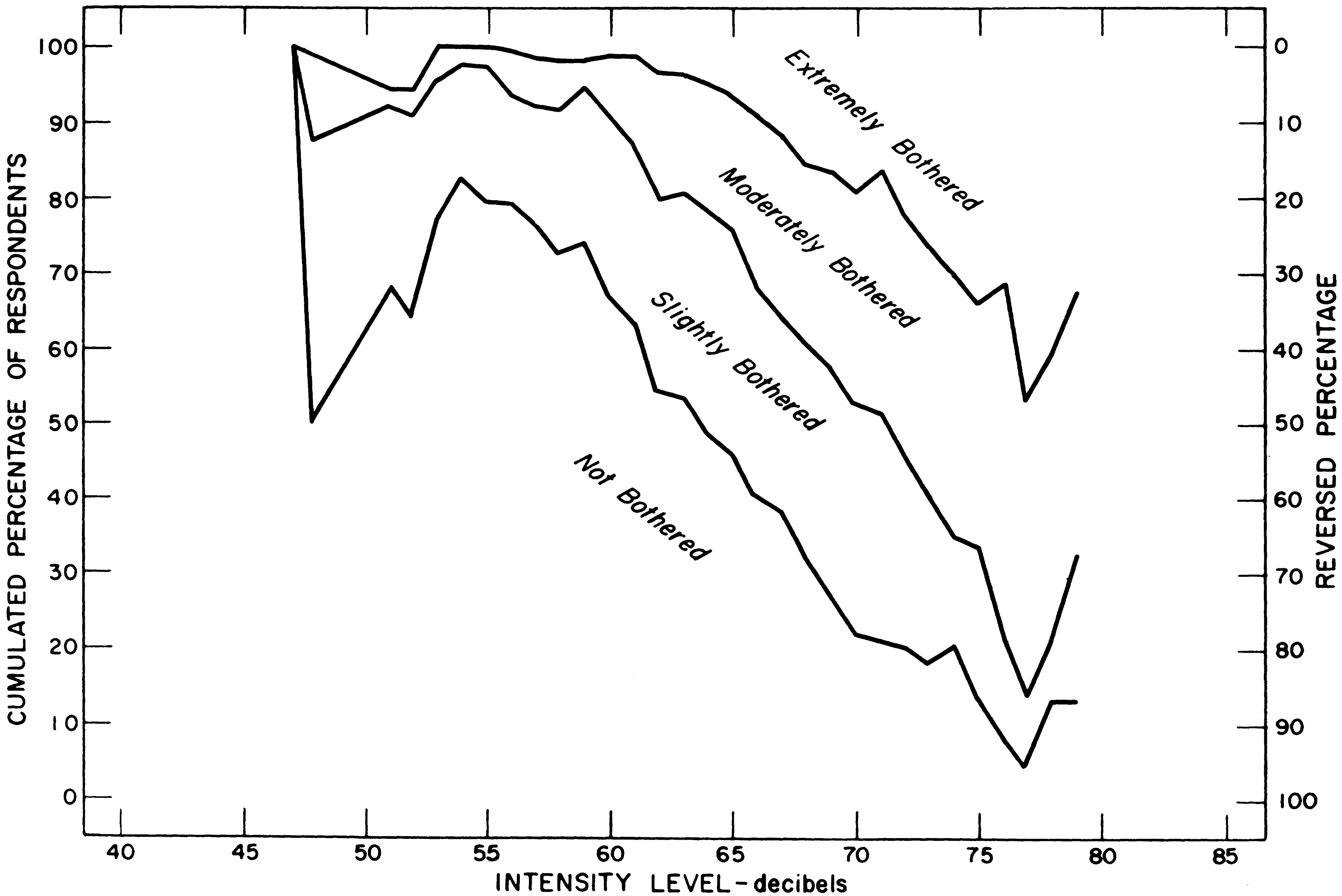


Figure 28. Washer, agitate cycle; all sources merged.

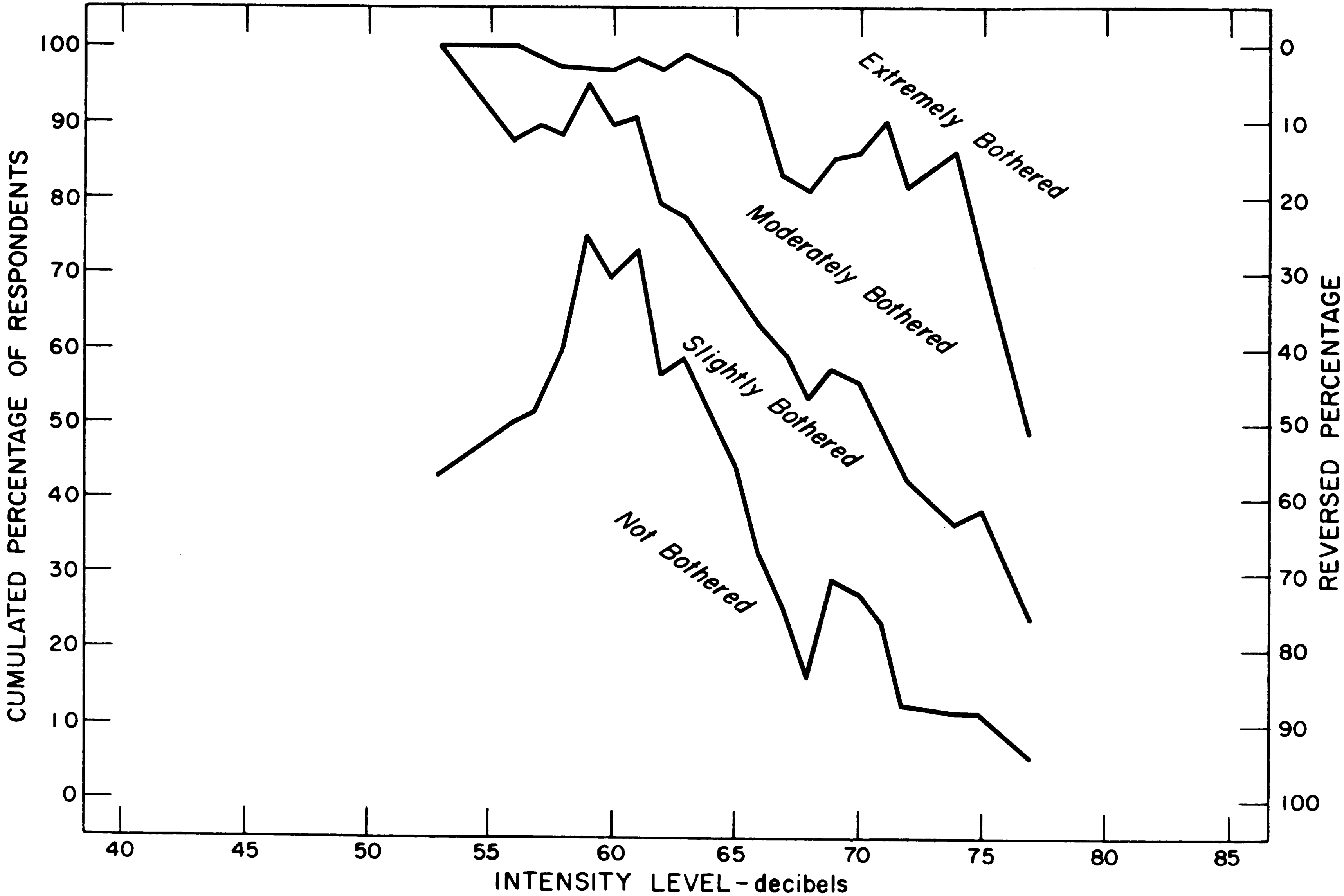


Figure 29. Washer, agitate-filtered; source in living room, unfiltered.

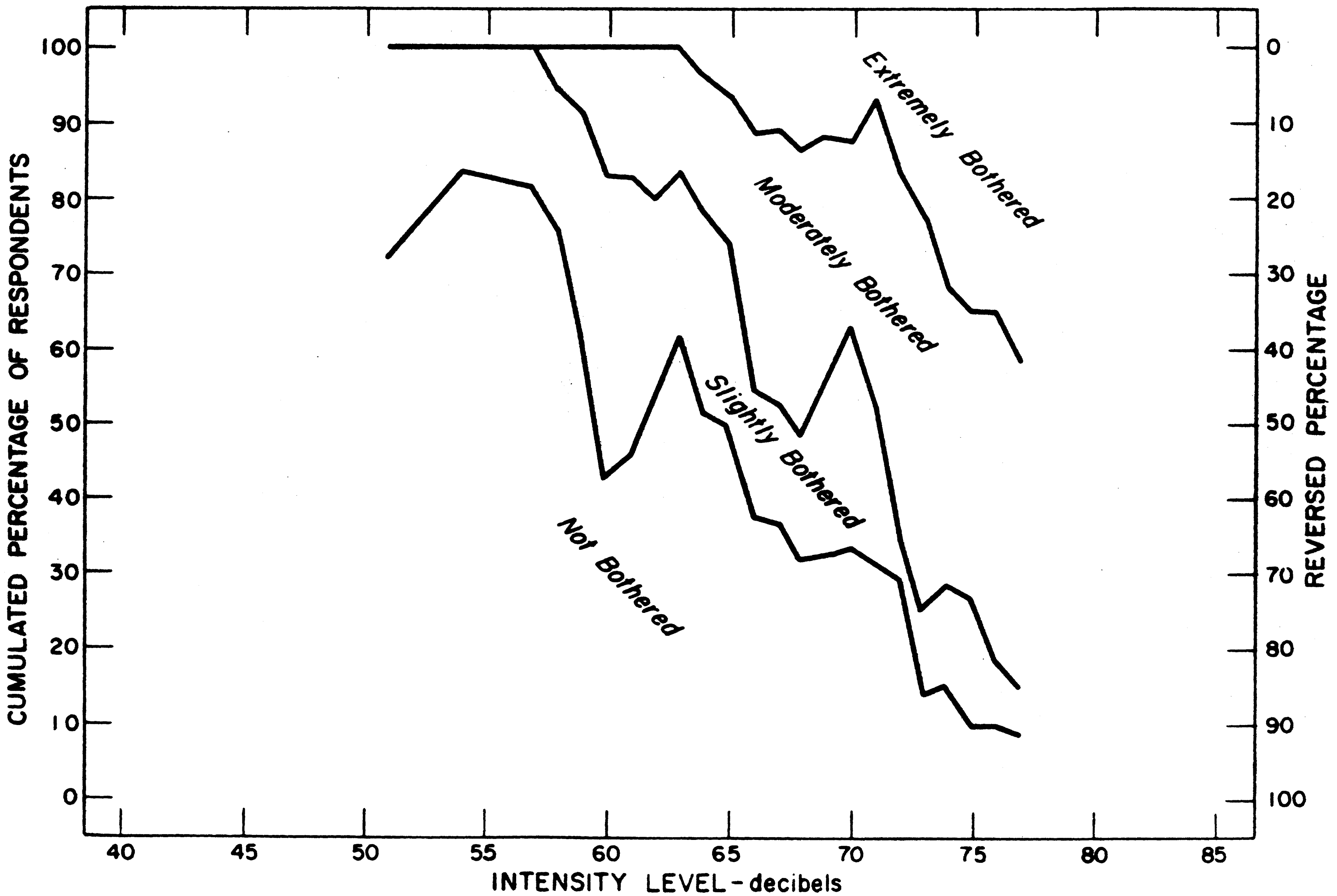


Figure 30. Washer, agitate-filtered; source in kitchen.

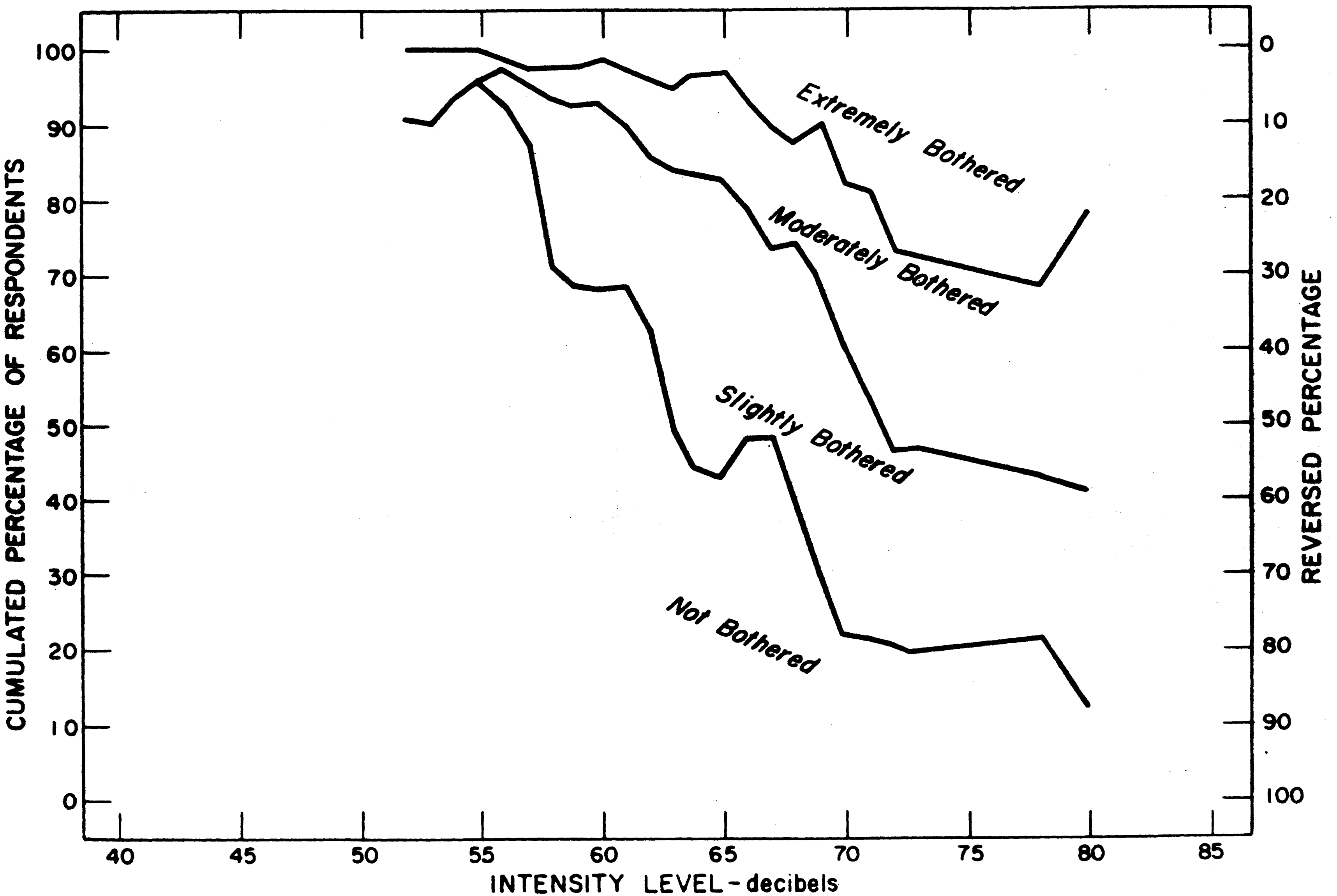


Figure 31. Washer, agitate-filtered; source in living room, filtered.

