COGNITIVE PSYCHOLOGICAL APPROACHES IN THE EVALUATION OF INFORMATION EXCHANGE PROCESSES

Marie van Melis-Wright, Mary Batcher, Deborah Stone, & Fritz Scheuren
Internal Revenue Service & Bureau of Labor Statistics
441 G Street, N.W., Washington, D.C., 20212

A woman seeking a divorce went to visit her attorney. The first question he asked her was, "Do you have grounds?". She replied, "Yes, about two acres". "Perhaps I am not making myself clear," he said, "do you have a grudge?" "No, but we have a carport," she responded. "Let me try again." "Does your husband beat you up?" he said impatiently. "No, generally I get up before he does," she said. At this point, the attorney decided to try a different tack. "Ma'am, are you sure you really want a divorce?" "I don't want one at all, but my husband does. He claims we are having trouble communicating." (Pat Swindall, in Wurman, p. 103, 1989).

Asking and answering questions is one of the most common human activities (e.g., Dillon 1982). In fact, with a few exceptions (e.g., formal lectures, speeches, etc.), question and answer sequences can be identified in most oral communications. Such interaction patterns, which require at minimum two participants, serve both sociological and psychological functions in that they facilitate the exchange of many types of information, foster social bonding, and, in many social contexts, monitor group and individual behaviors.

This study focuses on a special type of information exchange situation, namely, telephone conversations--a unique form of oral communication in that it is devoid of visual non-verbal cues (Farb 1973)--conducted to obtain/confirm domain-specific factual knowledge, which takes place between two individuals who "know" each other only in terms of their functional relationship, specifically a U.S. taxpayer and an Internal Revenue Service (IRS) taxpayer assister.

In this paper, we first describe the IRS Integrated Test Call Survey System or ITCSS, which forms the background leading to the study. Second, we propose a provisional model of a special case of taxpayer-tax assister exchange to illustrate some of the complexities and unique aspects characterizing this type of dialogue. Third, we discuss theoretical and pragmatic reasons for the particular variations that were investigated in this study. This is followed by a discussion of the study design and methodology. Finally, we conclude with a presentation of our findings and outline future research plans.

BACKGROUND

The Internal Revenue Service has offered taxpayers free telephone assistance on income tax related matters since 1965. At present, this service is available at 31 toll-free telephone sites located throughout the United States. The volume of inquiries is staggering; for example, in 1989, the IRS received approximately 50 million calls. In 1990, between January and mid-April alone, IRS assisters responded to 17.5 million taxpayer requests for information (Batcher & Scheuren 1989; 1990). This large call volume necessitates continued monitoring of the quality (e.g., accuracy and completeness) and type of information exchanged, as mistakes could impact significantly on individuals' lives and society as a whole.

In response to public and Congressional requests for an assessment of the accuracy of information given by the toll-free assistance program, the IRS designed the Integrated Test Call Survey System in 1987. Briefly, the ITCSS consists of a series of approximately 1300 weekly calls made by eight trained test callers. The questions used by the test callers cover twenty different categories of tax law and can be answered by the assister using the IRS publication 17 or the Form 1040 instructions. During the conversation with the tax-assisters, the test callers are required to record the presence or absence of certain key information. Actual scoring, which occurs later, is designed to assess whether particular response and probe combinations are present (for a detailed discussion of the ITCSS program, see Batcher and Scheuren 1989).

While the program is largely successful, the IRS has always been and remains sensitive to its limitations (e.g., Jabine 1989; Scheuren 1989). Of particular concern is the degree of artificiality and rigidity of the test calls (necessitated by measurement requirements) relative to the fluidity that is more characteristic of "real" taxpayer/tax assister interactions. For example, while the content of test calls and taxpayer inquiries overlap significantly in terms of tax law categories, the ITCSS test questions are strictly fixed in: wording; placement in the conversation (up front); amount of background information embedded in the question; presentation; type and amount of background information available in response to tax assister probes; and required (or prohibited) test caller behaviors. It is unknown to what extent this imposed structure influences the process and outcome (particularly in terms of information accuracy) of the interaction.

