THE IMPACT OF DESIGN ON INTERACTION IN THE COMPUTER ASSISTED SURVEY INTERVIEW

Sue Ellen Hansen
Survey Research Center, University of Michigan
426 Thompson Street, Ann Arbor, MI 48106-1248

Key Words: CAI, Design

Introduction[1]

Although computers have been used to conduct surveys since 1971 (Couper and Nicholls, 1998), only recently have researchers begun to focus on user centered design (UCD), that is, design from the computer user's perspective (see, for example, Sperry *et al.*, 1998). For face-to-face surveys this acknowledges the role of the interviewer as intermediary between the respondent and the computer, and the potential impact of the computer on interviewer-respondent interaction. In order to understand how the computer may affect interaction, it is necessary to understand how computer assisted personal interviewing (CAPI) is similar to and different from paper and pencil personal interviewing (PAPI).

An important difference in interaction between computer assisted and paper interviews is that a computer assisted interview may be viewed as involving *two* interactions: one between the interviewer and respondent, and one between the interviewer and the computer (Couper, 1997). The interviewer must attend to and manage her interactions with both respondent and computer. Each interaction may be influenced by features of the context of the interview, such as whether by telephone or face-to-face, as well as many other related factors and proximate activities. In addition, each interaction may influence the other.

Standardization of the survey interview is generally viewed as necessary to minimize measurement error attributable to the interviewer (Fowler and Mangione, 1990). Whether computer assisted or paper and pencil, the standardized survey interview involves the administration of an instrument comprised mainly of a schedule of scripted questions. It has been argued that standardization may not provide adequate resources for interviewers to recover from interactional difficulties with the respondent, and thereby threatens the validity of survey responses (Suchman and Jordan, 1990). Primarily to reinforce standardization but also in part to help the interviewer avoid interactional problems, most survey instruments provide interviewer instructions, definitions, and other information meant to guide the interviewer through the question-answer process. Success of implementation of these features can be expected to vary, affected by factors such as design, interviewer training and experience, and supervision.

The standardized survey instrument then is a tool, to which the interviewer refers during interaction with the respondent and on which she records the respondent's answers. Although the use of paper and pencil may allow more ecological mobility, giving the interviewer more flexibility on where and how to conduct the interview, both computer assisted and paper interviews place interviewers in a *tool saturated* environment, with laptops or paper instruments, supplements, show card booklets, laminated calendars, and so on. Such an environment places demands on interviewer attention that may affect interaction with the respondent (see Greatbatch *et al.*, 1993, for parallel similarities and differences in paper and computer assisted medical consultations).

The introduction of a computer in computer assisted interviewing (CAI) may further influence interaction. A paper questionnaire allows an interviewer to flip back and forth among pages of the instrument, to see more than one question at a time, and to write anything she needs to anywhere on the questionnaire. In contrast, the computer assisted instrument controls the delivery of questions and flow through the questionnaire, segments the questionnaire by displaying only one screen at a time, and restricts options for the recording of responses (Groves, Berry, and Mathiowetz, 1980). If programmed to tailor instructions and questions to reflect previous answers, to conditionally display probes, or to check the content or consistency of recorded responses, the computer also may influence the asking of questions, and react directly to interviewer actions and indirectly to respondent answers.

Interactional difficulties, that is, problems of shared resources and understanding, or common ground, may occur both in interviewer-respondent interaction (Clark and Schober, 1992) and in interviewer-computer interaction (Brennan, 1998). In addition, variation in interviewer behavior within and across modes may reflect differences in the interactional substrate, that is, differences in the requirements of the question-answer process in CAPI and PAPI standardized survey interviews, and the ways different interviewers handle them (Schaeffer and Maynard, 1996).

Thus, standardized interviewing, instrument design, and computer assistance all have a potential impact on interviewer-respondent interaction and its outcome, which may be positive, neutral, or negative, and which may vary

[1]This research was supported through funding from the National Center for Health Statistics, under Cooperative Agreement #S278-15/15 with the Survey Research Center (James M. Lepkowski, Principal Investigator).

throughout the interview. The actual impact may depend on how easy or difficult each makes it for interviewers to elicit from respondents at each question the information sought by the researcher. It is widely recognized that question wording and order may contribute to measurement error (see Sudman, Bradburn, and Schwarz, 1996, for an overview). It has also been argued that format and layout may affect the response process in self-administered paper interviews (Jenkins and Dillman, 1997). Less attention has been paid to the impact of instrument design in interviewer administered surveys, which involve interviewer-respondent interaction (for an exception see Sanchez, 1992). CAI raises further design issues, since it also involves interviewer-computer interaction. This draws interviewer attention away from the respondent and her responses, and may influence interviewer-respondent interaction, data quality, and the efficiency with which data are collected.