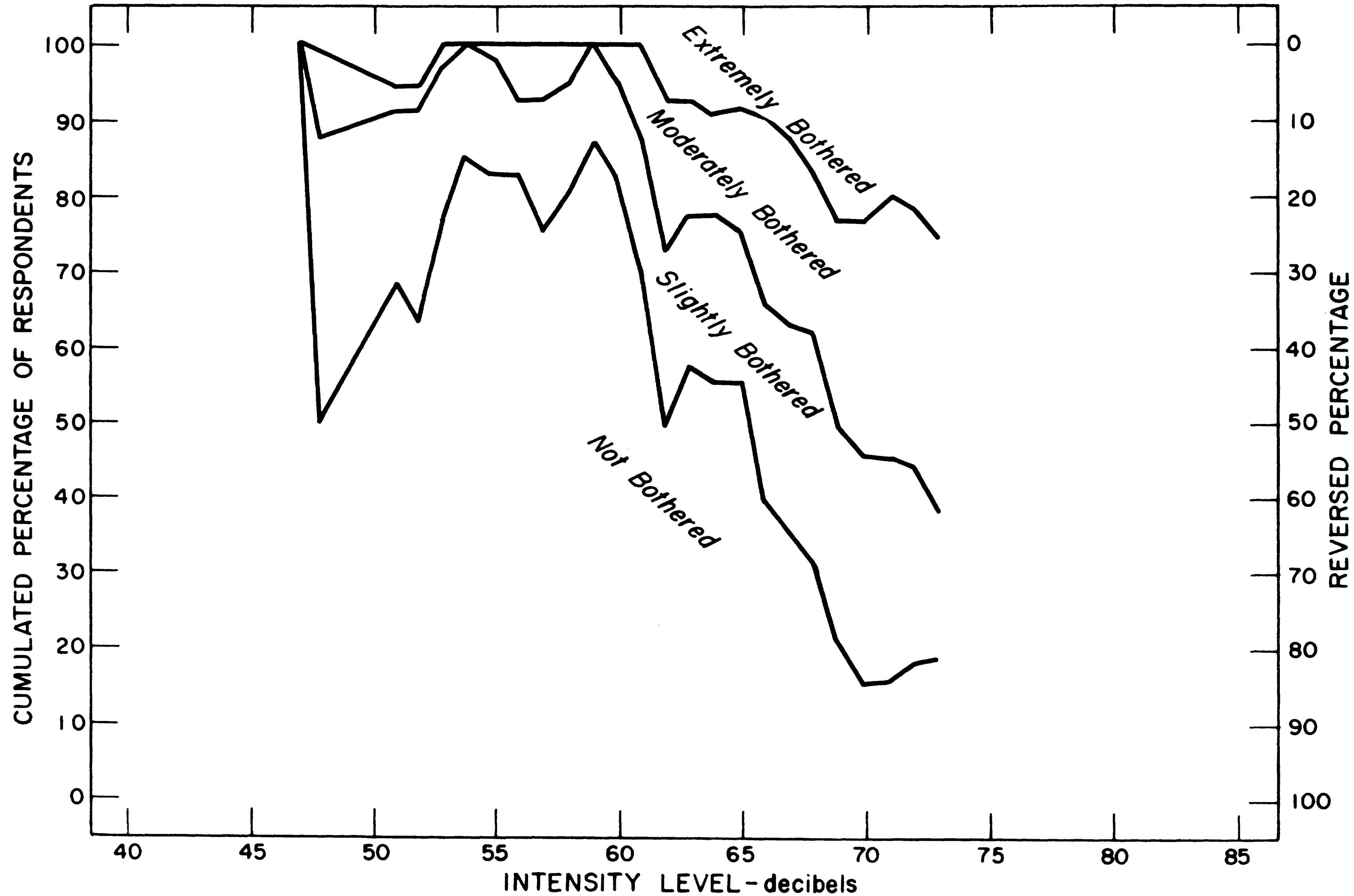


Figure 32. Washer, agitate-filtered; source in basement.

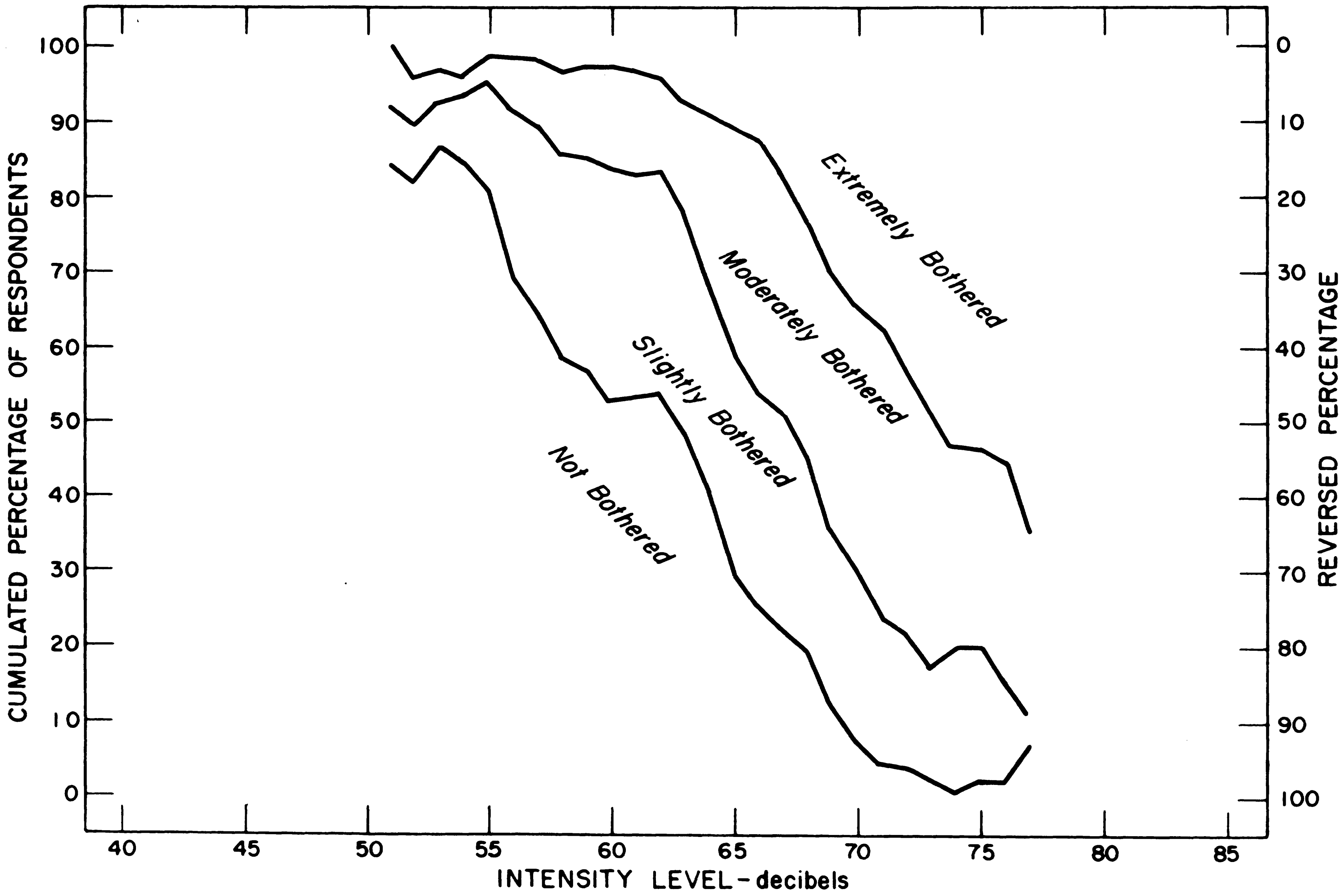


Figure 33. 1/3 octave band; all source locations merged.

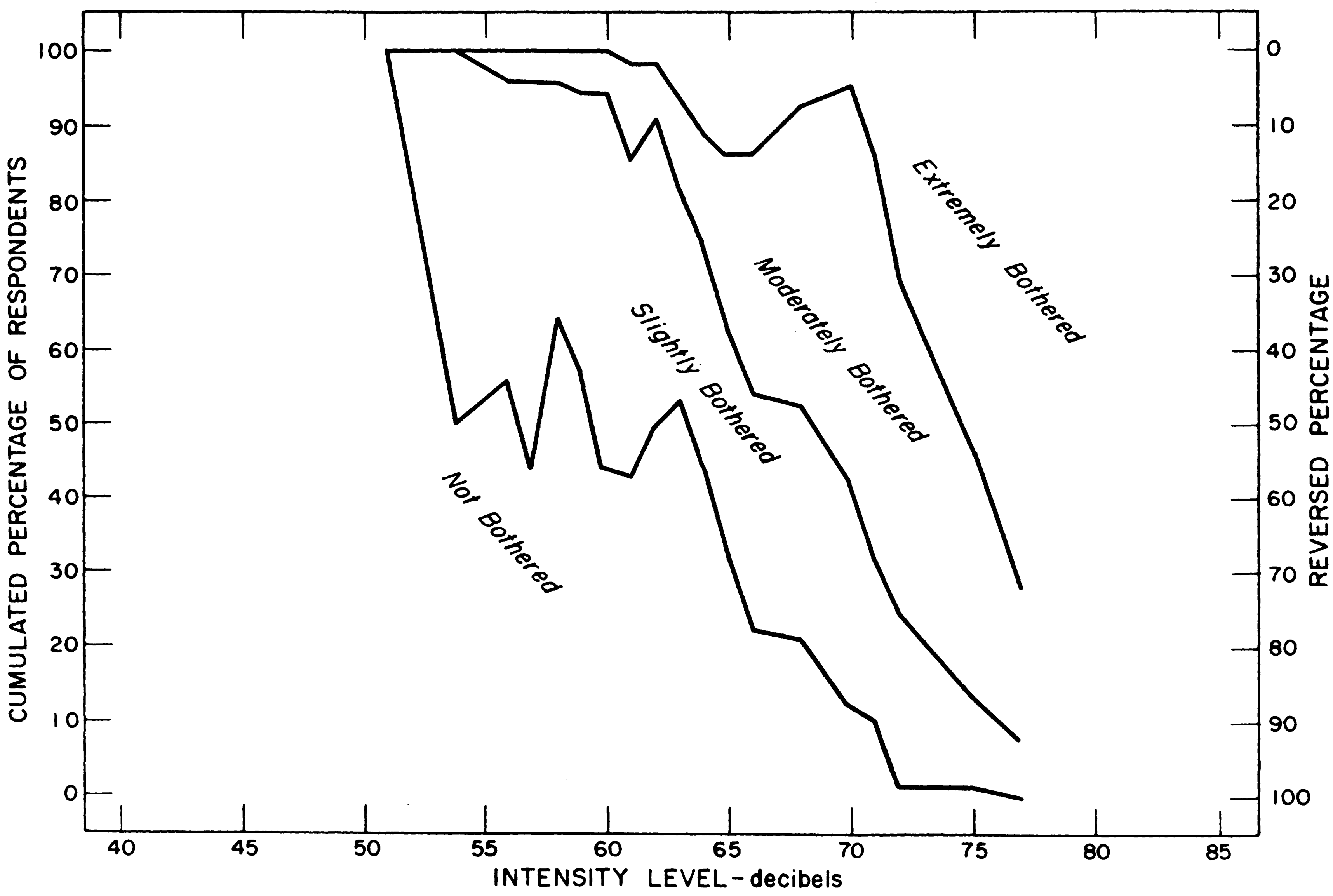


Figure 34. 1/3 octave band; source in living room, unfiltered.