To address some of these concerns, the IRS commissioned the Bureau of Labor Statistics' Collection Procedures Research Laboratory (CPRL) to conduct a number of studies. This study is the first of a series designed to shed light on taxpayers' interactions with the IRS. The laboratory, housed within the Office of Research and Evaluation, has as its major mission to assist the members of various government agencies in the design and execution of surveys, conduct research aimed at improving particular surveys and survey technology in general, and conduct basic research on topics related to information exchange processes. The CPRL's continuing aim is to develop a well-integrated theory of how people ask and

answer questions, with the ultimate goal to improve both theoretical understanding of the survey process and the practical applications of such knowledge.

ITCSS INFORMATION EXCHANGE MODEL

For a period of four months test caller/tax assister exchanges were monitored on a regular basis and subsequently were analyzed for conversational structure. Figure 1 presents a provisional and simplified model of the typical test caller/ tax assister exchange that emerged from this analysis. While it fairly accurately reflects the various paths of this interaction, the authors have strong reasons to believe that a model of all possible real life taxpayer/tax assister interactions will be much more complex. Nevertheless, the figure 1 model delineates the various tasks faced by both parties, identifies some of the major components, and will allow us to indicate where we believe it deviates from real life taxpayer/tax assister interactions. We will now briefly describe some of the salient aspects of this interaction. Space limitations and the particular focus of the paper force us to keep the description of the model brief, and therefore, incomplete.

As figure 1 shows, a question regarding tax matters is presented by the test caller to the assister. The ITCSS test questions are carefully designed with respect to content, wording, length, amount of embedded information, etc., to maximize clarity, and to minimize assister confusion. (For an extensive review of the test question development process, see Batcher and Collins 1989). For the purpose of illustrating the process we will, at this point, assume that the question is comprehended by the assister. The authors are, of course, aware that there is considerable research demonstrating that comprehension processes often fail, even when question quality is close to optimal (e.g., Cannell, Miller and Oksenberg 1981; Dippo and Herrmann 1990).

After the question has been posed and assuming the assister has comprehended the question at least to some degree (box 1), there are a number of options. In most cases, the assister will (and generally should) probe for further relevant information (box 2). Within this context, probing is defined as exploratory questioning designed to establish whether or not the caller's situation meets certain conditions. The probing is systematic, in that the expected probes should reflect the assister's prior training, knowledge, and experience (Craik and Tulving 1975) regarding applicable areas of tax law. For example, a question about whether or not to file should elicit probes regarding age, marital status, amount of income, etc.

Once the probing has been completed, the assister again has several choices. Given that a number of conditions are met (e.g., the question has been comprehended, caller's responses to probes have yielded satisfactory background information, and the assister believes he or she has or can find the appropriate information in Publication 17), an answer to the original question will follow (box 7). In some cases however, factors such as uncertainty about the tax law on the part of the assister and/or a real or perceived challenge by the caller, may elicit a referral to a more highly trained technical back-up person (box 8). Finally, situations occur under which the assister may opt for an offer of a write-up (box 9) For example, the assister may not be able to provide an answer because of lack of training, back-up assisters may not be available, etc. In this case, the caller's question is recorded, and a written answer will be sent to the caller at a later time.

Still assuming that the question is clear to the assister there are, of course, a number of other possible paths. One is to skip the probing process completely and to provide an immediate answer (from box 1 to box 3). This is appropriate, however, only for certain types of questions, e.g., Where do I get a Form 1040? Another option the assister may take, is to transfer the caller to a technical back-up person (i.e., from box 1 to box 4), or the assister may elect to offer the caller a write-up (box 1 to box 5). We believe the reasons for choosing one of these options are essentially similar to the ones quoted above.

Finally, the assister may not comprehend the caller's question (box 10). In this case, the assister may attempt to get clarification from the caller (box 11). For example, he or she will ask the caller to repeat the question, to elaborate, or to rephrase. (Note however, that within the context of the ITCSS program, test callers can repeat, but are prohibited from rephrasing the questions). In such instances, the most common approaches taken by assisters are to transfer the call to a technical back-up or to request the caller to put the question in writing and mail it to the IRS (box 12).