Research on human-computer interaction (HCI) suggests that in computer assisted interviews ease of use or usability may be determined in large part by the design of the computer interface (Couper, 1997). This includes the layout and format of information displayed on the computer screen (Tullis, 1997), the availability, clarity, and degree of consistency of design features and system functions (Shneiderman, 1992), and the ways the computer communicates through commands and instructions, and feedback following user actions (Brennan, 1998).

HCI focuses on the cognitive and interactional processes involved in a person's interaction with a computer (Carroll, 1997). It forces extension of traditional models of interaction in the survey interview (Schaeffer, 1991) that focus primarily either on question structure, social psychological aspects of interaction, or respondent cognitive processes. HCI suggests that a model of interaction in the computer assisted interview needs also to account for the influence of the computer through the design of CAI software and survey instruments.

The remainder of this paper presents and discusses findings from a comparison of interaction in CAPI and PAPI interviews, and their implications for survey measurement. Analysis of coded interviewer behavior from videotaped interviews identified similarities and differences in interviewer behavior in the two modes.

Data and Methods

The primary source of data for this study is a set of 52 videotaped laboratory interviews, averaging 50 minutes in length. Thirty-eight of the interviews were CAPI and 14 were PAPI. All interviews took place as part of an evaluation of the 1997 CAPI instrument of the National Health Interview Survey (NHIS), conducted for the National Center of Health Statistics (NCHS). The 1996 paper NHIS instrument was used for the PAPI interviews. Nine U.S. Census Bureau interviewers each conducted two to six CAPI and one or two PAPI interviews during the Spring and Summer of 1997. All interviewers were experienced in both PAPI and CAPI NHIS interviewing. Respondents were recruited from the local community and were paid for their participation.

Interaction Event Codes. Each CAPI screen or PAPI item in an interview was coded to indicate the occurrence of specific events, such as not performing required tasks, data entry problems, and muttering. Table 1 lists the event codes. The codes were designed to capture interviewer behavior related to the use of the survey instrument or the computer, in order to identify usability problems related to instrument design (cf. Dumas and Redish, 1994). Although they may also capture events related to interviewer-respondent interaction, they differ from traditional behavior codes (Cannell, Miller and Oksenberg, 1981), which capture behaviors involved in the question-response process for evaluation of interviewer performance and/or the survey questions themselves.

Table 1. Interaction Event Codes

Code	Description
C	Task-related comment
E	Error or data entry problem
F	Filler word or phrase
L	Laughter
M	Muttering or self-talk
N	Task not performed
S	Complete interviewer silence
T	Problem reading text

Every item encountered in an interview has one record in the data set with nine codes, each set either to "0" if the event did not occur at all, or to "1" if the event occurred at least once. There may be multiple records for an item, either because an item was asked of multiple persons in a household, or because an interviewer revisited items during the process of backing up to review or change prior answers.

As with traditional behavior coding, such event coding is subject to reliability problems, due to variation among coders in their understanding and consistent application of the codes. To check the reliability of the coding, five interviews (1,418 items) were double coded during an early phase of coding, and kappa statistics were calculated (Fleiss, 1981). Agreement on the presence or absence of an event at each item ranged from 84.4% to 99.4%, and kappas ranged from .41 to .73, suggesting moderate reliability of the codes and confidence that the proportions of events obtained are not due to chance alone.

The unit of analysis is the CAPI screen or PAPI questionnaire item, generally referred to as an item. This is typically a single question and space for recording an

answer. Each represents an interactional exchange at which events may occur. The final data set has 11,258 CAPI items and 3,765 PAPI items.

Standardized scores were computed for each question in order to identify items that may pose difficulties for interviewers (cf. Lepkowski *et al.*, 1998). They were calculated as the difference between the percentage p_i for an event at an item and the mean of percentages $\bar{p}$ for the event across all items, divided by the standard deviation σ_p of percentages for the event across all items. Those items with standardized scores of 2.0 or greater on any of the eight events were selected for further examination. This process produced lists of 49 unique CAPI items and 44 unique PAPI items. Some items had high percentages of more than one event.