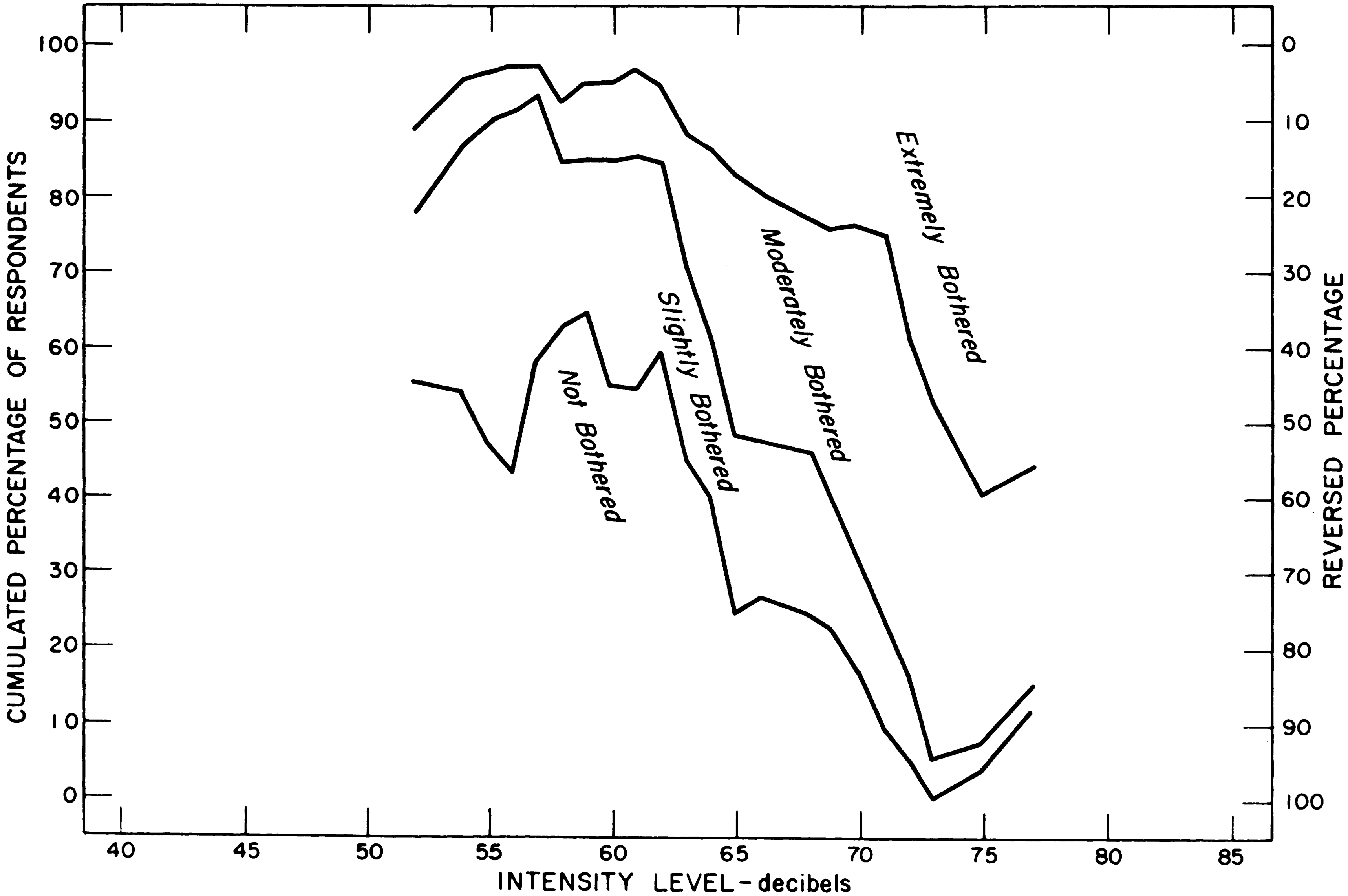


Figure 35. 1/3 octave band; source in kitchen.

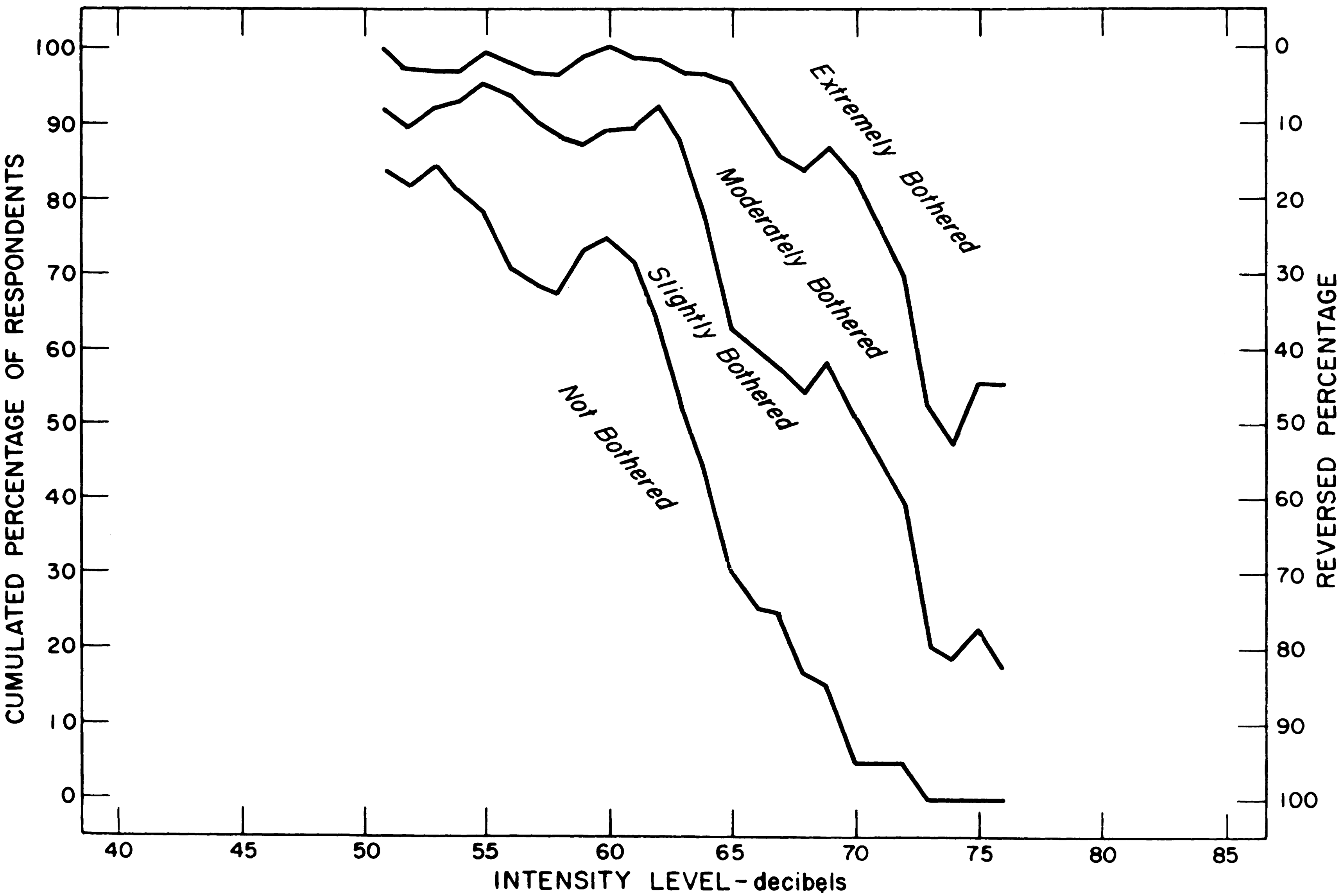


Figure 36. 1/3 octave band; source in living room, filtered.

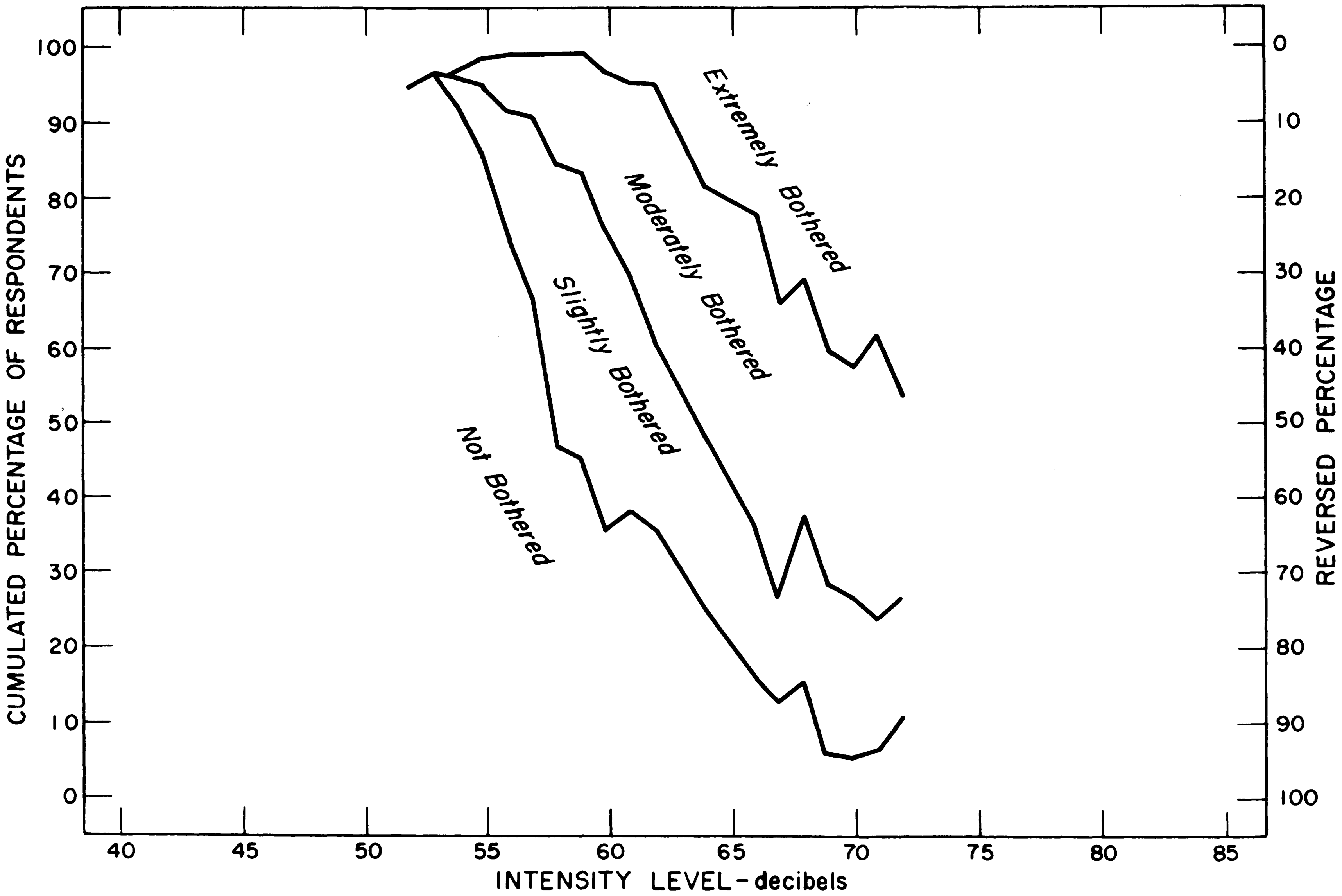


Figure 37. 1/3 octave band; source in basement.

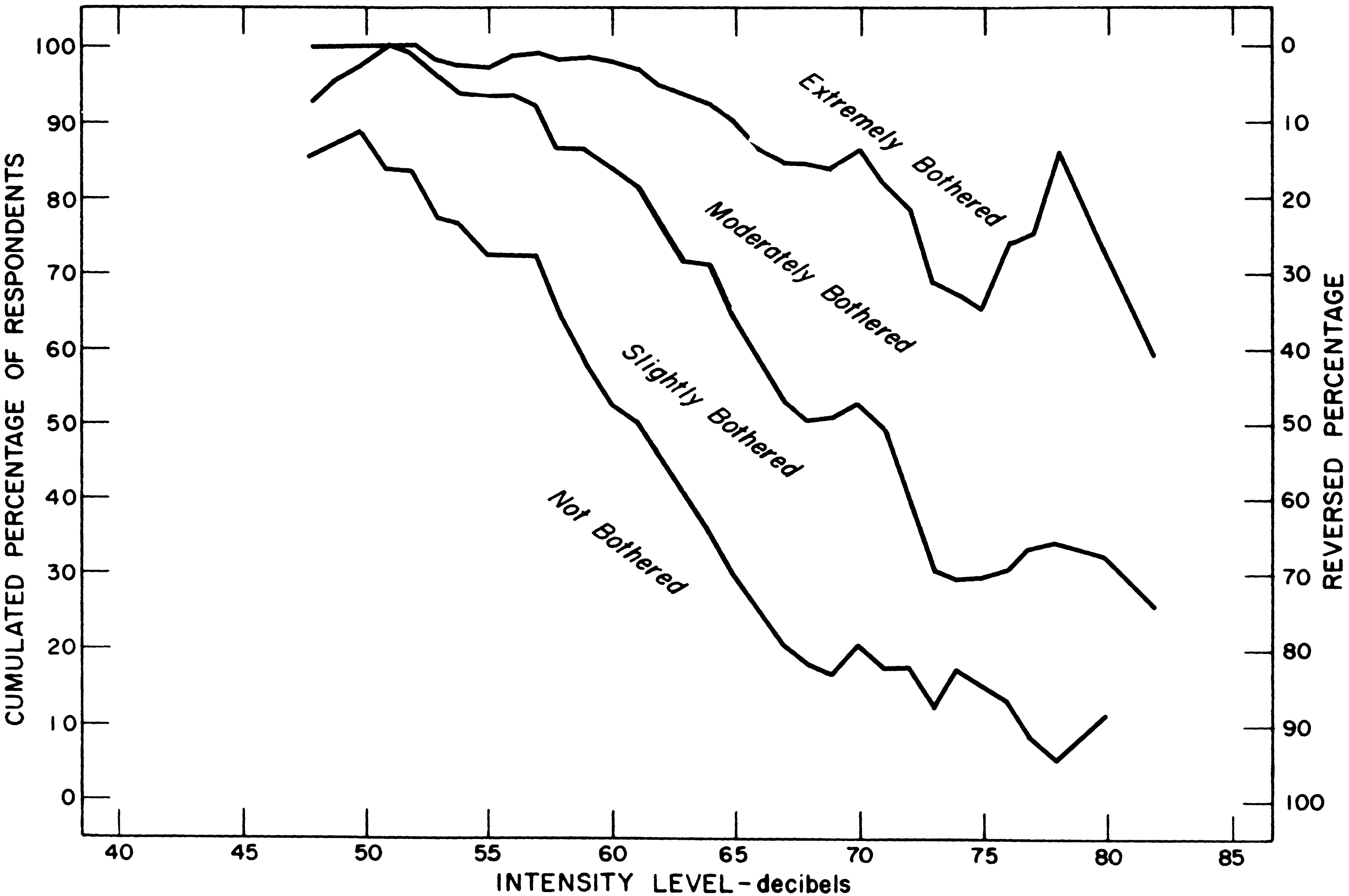


Figure 38. 1/3 octave + pure tone; all source locations merged.