As was noted earlier, the IRS is quite aware that the information exchange between real taxpayers and assisters is more complex than the model described above and has commissioned the Bureau of Labor Statistics' CPRL to conduct a study designed to map structure, content, and flow of these interactions. This study is just beginning, but already we have uncovered some fascinating results. Most striking is the incredible variation in taxpayers' question formulations. For example, a preliminary analysis of transcribed (anonymous) protocols includes well-formulated questions phrased as statements, incomplete or extremely vague inquiries, questions which are, in fact, multiple questions, etc. Some possible variations in the way in which taxpayers may ask questions are presented in figure 2. This figure contains a series of fictitious information exchanges that illustrate the variety of possible opening statements by the taxpayer upon initial contact. This series of five examples, although altered for the sake of confidentiality are, nevertheless, unaltered in structure. Also noteworthy, but not surprising, is that taxpayers' questions frequently are not posed right at the beginning of the conversation. Other points worth noting are taxpayers' tendency: to spontaneously paraphrase answers given by assisters, to ask for clarification when answers seem unclear, and, on occasion, to challenge the assister's information. Gathering systematic knowledge about patterns characterizing this kind of interactions is crucial, as it will allow us to eventually align the ITCSS program closer in content, process, and structure to what takes place in real life.

RATIONALE FOR THE STUDY

The current study was designed in response to IRS concerns about the potential effects of various factors (e.g., question length or placement, embedded background information, etc.) on assisters' responses to test questions. During the development stage of the ITCSS program, opinions and beliefs about the effects and direction of such effects frequently influenced question construction and also defined the parameters of allowable test caller responses, yet none of these beliefs have been formally tested.

Although a certain amount of artificiality is inherent in the test call situation, understanding the differences in the structure of test caller/assister and taxpayer/assister interactions may allow the construction of test questions

that more closely mimic real-life conversations. Other potential benefits are improved training of assisters and the capability to detail more effective question procedures for taxpayers to use when making their telephone inquiries.

Out of a number of research ideas generated, two were selected for this pilot study. The first addresses the potential effects of asking the assister to spell his or her name (ostensibly, for record keeping purposes by the caller, although no names were written down). The second involves having the caller ask the assister to answer the question in a step-by-step fashion. In both cases, the dependent variables were accuracy and completeness of the answers and the assister's probing behaviors.

The request for the assister's name was included for two reasons. First, under IRS regulations, if the examination of a tax return reveals a mistake, fines will generally be waived when the error is a result of incorrect information given by the assister (although the taxpayer remains responsible for taxes owed). As stated in the Form 1040 however, this rule only holds if the taxpayer is able to provide the assister's name. Second, field workers have voiced a suspicion that there may be individuals who seem to be shopping around for the wrong answer, perhaps for the purpose described above. Tax assisters are familiar with this rule, and the authors believe the explicit request for the assister's name may "trigger" this knowledge. Under this condition then, one may see greater accuracy, an increased tendency to transfer to the technical back-up, or a rise in offers to the caller for write-ups.

The decision to include the caller request for an answer in a step-by step fashion was based on research from the field of education. For example, elaboration of facts tends to increase knowledge acquisition and enhances product quality (e.g. Hendrickson 1980; Stein, Littlefield and Bransford 1984). Similarly, Dillon (1982) notes identifying relationships between objects, events, etc. may result in increased comprehension.

Within the context of the ITCSS, we hypothesized requests for step-by-step instructions might stimulate the assister to apply tax law information to the specific situation of the caller. This would not only improve accuracy, but also increase the degree to which the answer is helpful to the caller. This manipulation is deemed of particular importance in light of the fact that some proportion of tax assisters rountinely cite tax rules taken verbatim from IRS publications. While such an answer (reading the tax law to the taxpayer from IRS publications) is technically accurate, there may be callers who lack the ability to apply that information to their particular situation.

Additional goals of this pilot study were (1) to test experimental procedures, (2) to identify the feasibility of quantifying hitherto not considered response variables (e.g., length of call, changes in interpersonal behaviors, etc.), and (3) to obtain some estimate of effect size in order to determine sample sizes for later studies.

METHODOLOGY

The experimental treatments were structured using a test question that had been dropped from the ITCSS question pool. The question read:

In 1987, I bought an old painting. I paid to have it framed. Then I sold it last year (1989) and got $1,500 for it. What do I have to show on my return?

Three variations of the question were developed to reduce the risk that the question would be identified as a test question. The only difference between the versions was the object under consideration (i.e., old painting versus antique lamp versus antique table).

The experimental calls were placed by the permanent staff of eight test callers. These callers had already been fully trained on test call procedures within the context of the ITCSS program. That training included role playing, with callers playing the part of assister and test caller; the development of probing skills, to help identify and respond to probes for background information; tax law training to allow them to distinguish the various forms that responses can take; and training in the mechanics of the system.