Item Characteristics. There are 455 unique CAPI screens and 238 unique PAPI items. It was hypothesized that certain interactional difficulties may be associated with particular screen or item characteristics. For example, an interviewer may be more likely to stumble on a question with a name fill, to use filler words when orienting to a screen with several instructions, or to mutter or be silent at an interviewer checkpoint, which does not require interviewer-respondent interaction. For this reason, each item was coded for characteristics such as question and input types, features of question text, and interviewer instructions. Table 2 summarizes the question characteristic codes.

Table 2. Item Characteristics

Category	Characteristics
Question/code text	Emphasis, dates, name fills, parentheses, slashes
Response/input	Interviewer checkpoint, introduction, open or fixed, multiple input, multiple response
Instructions	Hand card, verify, next question, data entry
Other	Help, context information, household roster

Findings

To test the null hypothesis that the proportions of events in PAPI and CAPI interviews were equal, z-scores for the differences in proportions were calculated for each event (Fleiss, 1981). Between the two modes there are small and non-significant differences in proportions for task- or affect-related comments, filler words, laughter, and errors or backing up in the instrument. Interviewers were more likely to be silent (X^2=233.48) or mutter (X^2=11.66) while working with the PAPI instrument, whereas they were more likely in CAPI *not* to ask an explicit question or complete a task as required by explicit instructions in the instrument (X^2=45.91).

Muttering and Silence. The higher proportions of muttering and silence in PAPI interviews may reflect the complexity of the PAPI instrument, which has many more interviewer checkpoints, items that require only interviewer data entry and are often used to determine the text fills and routing through the questionnaire. At such items, interviewers tend to be silent, or to mutter. The latter may reflect the interviewer's self-talk while processing information, or may be a means of filling conversational gaps. Of the 3,765 total PAPI items, 1,057 (28.1%) are interviewer checkpoints, 550 at which interviewers were silent (52%) and 98 at which they muttered (9.3%). In contrast, of the 11,258 total CAPI items, 662 (5.9%) are interviewer checkpoints, 370 (55.9%) and 134 (20.2%) of which involve silence and muttering.

By programming most checkpoints and controlling skip patterns in the questionnaire, CAPI appears to have greatly minimized problems interviewers may have with the complex paper NHIS instrument. However, CAPI has not yet eliminated them. That 5.9% of items encountered in these CAPI interviews are checkpoints, and that they have a higher incidence of muttering (20.2% versus 9.3% in PAPI interviews) suggests that the NHIS CAPI checkpoints could be evaluated further to see if any can be eliminated through additional programming, or improved if they must remain in the instrument.

That interviewers in both modes tend to be silent at checkpoints is not surprising. In fact, all PAPI items at which there was an unusually high incidence of silence are checkpoints, as are many of the CAPI items. In order to try to isolate other factors that may increase silence and muttering, proportions of non-checkpoint items were compared (n=10,596 in CAPI, 2,708 in). Without the checkpoint items, there are significant differences in percentages of filler words, muttering, tasks not performed, and silence. PAPI items had higher percentages of filler words (X^2=64.78) and muttering (X^2=4.51); CAPI items had higher percentages of tasks not performed (X^2=13.90) and complete silences (X^2=8.35). In CAPI the higher proportion of silence may be associated with the increase in tasks not performed, which may include questions not asked as required. That muttering is still high in PAPI on non-checkpoint items, may reflect the overall complexity of the PAPI instrument and that interviewers must manage skip instructions, frequently moving between the main questionnaire and information in the household roster and complex checkpoints.

Summarized in Table 3, these differences are not easily explained without more sophisticated analyses, which are beyond the scope of this paper. However, combined with the fact that there is more silence and muttering at checkpoint items in CAPI interviews, they suggest that some differences in interaction may reflect the influence of the computer and/or the design of the computer assisted instrument.

Tasks Not Performed. Of all CAPI items 13.6% (979) were coded as having some kind of task not

performed (N). This code indicates the interviewer did not show the respondent a response options card, did not read a question or an introduction to a question or question series, or did not follow some other explicit instruction. Of the items coded N, 56.9% (557) had hand card instructions. Of the 348 PAPI items coded N, less than 1% (17) had such instructions.

Table 3. Summary of Events in which CAPI Differed Significantly from PAPI

Compared to PAPI	Including Checkpoints*
CAPI overall had	Less muttering Less silence More tasks not performed More problems reading text
but at checkpoints had	More silence and muttering

*5.9% of CAPI and 28.1% of PAPI items were checkpoints.