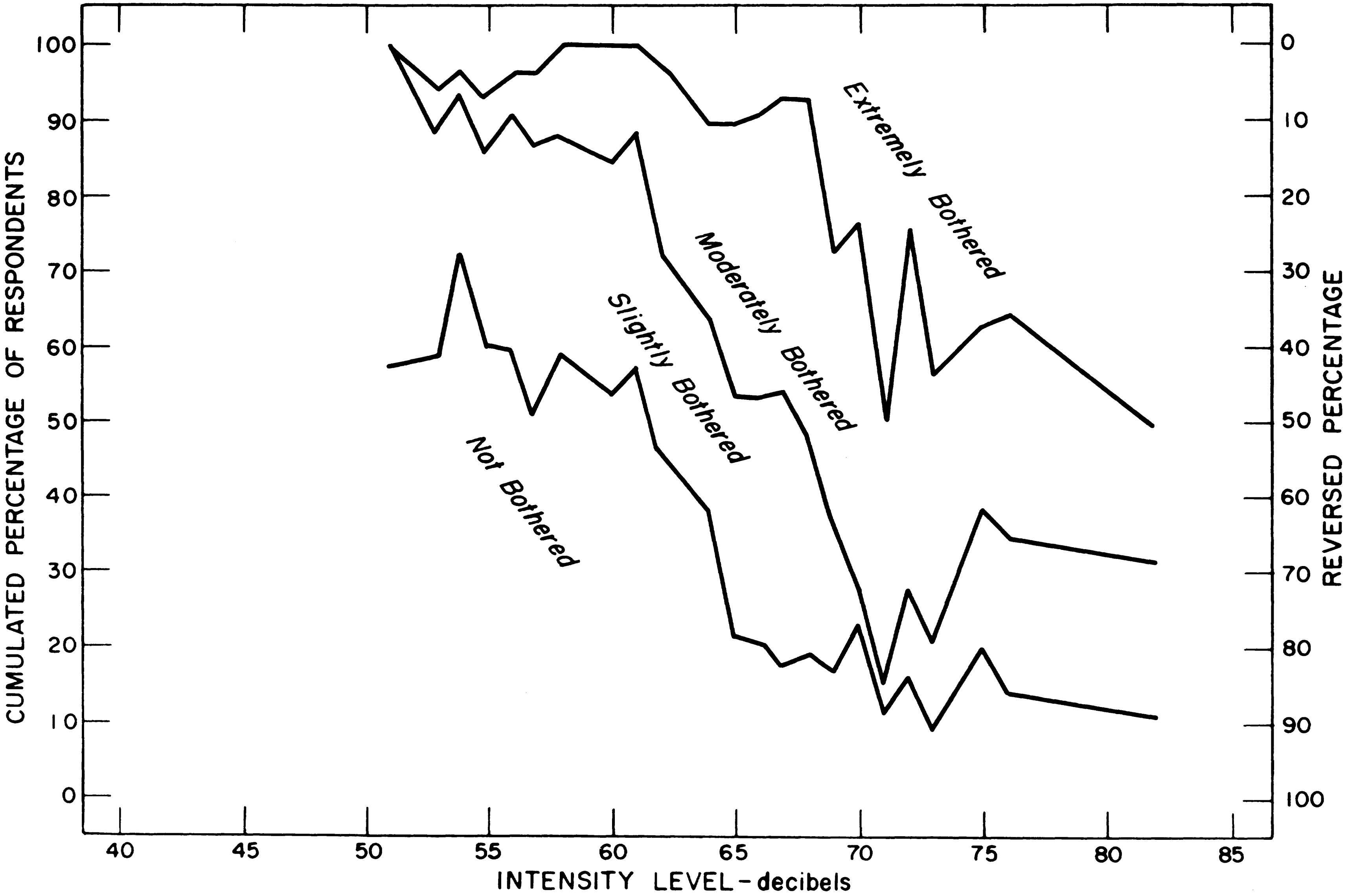


Figure 39. 1/3 octave + pure tone; source in living room, unfiltered.

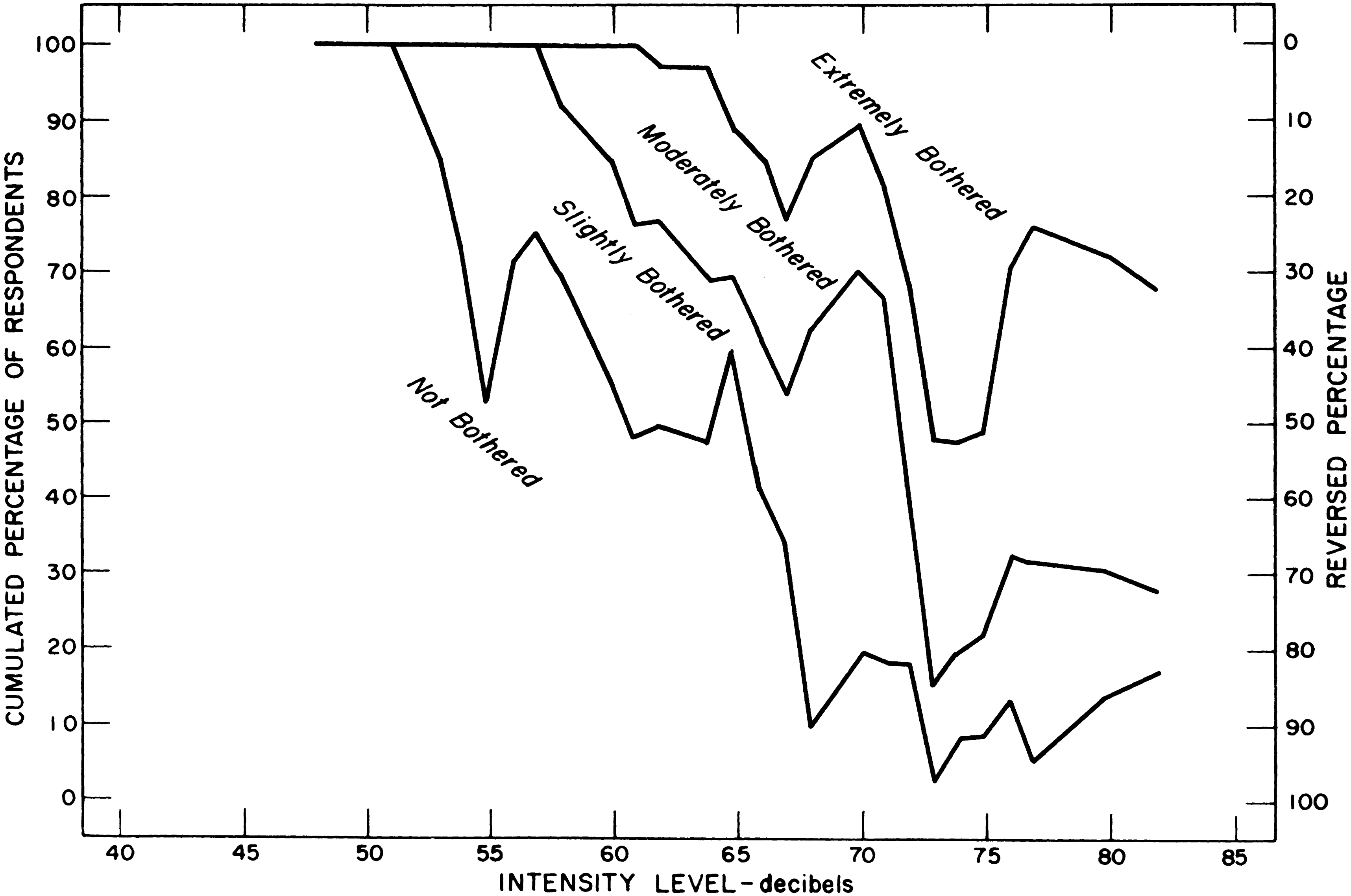


Figure 40. 1/3 octave + pure tone; source in kitchen.

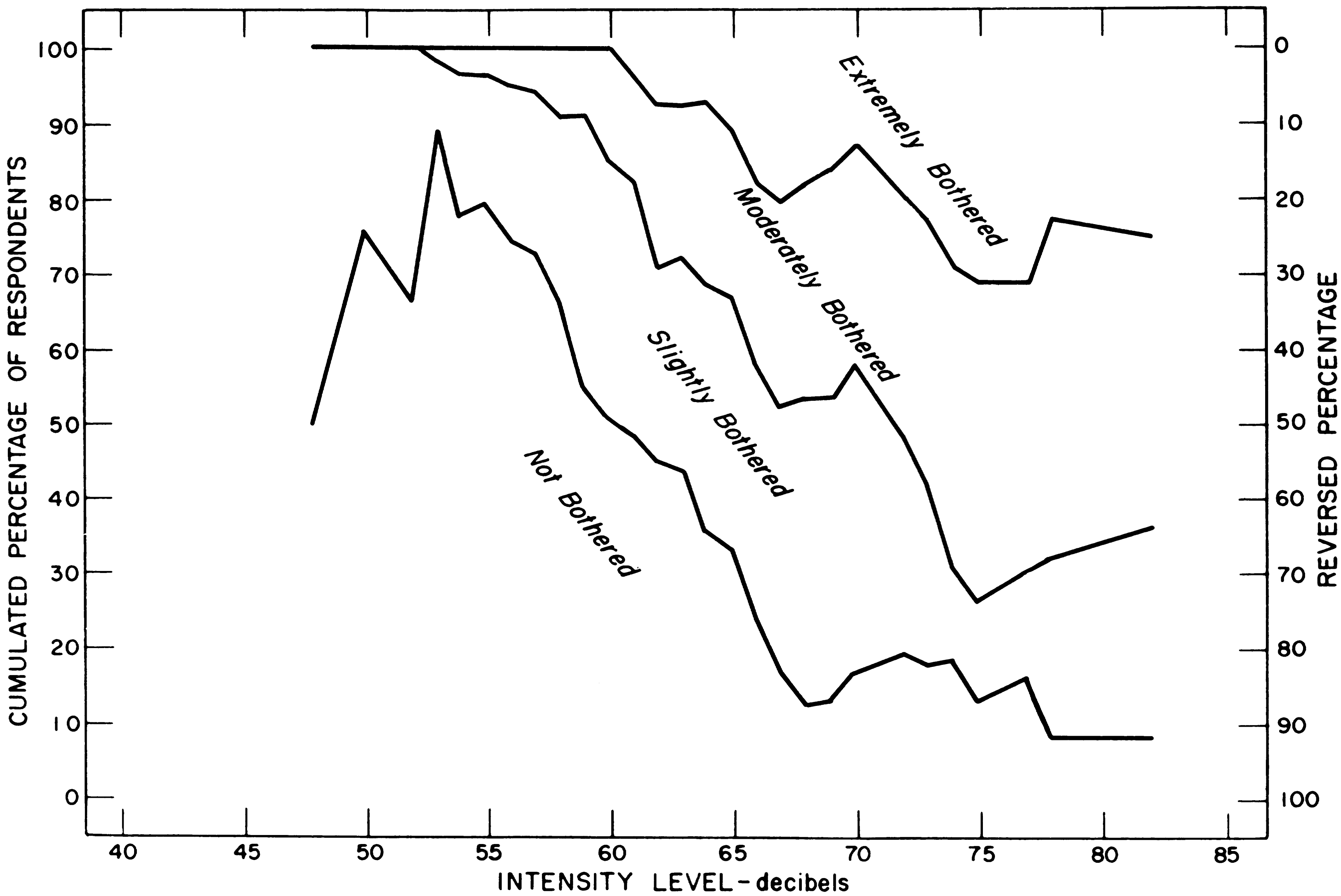


Figure 41. 1/3 octave + pure tone; source in living room, filtered.