Although the regular test call operation is computerized, the special experimental testing was a paper and pencil operation. However, the test questions and response sheets were structured to resemble the written version of the questions that the callers were accustomed to using in conjunction with the computer version. Except for absence of the computer assisted system, we asked the test callers to follow the procedures used during the ITCSS calls. As with the regular calls, callers were provided with background information to allow them to respond to probes from the assisters. They were also instructed to follow the script but to try to be as natural as possible in tone.

There were, of course, unanticipated questions from assisters which required the test callers to use their judgment in their response. Researchers met frequently with the test callers as a group to discuss these unanticipated probes and ensure a unified response. These meetings also allowed the test callers to provide feedback and suggestions for improving later experimentation.

Testing occurred during a three-week period in late April and early May. Calls were assigned evenly across test callers, weeks, and time of day (morning or afternoon), but included no weekend calls.

Treatment Conditions

Two treatment conditions were studied. In the first the test caller requested the assister's name using the question:

Would you please spell your name for me, I would like to write it down in my notes. (pause, as if writing down the assister's name, but DO NOT WRITE IT DOWN).

In the second condition, the caller requested step-by-step presentation of the answer using the question:

Well, I find this very difficult. Could you please walk me through the steps so I know exactly what to do?

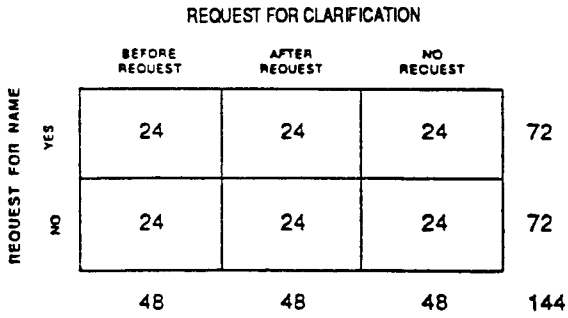
The request for name occurred at two levels, present or absent; the request for step-by-step presentation occurred at three levels: request for clarification up-front in the conversation, request following an answer given by the assister, and no request.

Blocking Conditions

Although there are several potential blocking variables, the sample size allowed only for one blocking variable--site accuracy as measured by the 1990 ITCSS. The experimental conditions were replicated in high and low accuracy site groupings.

Experimental Design

The two factors were completely crossed and fully replicated within blocks for a 2 by 3 design with blocking at 2 levels. Twenty-four calls were scheduled for each cell of the design, for a total of 144 calls.



RESULTS

We have completed preliminary analyses, focusing on two dependent variables, the **proportion of correct responses** and the **length of time** needed to complete the call.

Percent Correct--A chi-square test for independence was used to test whether the proportion of correct responses differed by the experimental conditions, **request for name** and **request for clarification.** The observed chi-square results were 0.25 with 1 degree of freedom for the request for name and 1.17 with 2 degrees of freedom for the request for clarification. Neither of these is significant.

Length of Call--An analysis of variance was performed on the logarithm of the length of call. The results are given below; only the request for clarification was significant. This should not be a surprising result since such a request might be expected to increase the length of the call.

Analysis of Variance for Length of Call

Source	df	Sum of Squares	Mean Square	F-ratio	Prob
Name	1	0.002524	0.002524	0.050	0.8229
Clarification	2	0.395830	0.197915	3.95	0.0221
Error	110	5.51571	0.050143		
Total	113	5.92191			

CONCLUSIONS

The statistical results observed were, at best, inconclusive. This may be a reflection of small sample size, as the callers were not able to place calls for as many weeks as originally planned. In addition, because of the pilot nature of this study, the procedures and questions used were tested minimally in advance and probably introduced a large amount of noise into the data.

Probably the most useful information we obtained was on the question we used, the scoring sheet, caller procedures and other procedural matters. Based on informal feedback from the test callers, the question was too form-specific. In most cases, the request for further information brought a line-by-line response and sometimes some confusion if the assister had already provided a line-by-line response but then was asked to step the caller through the response.

The question may have been too complex as well. Transfers to technical referral were common, occurring in 52 percent of the calls, and, while technical backup is a part of the system, generally less than a third of taxpayer calls are transferred.

Other procedural components generally worked out well, but small improvement should be realized, based on this pilot.

Another area where we can improve our process is to identify more sensitive outcome measures. Although the accuracy of a response is important, it provides minimal information about what actually goes on in these conversations.