Examination of the 47 unique CAPI hand card items coded N reveals that four items with the highest number of code Ns are items asked of every or nearly every person in the household, and together represent 35.5% of the 557 instances of hand card items coded N. They are relationship to reference person, race, national origin, and education. Event coding identified one of these, relationship to reference person, as problematic with 68% coded N. As already noted, interviewers may tend not to refer to a hand card on repeated visits to these questions. This could be appropriate when items repeated are in sequence, as with education, where the interviewer asks level of education about six people on six separate but contiguous screens. However, it is not appropriate for the other three questions, repeated as part of a series of questions asked of each respondent, so that repeated items are not contiguous and other hand cards may have been referred to in between (as with relationship to reference person).

The much lower number of code Ns on hand card items in PAPI may reflect the fact that race and national origin are presented by topic, not by person as in CAPI, so that interviewers obtain national origin of all members at the same time, requiring one reference to the hand card, and then move on to the race question. It also likely reflects the fact that in PAPI two of these items, relationship to reference person and education, do not have corresponding hand cards.

That PAPI does not have hand cards for these questions may increase interviewers' tendency not to use the hand cards, since all interviewers in the study had interviewed with the NHIS PAPI instrument prior to the transition to CAPI, and they were not accustomed to referring to hand cards on these items. In the case of the relationship to reference person item, the unusually high percentage of code Ns may indicate that interviewers either do not think it necessary to use a hand card or do not need to ask the question. Respondents generally would not need to see a list of response options in order to report the relationship of another person in the household, and interviewers would often have obtained the information earlier in the household listing question series.

Another factor contributing to interviewers' tendency not to refer to hand cards in CAPI may be placement of hand card instructions. Review of the CAPI items shows that while most hand card instructions appear at the top of the screen, prior to question text, some appear below, and sometimes with other instructions. This may cause interviewers sometimes not to see the instruction and thus not refer to the hand card.

In CAPI interviews, of the non-hand card questions coded N because the interviewer did not perform a required task, such as asking a question or verifying information, many were questions at which the interviewer may have felt the required task was at conflict with conversational norms. For example, nine of the 15 times an interviewer was required to ask the "other name" a household member went by, it is likely that the she already was given the information when she asked if the household member went by another name (for example, "Does, Ryan Smith go by another first name", "Yes, Mike", making the question "What is this other first name" an awkward followup).

Approximately half of the CAPI items at which a task was not performed asked for information about a family member in the household, such as relationship to reference person. As with the first name followup, the interviewer may not ask such questions because she already has the information. For example, she may have been told "we're all Puerto Rican," after asking the national origin followup question about the third member of a six-person household. However, these screens appear separately for each person. In a six-person household, for instance, the race question appears as a one of several items for each household member, and relationship to reference person appears for five of them. Even when it is necessary to verify information for every person, presenting separate questions for each makes the question-response process tediously redundant for both interviewer and respondent, and interviewing less efficient. It also may lead interviewers to not ask some of these items, reducing standardization.

Filler Words. There are large percentages of the use of filler words in both modes, although it is significantly higher at non-checkpoint items in PAPI. Both modes require some cognitive processing on the part of the interviewer between questions. In CAPI this involves orientation to a new screen, and interpretation of what to do at the screen. In PAPI the interviewer has to determine

where to go next, take time to get there, and then interpret what to do at the new item. In both modes of interviewing there are similar percentages of filler words at checkpoint items (21.1% in CAPI and 23% in PAPI). Filler words in either mode may be uttered automatically, as part of a transition from item to item, as the interviewer finishes recording a response to one item and moves on to the next, or may fill conversational space as the interviewer tries to figure out what to do at the next item. The significantly lower occurrence of filler phrases at non-checkpoint items in CAPI (44.8% versus 53.5%) may be due to the fact that the computer controls skip patterns in CAPI, which overall should place a lower cognitive burden on the interviewer. Although beyond the scope of this paper, detailed examination of available transcripts from interviews in both modes, which would reveal the location of filler words in interaction, as well as pauses and lengths of pauses in conversation, could clarify the ways in which interviewers use such discourse markers in the two modes of interviewing.

Problems Reading Text. CAPI shows significantly fewer problems reading question text, but this difference reverses when checkpoint items are excluded. This suggests that the original negative difference is due to the large PAPI percentage of checkpoints, which do not have text to read. Once checkpoints are removed, CAPI has more text reading problems, although the difference is no longer significant. The higher percentage may be due to the use in CAPI of upper case for both name fills and emphasis in question text.