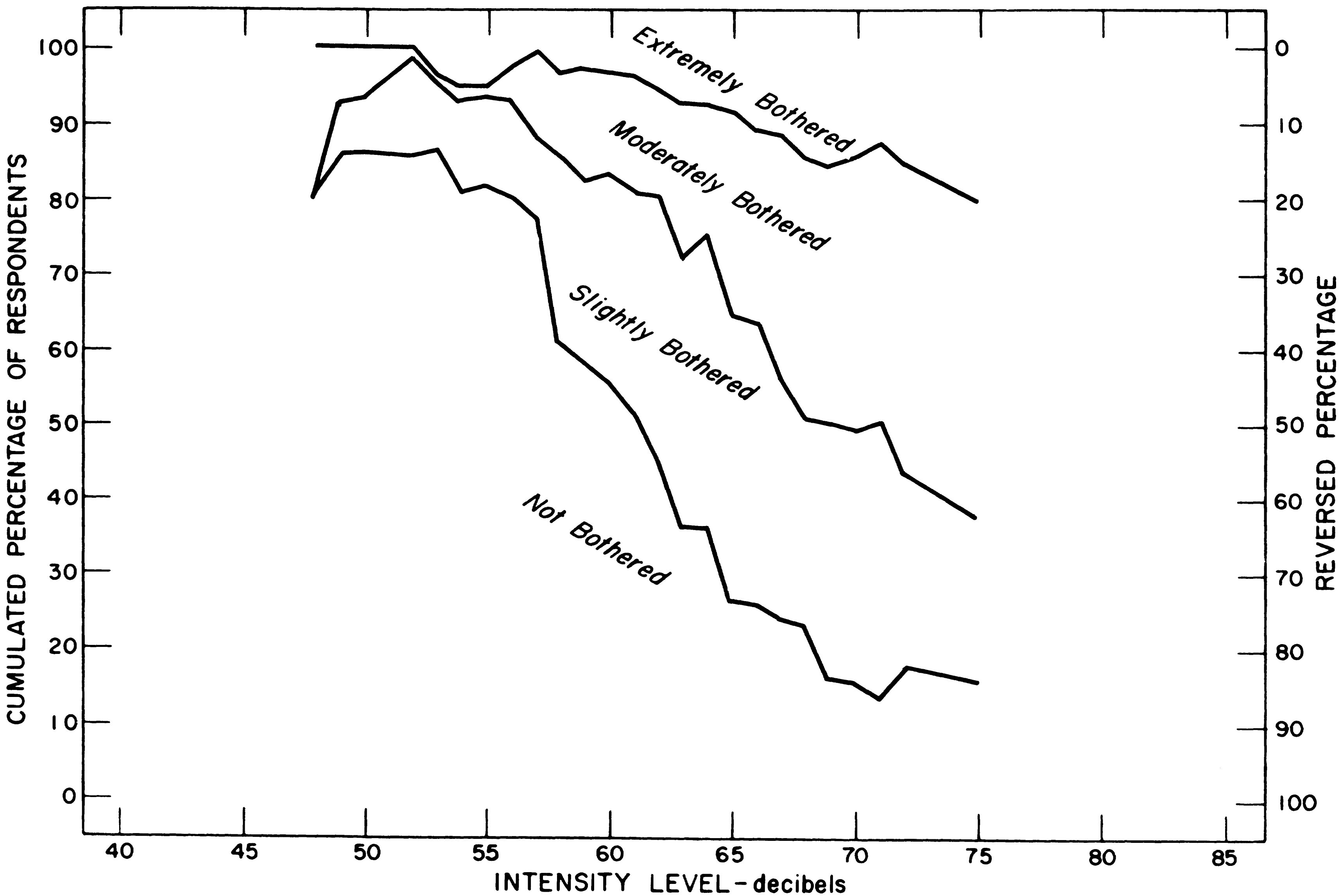


Figure 42. 1/3 octave + pure tone; source in basement.

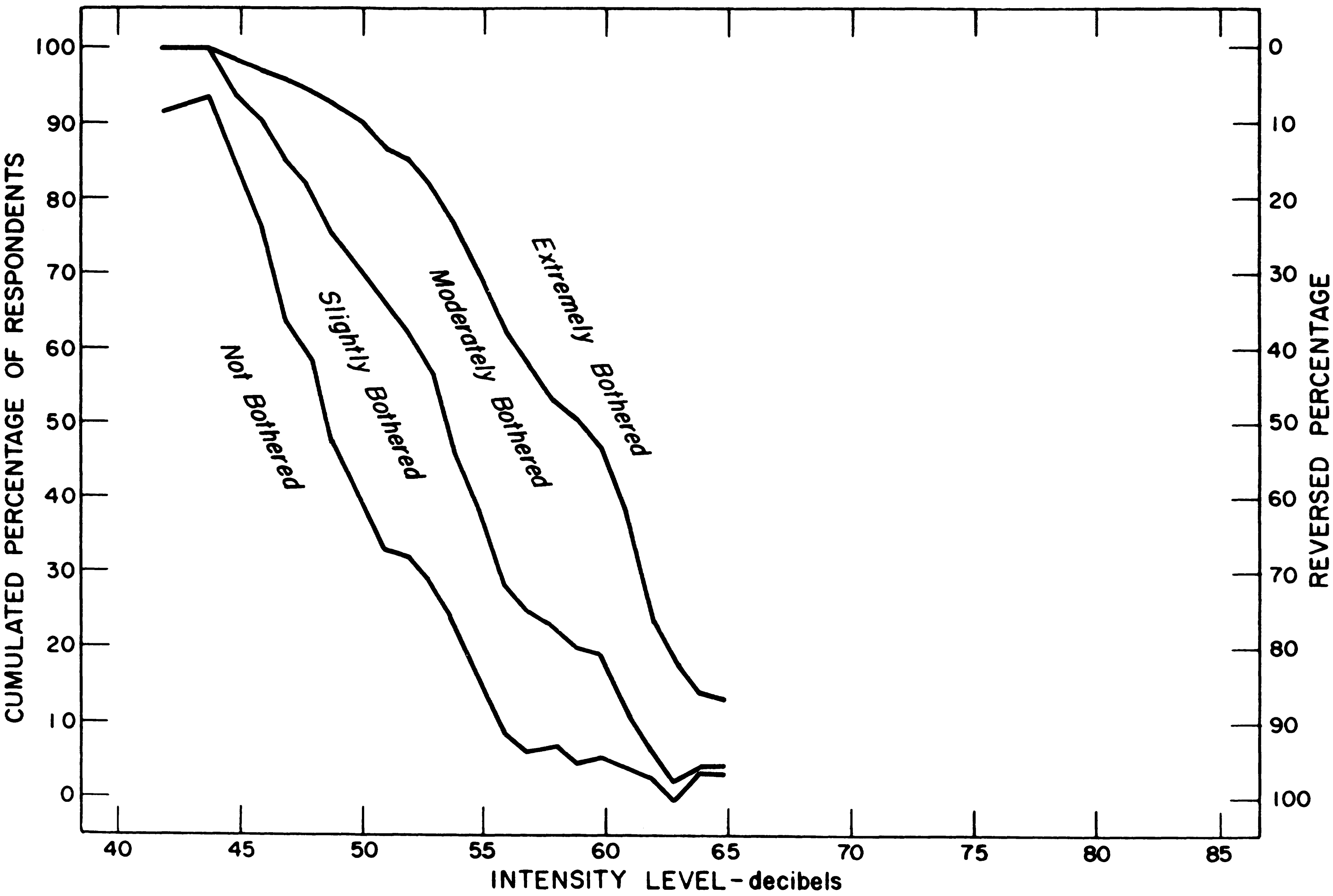


Figure 43. White noise; all source locations merged.

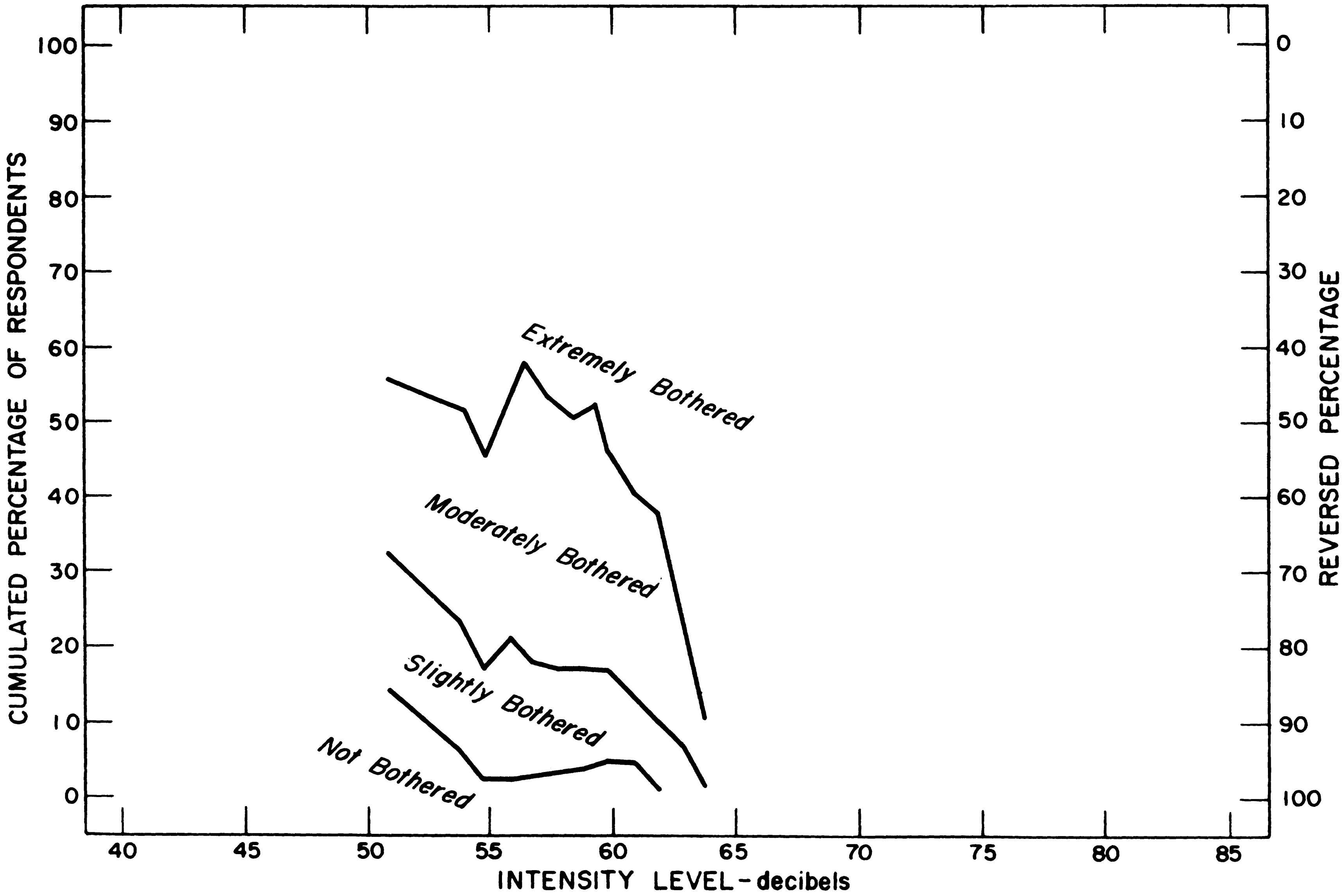


Figure 44. White noise; source in living room, unfiltered.

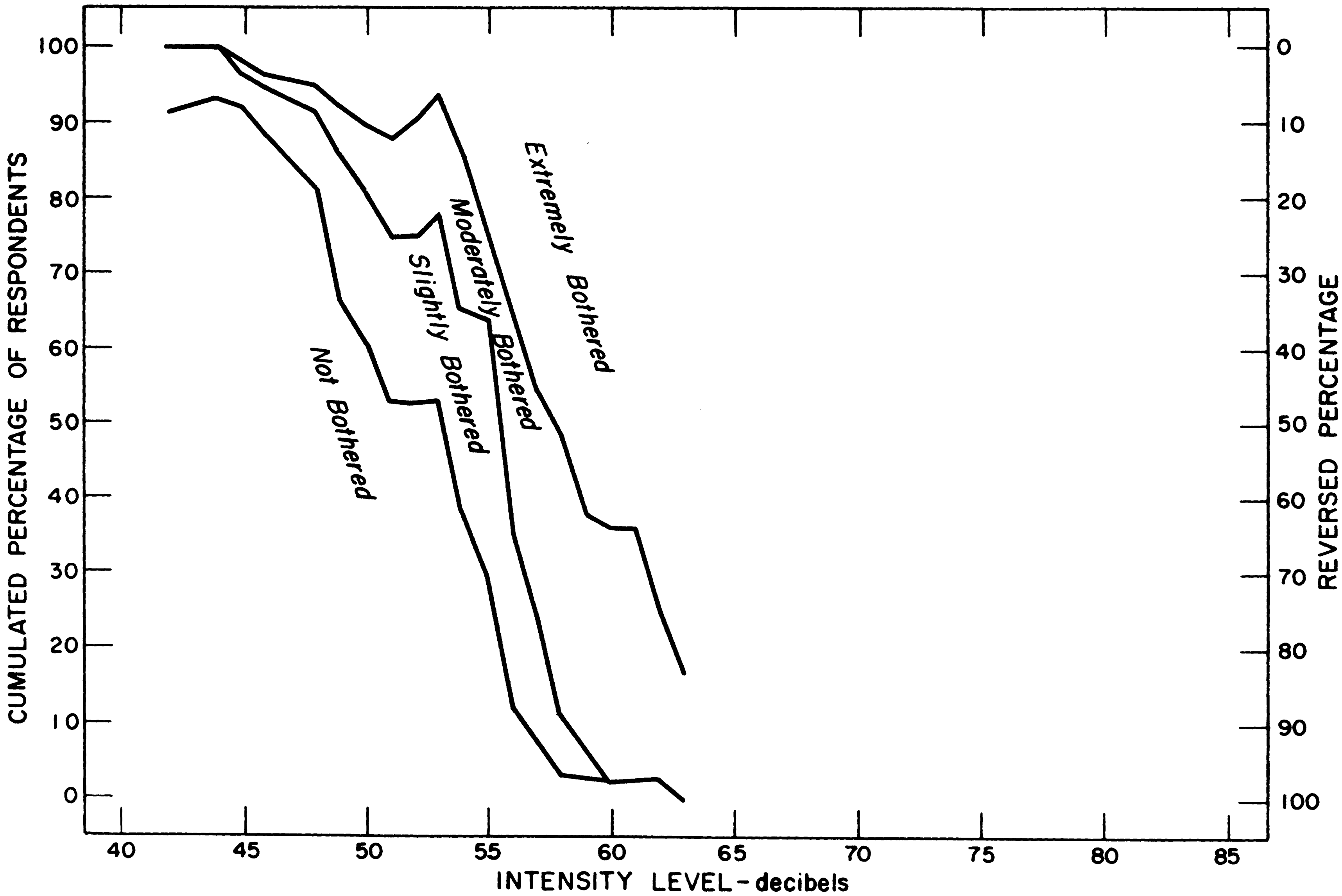


Figure 45. White noise; source in kitchen.