We expect to conduct further testing with these and other experimental conditions, building on what we have learned from this small pilot to do a series of larger, better structured experiments on effect of the way we ask questions on accuracy and other indicators.

NOTE: We gratefully acknowledge the Internal Revenue Services' Taxpayer Services' managers and test callers for their cooperation in making the test calls. We also want to thank Doug Raybeck and Douglas Herrmann for their review of our initial draft and for their thoughtful suggestions and comments.

BIBLIOGRAPHY

Batcher, M. (1989). Integrated test call survey system: Statistical working note. Series 1. Memorandum in Scheuren, F. 1989 Integrated Test Call Survey System. Washington, D.C.: Government Printing office.

Batcher & Collins (eds.). (1989). The 1989 Integrated Test Call Survey System. Volume I: Design and Development., Volume II: Implementation. Statistics of Income Division. The Internal Revenue Service. Washington, D.C.: Government Printing Office.

Batcher, M. & Scheuren, F. (1989). The IRS Test Call Program. Paper presented at the 1989 Annual Meeting of the American Statistical Association in Washington, D.C.

Batcher, M. & Scheuren F. (1990). The IRS Test Call Program: Design and Estimation Issues. Paper presented at the 1990 Annual Meeting of the American Statistical Association in Anaheim, California.

Cannell, C.F., Miller, P.V., & Oksenberg, L. (1981). Research on Interviewing Techniques. ISR. University of Michigan.

Craik, F.I.M. & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. Journal of Experimental Psychology, 104, 268-294.

Dillon, J.T. (1982). The effect of questions in education and other enterprises. Journal of Curriculum Studies, 14(2), 127-152.

Dillon, J.T. (1982). The multidisciplinary study of questioning. American Educational Research Journal, 19(4), 540-551.

Dippo, C.S. & Herrmann, D. (1990). The Bureau of Labor Statistics' Collection Procedures Research Laboratory: Accomplishments and Future Directions. Paper presented at COPAS, May 21.

Farb, P. (1974). Word Play: What Happens When People Talk. New York: Bantam Books.

Good, T.L.; Slavings, R.L.; Harel, K.M., & Emerson, H. (1987). Student passivity: A study of question asking in K-12 classrooms. Sociology of Education, 60, 181-199.

Groves, R.M., et al. (eds.). (1988). Telephone Survey Methodolgy. New York: John Wiley & Sons.

Hendrickson. (1980). Procedures and results of an evaluation of writing. Educational Evaluation and Policy Analysis, 2, 19-30.

Jabine, T. (1989). In Batcher & Collins (eds.). 1989 Integrated Test Call Survey System. Volume II: Implementation. Washington, D.C.: Government Printing Office.

Leech, G. (1981). Semantics: The Study of Meaning. 2nd edition. New York: Penguin Books.

Searle, J.R. (1969). Speech Acts: An Assay in the Philosophy of Language. Cambridge: University of Cambridge Press.

Slobin, D.I. (1979). Psycholinguistics. 2nd edition. Glenview, Illinois: Foresman & Co.

Stein, B.S.; Littlefield, J., & Bransford, J.D. (1984). Elaboration and knowledge acquistion. Memory & Cognition, 12(5), 522-529.

Wurman, R.S. (1989). Information Anxiety. New York: Doubleday.