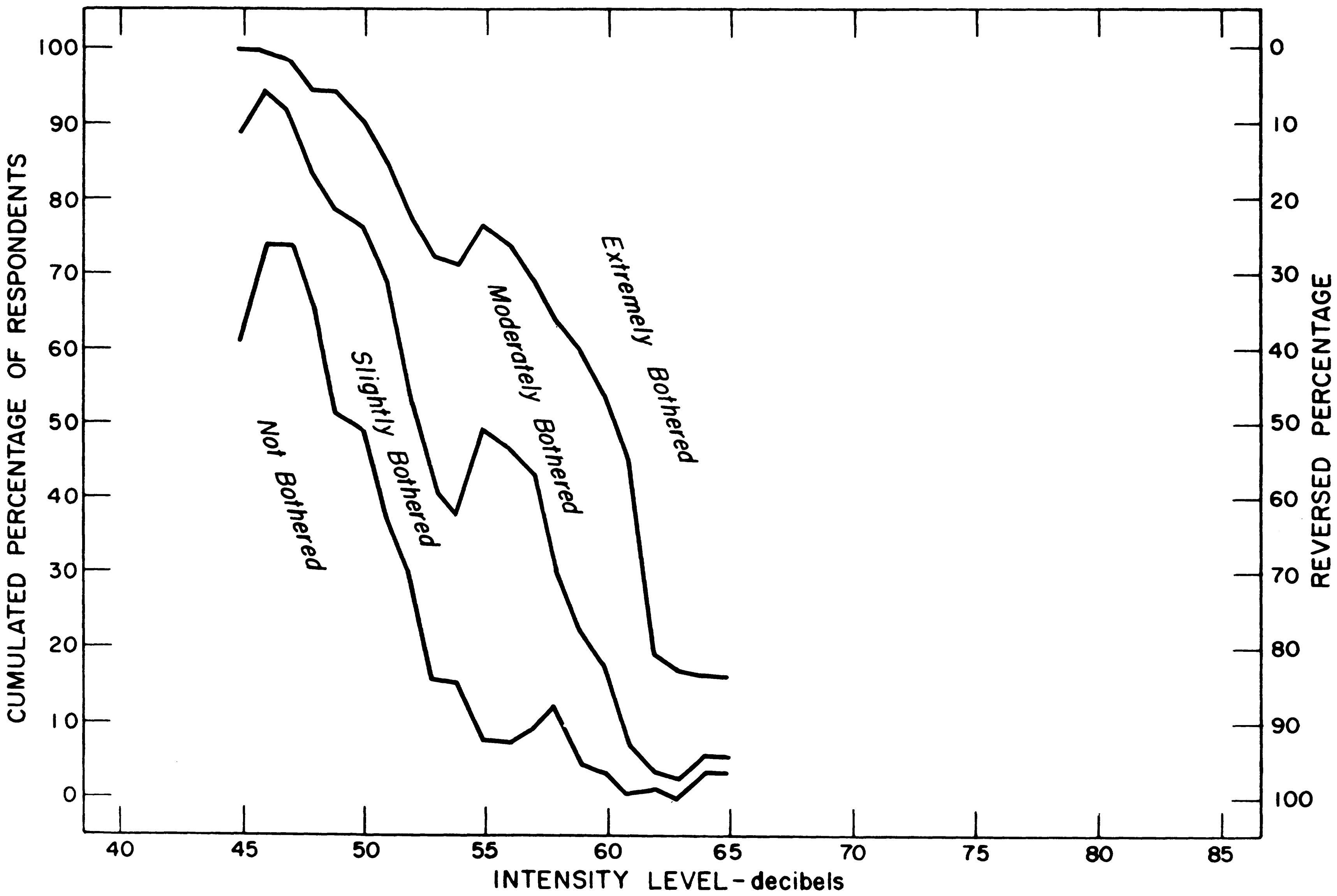


Figure 46. White noise; source in living room, filtered.

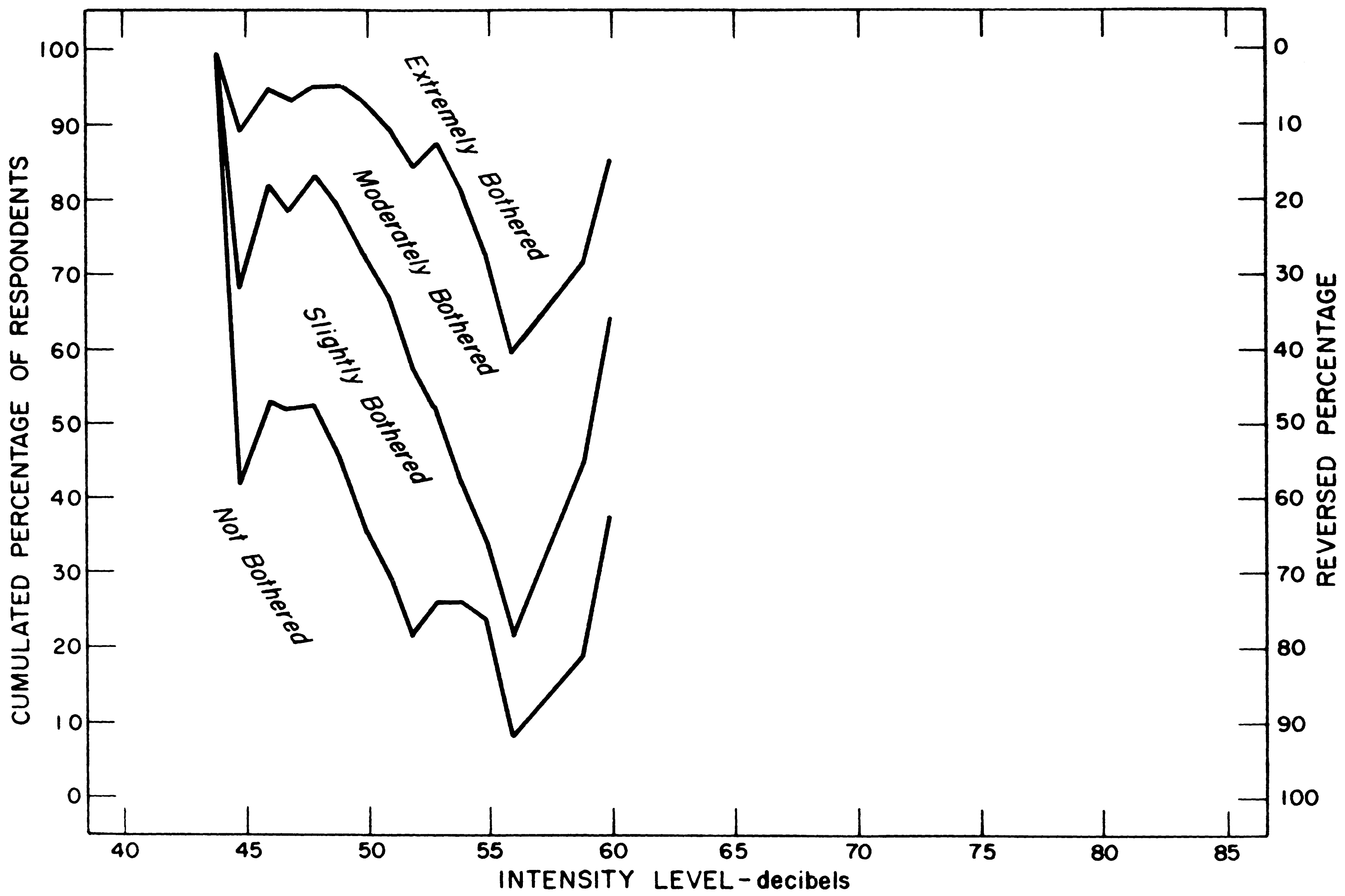


Figure 47. White noise; source in basement.

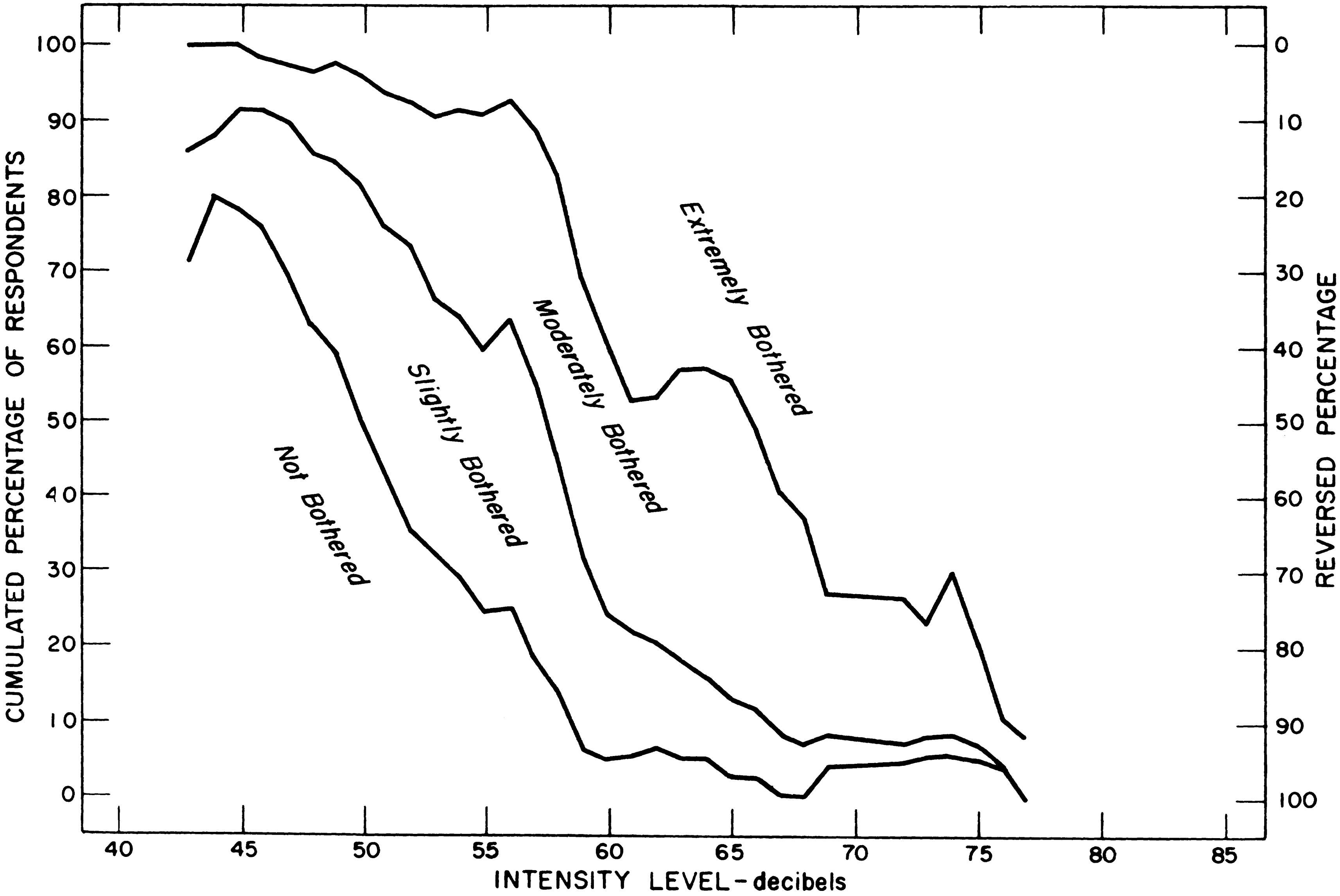


Figure 48. Standard band of noise; all source locations merged.

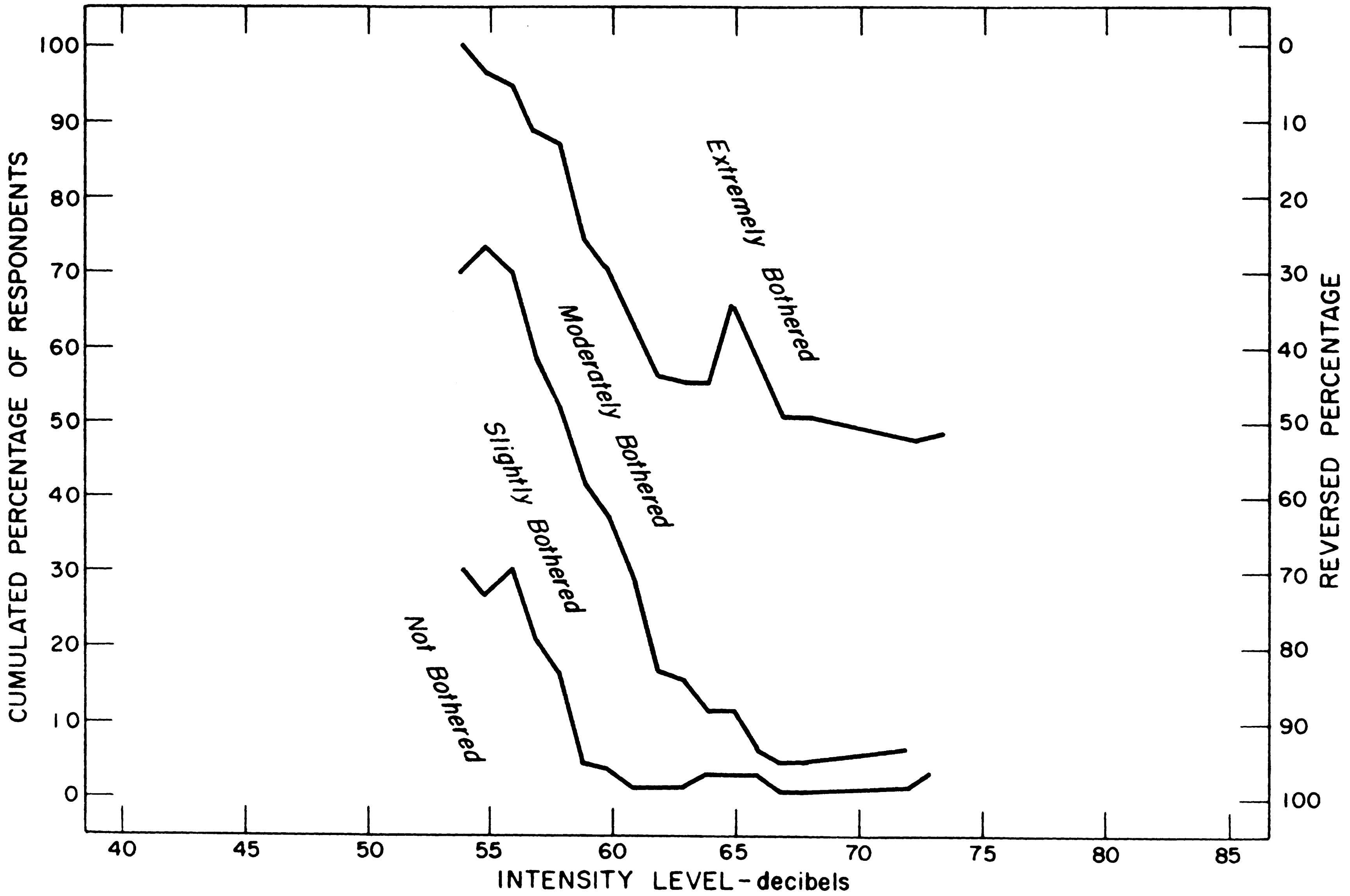


Figure 49. Standard band of noise; source in living room, unfiltered.

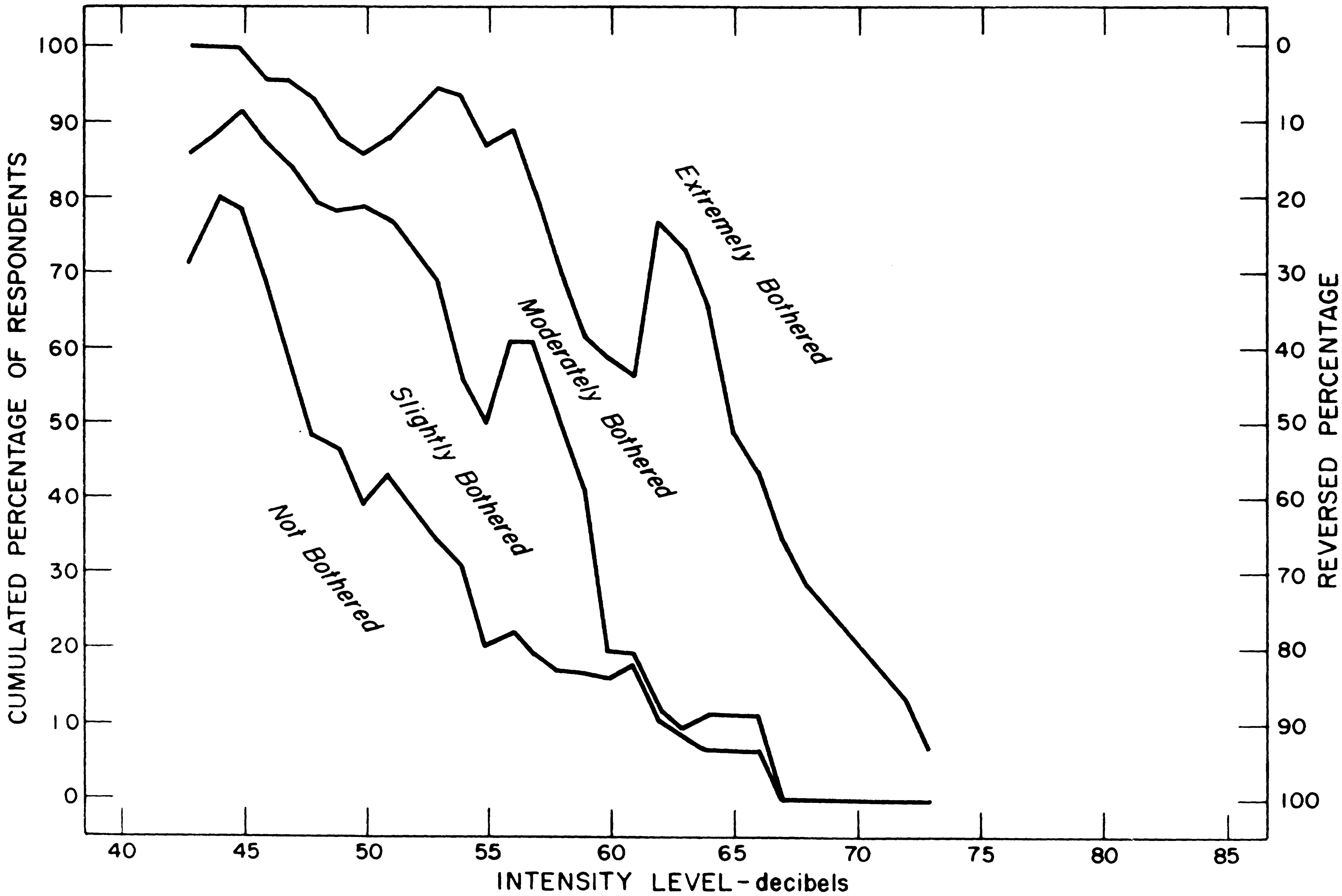


Figure 50. Standard band of noise; source in kitchen.

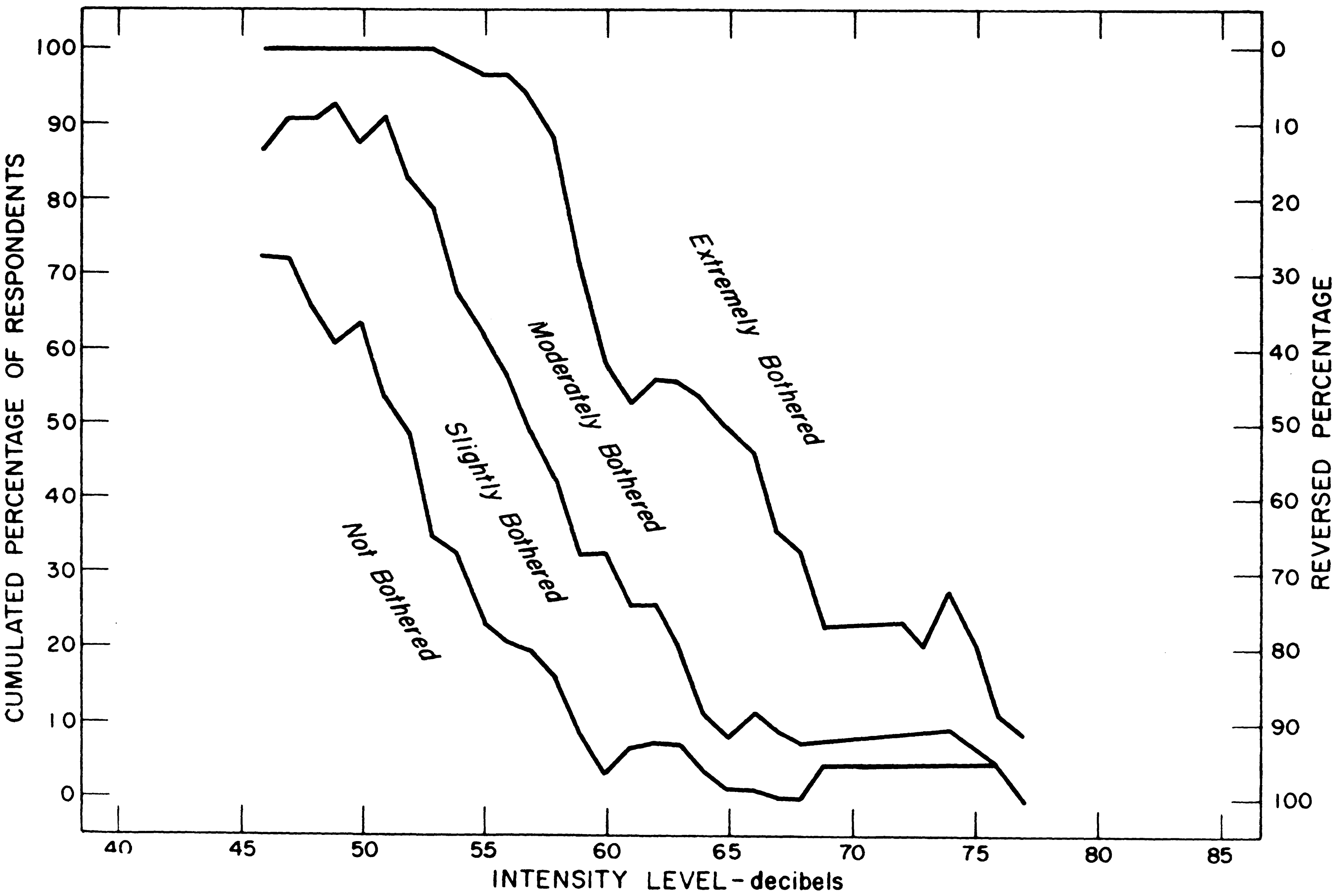


Figure 51. Standard band of noise; source in living room, filtered.

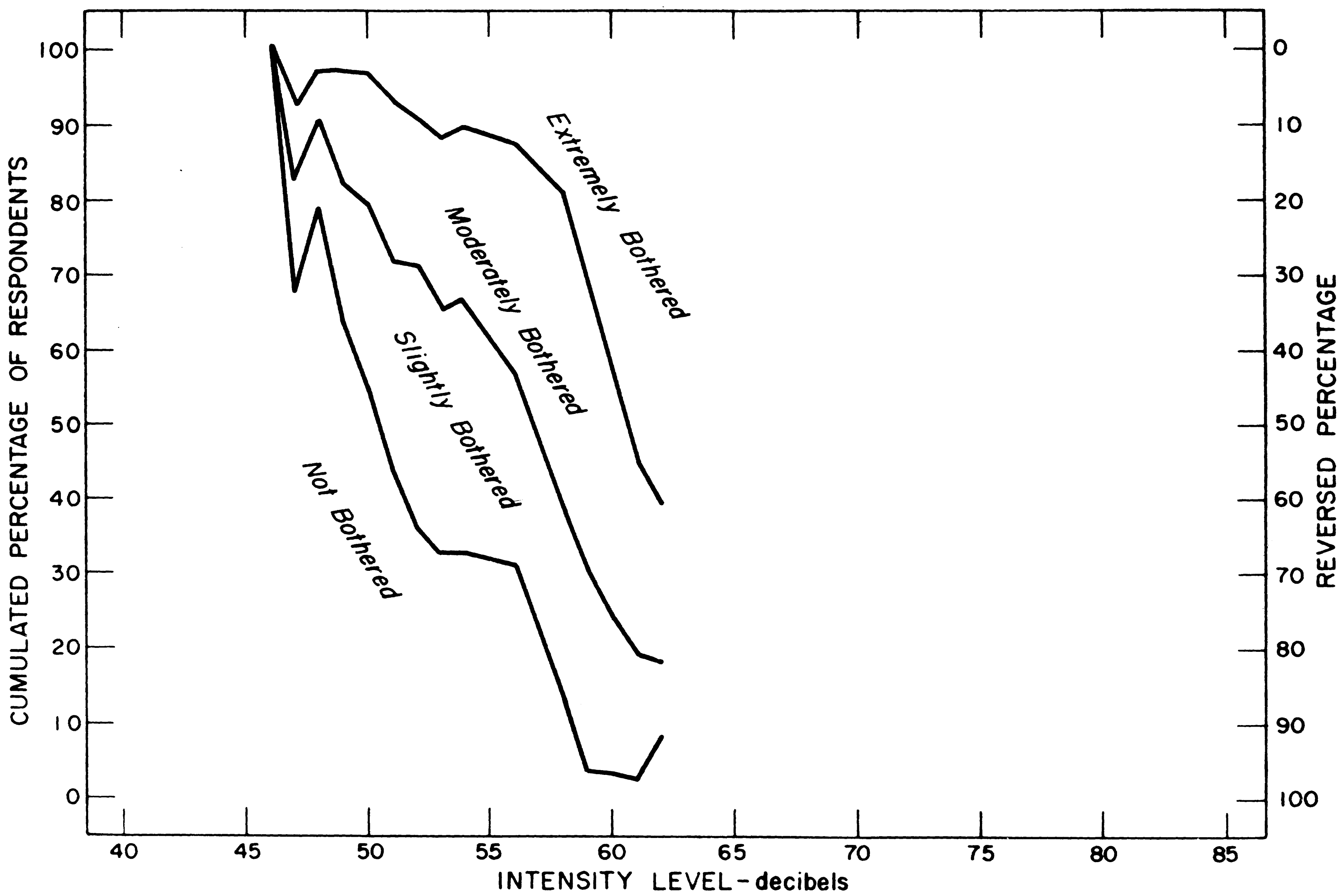


Figure 52. Standard band of noise; source in basement.