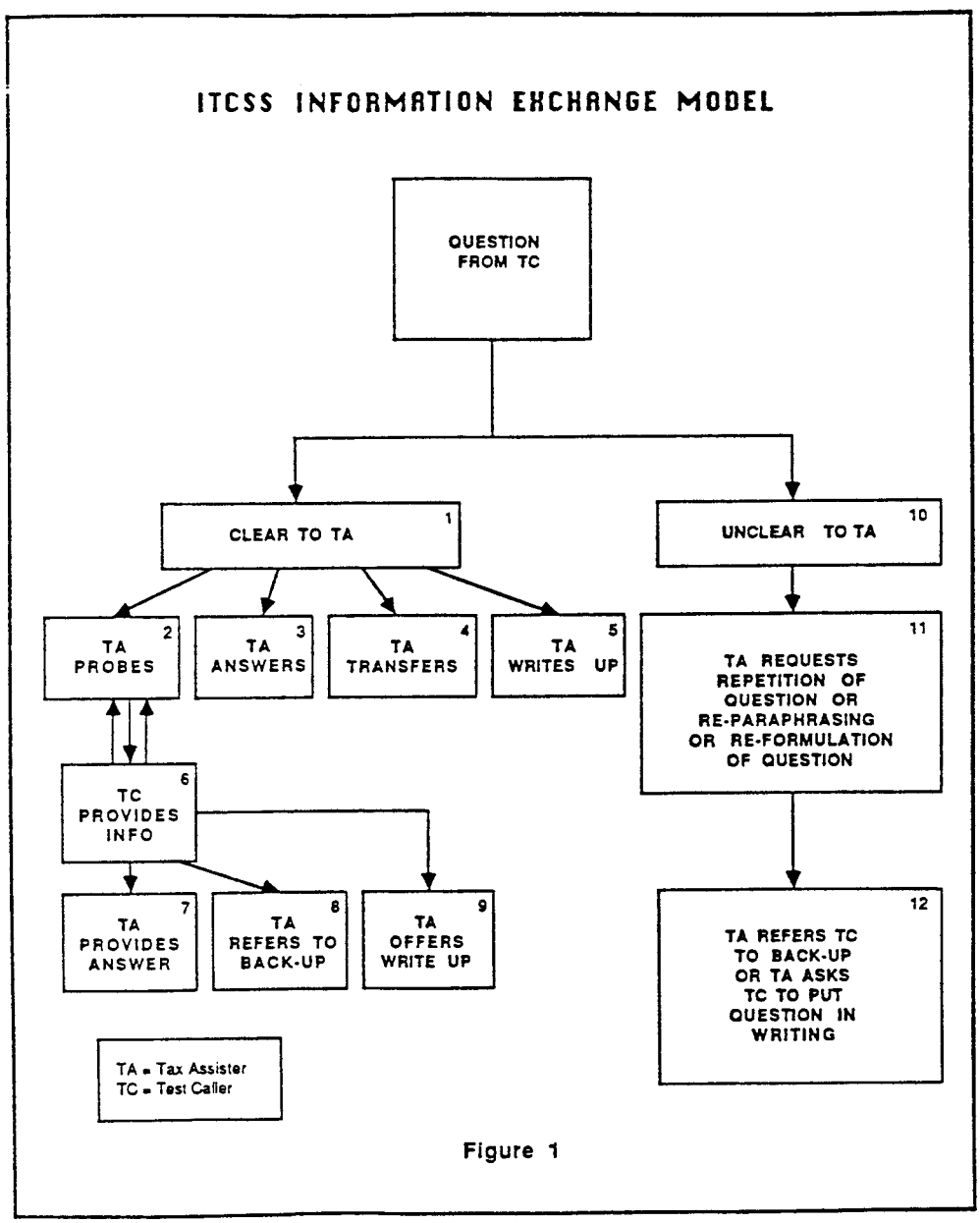


Figure 1

TEST CALLER/TAX ASSISTER INITIAL INFORMATION EXCHANGE

TA: Internal Revenue Service, _____ speaking. May I help you?
TP: My husband and I are fixing to file a return and we will be moving March 30th. What address should I put on my return?
TA: What date will you be moving?
TP: March 30th, assuming everything goes well.
TA: You're sure?
TP: Well, if everything goes well.

TA: Internal Revenue, Taxpayer Service. This is _____. How may I help you?
TP: Well, this is _____. I have a problem with my taxes. I haven't filed for '84 and '85 and I got a phone call this morning that says I owe a bunch of money that I don't think I owe.
TA: Who called you?
TP: Somebody in _____.
TA: Okay.

TA: Internal Revenue Service, _____ speaking. How may I help you?
TP: I need to figure out a little bit about bankruptcy . . . how that is treated as far as taxes go and if it is tax deductible . . . What form do I use?
TA: What do you mean deductible?
TP: The legal fees I paid.

TA: Internal Revenue, Taxpayer Service. This is _____. How may I help you?
TP: Do you have a number for the Forms Division. I get a recording on that number that it has been disconnected.
TA: So you need a telephone number of the Form Distribution Site?
TP: Yes, I need some forms.

TA: Internal Revenue Service, _____. How may I help you?
TP: Yes, I was talking to my neighbor, and she said my little girl probably had to pay taxes on her savings account. If a little 3 year old girl has a savings account, and she has a social security number, will she need to pay income tax on that ?
TA: Is she your dependent?
TP: Yes. See it was a college fund set up by her uncle.

Figure 2