APPENDIX C

STATISTICAL APPENDIX
(Correlation, Regression and Fiducial Limits, and Graphical Methods)

In this report extensive use is made of regression and correlation methods. Standard statistical texts discuss the usual applications of these methods, and there is no intent to duplicate that information here. Some of the salient features as they apply to the applications here will be mentioned, and some unique uses discussed. Of course, for full interpretation of the materials it would be desirable to consult texts and/or experts.

"Regression" is the misleading term applied to constructing the straight line that best approximates the functional relation between two variables. The regression of y on x (e.g., the regression of degree of bother on intensity level in db) has several meanings. The straight line described by the regression equation (ordinarily using the slope-intercept equation of a straight line) comes as close as possible (in the least-squares sense) to the mean degree of bother at each level in db. The mean of regression line (or the value of y at some specified value of x) represents the overall level of bothersomeness, while the slope of the line represents the rate of change of bothersomeness with change in dbs.

The coefficient of correlation is a single number that expresses how closely the regression line describes the data. The correlation coefficient ranges in size from zero to one. If all data points fell exactly on the regression line, the correlation would be one; if the regression is no closer to the points than the mean score is, then the correlation is zero.

Since correlation and regression are "least-square" statistical techniques, it is also possible to make interpretation in terms of variance (i.e., the squared deviation) accounted for by various influences. For example, the square of the correlation coefficient is the proportion of variance in the dependent variable accounted for by the independent variable.

Multiple regression and multiple correlation are techniques in which a single dependent variable is predicted from a linearly weighted combination of a number of dependent variables. Coefficients are interpreted in the same fashion; for example, a multiple correlation of .7 between three variables and the bothersomeness judgment indicates that .7 x .7 = .49 or about 1/2 of the variance in bothersomeness is accounted for by the linear combination of the three variables. The "stepwise multiple regression" program used in some cases checks every possible variable at each step, selectively picking out those variables that account for the maximum amount of remaining variance at each step.

Each of these methods is a technique of descriptive statistics; that is, each is a data reduction technique that reduces the information in the data to manageable form. A correlation coefficient is a single number summarizing the (linear) relation between two variables; a regression formula uses two parameters and multiple regression, many parameters to describe the data. We can ask questions that go beyond merely summarizing the data—we can generalize beyond the data. This is the subject matter of inferential statistics. Examples of the uses of inferential statistics in this study include tests of whether observed correlation coefficients might have occurred by chance in sampling from a population with a correlation of zero; setting probability confidence limits (fiducial limits) around regression lines to see if different sounds differ significantly or only by chance in the amount of bothersomeness they cause; and using a chi-squared test of association to test whether the cross-tabulation of two variables differs significantly from what is expected on the basis of marginal frequencies.

One rather special application of standard statistical techniques involves the use of Canonical Correlation methods to construct a new metric for an independent and dependent variable—the new metric relation being such as to maximize the correlation between the two variables.[5] Thus, with the appropriate stretching and shrinking of the two scales indicated in Figure 53, we can transform Figure 53 into Figure 54 which indicates a perfect linear relation.

Of course, the scales for the axes in Figure 54 are not the same as the scales in Figure 53. For example, on the Y-axis it was necessary to expand the scale in the 6-7 region and to contract the scale in the 1-5 region in order to bring the points into a straight line. With real data, it is not usually possible to make the relation perfect, but a maximum linear correlation is achieved.

Such rescaling of the metric can be especially valuable in cases where there is no a priori metric—the Canonical Correlation method creates a scale which is optimal in the sense of maximizing the linear regression. In our study we used this method to scale the judgment of "bother" made by our subjects. We knew that "not bothered," "a little bothered," "moderately bothered" and "extremely bothered" were in ascending order, but had no prior basis for assessing the size of the successive steps on the scale.

Another special technique used (which, like the Lingoes Metric, is too recent to appear in standard textbooks) is the Automatic Interaction Detector (AID), developed by Morgan and Sonquist of the Institute for Social Research, University of Michigan. This technique is an analog of stepwise-multiple-regression in that it seeks out and uses those variables which successively account for the maximum remaining variance. However, AID, as its middle name suggests, seeks these variables in an interactive fashion, rather than combining all variables by linear weight into a single equation. Thus, it might turn out, if loud sounds vs quiet sounds (divided at 65 db) was the first best prediction of bothersomeness, that for quiet sounds the extent of hearing loss is the second best predictor, while for loud sounds the psychoacoustic sensitivity is the second best predictor, and so on.

In spite of the large number of complex and sophisticated statistical techniques used, it may still be true that the most efficient statistical device is the familiar one of graphic presentation. This report includes a large number of graphs

---

[5]Devised by Professor James Lingoes, University of Michigan Computing Center; therefore referred to as a "Lingoes Metric."

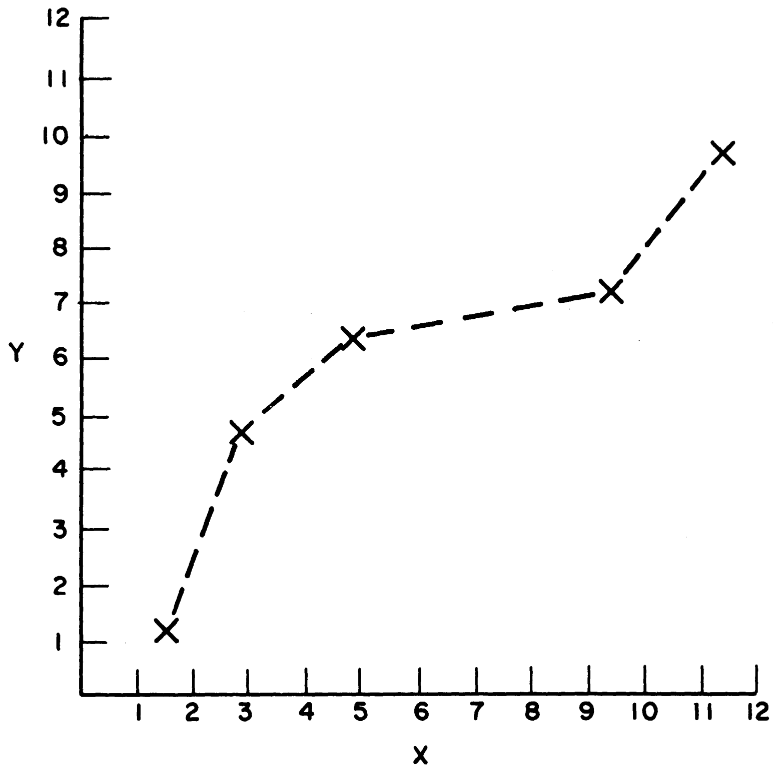


Figure 53. Hypothetical graph.

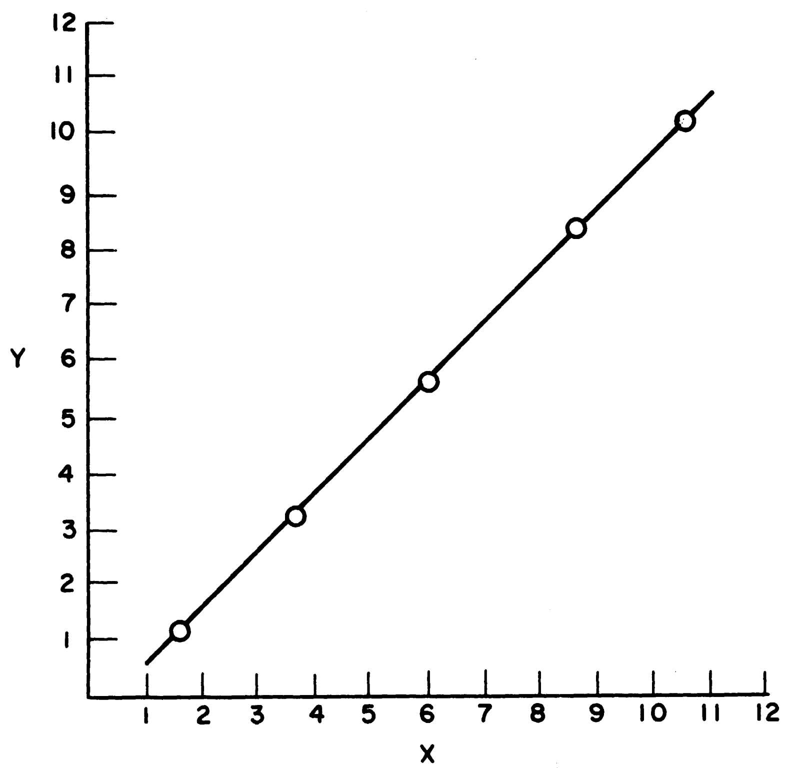


Figure 54. Rescaled hypothetical graph.

(Appendix B) which show the cumulated percentage of our respondents (or some sub-group of our respondents) who are "not bothered," "bothered a little," "moderately bothered," or "extremely bothered" as a function of the intensity level (or loudness level) of the sound. These graphs give directly the information discussed in the earliest proposal to Whirlpool; namely, a quantitative statement of the difference between sounds in terms of consumer subjective response.

Does the engineering department want to know the effect of increasing the dryer noise by 5 db? Then look at the dryer graph, displace it 5 db to the right, and check the subjective reaction at several critical points—e.g., typical levels for nearby, adjacent room, and basement-to-living room.

Does the product development group have in mind a new washing machine with a white noise sound spectrum? Use the white noise graphs to determine the appropriate intensity level for the desired freedom from complaint regarding noise.

Does the sales department want to know whether superior hearing or subjectively judged noise sensitivity is a better index of receptivity to a line of high quality products advertised as quiet? Check the graphs for the appropriate subgroups for significant differences in response.

REFERENCES

1. "Proposal for Research on Appliance Noise," ORA-61-943-B1, Institute of Science and Technology, The University of Michigan, Ann Arbor, Michigan, March 31, 1961.

2. N. E. Barnett, B. E. Erickson, and R. Hefner, "Review of Appliance Noise Problem," and including, G. Katona and J. Schmiedeskamp, "Complaints About Noise in Household Appliances," Institute of Science and Technology Report No. 4790-1-T, The University of Michigan, Ann Arbor, Michigan, July 1963. (Issued in draft form April 1962.)

3. N. E. Barnett, B. E. Erickson, and R. Hefner, "Appliance Noise Study, Second Technical Report," Institute of Science and Technology Report No. 4790-2-T, The University of Michigan, Ann Arbor, Michigan, March 1963.

4. R. A. Hefner, Jr., and S. J. Schultz, "Appliance Noise Study, Third Technical Report," Institute of Science and Technology Report No. 4790-3-T, The University of Michigan, Ann Arbor, Michigan, May 1963.

5. W. Spieth, "Annoyance Threshold Judgments of Bands of Noise," J. Acoust. Soc. Am., vol. 28, pp. 872-877 (1956).

6. H. Fletcher and W. A. Munson, "Loudness, its definition, measurement and calculation," J. Acoust. Soc. Am., vol. 5, p. 82 (1933).

7. D. W. Robinson and R. S. Dadson, "A Re-Determination of the Equal-Loudness Relations for Pure Tones," Brit. J. Appl. Phys., vol. 7, p. 166 (1956).

8. U. S. Census of Population and Housing 1960, Census tracts, Ann Arbor, Michigan, U. S. Dept. of Commerce, Luther H. Hodge, Secretary.

9. Dept. of Commerce, Bureau of the Census, p. 756, Table 1065.

10. Reference 2 above, Table II, page 13.

11. S. S. Stevens, "Calculation of the Loudness of Complex Noise," J. Acoust. Soc. Am., vol. 28, p. 807 (1956).

12. S. J. Rule and J. W. Little, "Importance of Instructional Set to Psycho-acoustical Scaling," J. Acoust. Soc. Am., vol. 34, p. 1974(A), (1962).

13. "Guides to the Evaluation of Permanent Impairment; Ear, Nose, Throat, and Related Structures," by The Committee on Medical Rating of Physical Impairment, J. Am. Medical Assoc., vol. 177, pp. 489-501, August 19, 1961.