As expected and as these findings show, in both modes there is a correlation between the question characteristic checkpoint and the event silence, stronger in PAPI (r=.59) than in CAPI (r=.47). There were no other characteristic-event correlations this high. In CAPI, there was a slight correlation between errors and comments (r=.22), suggesting interviewers may comment about what they're doing when they are correcting entries or backing up. There was also a slight correlation (r=.28) between a hand card instruction on an item and not performing a task, and some of the reasons for this have been discussed. In PAPI there is a slight correlation (r=.24) between open questions and the use of filler words. This might be related to the interviewer having to move back and forth between other portions and the top right portion of the questionnaire, where open items such as names, relationships, and health conditions are recorded. She may utter something like "Okay..." to fill a conversational gap as she moves between locations or tries to identify where to record the response.

Discussion

Some interviewer behavior does not differ between paper and computer assisted standardized interviews. Regardless of mode, instruments and interviewer-respondent interaction is sometimes complex, leading to events beyond basic questions and answers, such as laughter, interviewer comments about her tasks, and the editing of responses. The interviewer must attend to both the survey instrument and her interaction with the respondent, and the cognitive demands are great, as demonstrated her tendency in both modes to fill conversational space as she moves between the two, comprehending task requirements and interpreting and recording respondent answers.

However, there are some significant differences in interviewer behavior between the two modes. The most striking difference is the substantially higher proportion of tasks not performed in CAPI interviews, suggesting less standardization. This appears to contradict a widely held belief that CAI generally leads to greater standardization (de Leeuw and Nicholls, 1996) and prior research that found a reduction in interviewer variance with CAI, at least for computer assisted telephone interviewing (Groves and Mathiowetz, 1984). One explanation may be that when interviewers are faced with what they view as unnecessary standardization, or standardization harmful to interaction with the respondent, they choose not to standardize, although it also may be that Census Bureau interviewers use less standardization than others.

That muttering and silence occur more frequently at CAPI interviewer checkpoints, in spite of the fact that there are fewer checkpoints in CAPI, suggests that CAI instrument design has an impact on interaction.

Event coding identified many features of the design of NHIS CAPI instrument that led to tasks not being performed. However, event coding is limited in what it can reveal. Like behavior coding, it can detect that a problem occurred, but not why it occurred. Examination of transcripts of interaction in the interview would show the conversational context in which problems occur. This can provide greater insight into the nature of interactional difficulties and guidance in providing solutions. Although not as rich a source of information, automatically generated data such as trace files of interviewer and CAPI system would provide additional information on the interviewer's interaction with the computer. Both transcripts and trace files for the interviews in this study are available for future analyses.

Design *does* matter. In CAI, that includes design of the computer interface, and instrument design considerations that go beyond the structure, layout and format of questions. Standardization, instrument design, and computer assistance all may affect measurement. Further analysis and additional research is necessary to understand the nature and extent of the impact of computer assistance on interaction in the computer assisted standardized survey interview, and to provide guidelines for design of computer assisted instruments. This involves understanding the requirements of the two

interactions the interviewer manages, and the resources she uses or could use to manage them.

References

Brennan, S.E. (1998). The Grounding Problem in Conversations with and through Computers. In S.R. Fussell and R.J. Kreuz (eds.), *Social and Cognitive Approaches to Interpersonal Communication.* Mahwah, New Jersey: Laurence Erlbaum Associates.

Cannell, C.F., P.V. Miller, and L.F. Oksenberg (1981). Research on Interviewing Techniques. In S. Leinhardt (ed.), *Sociological Methodology* vol. 2. Reading, Massachusetts: Addison-Wesley Publishing Company.

Carroll, J.M. (1997). Human-Computer Interaction: Psychology As a Science of Design. *Annual Review of Psychology 48*:61-83.

Clark, H.H. and M. F. Schober (1992). Asking Questions and Influencing Answers. In J.M. Tanur (ed.), *Questions About Questions: Inquiries into the Cognitive Bases of Surveys.* New York: Russell Sage Foundation.

Couper, M.P. (1997). The Application of Cognitive Science to Computer Assisted Interviewing. Paper presented at the CASM II Seminar, Charlottsville, Virginia, June.

Couper, M.P. and W.L. Nicholls (1998). The History and Development of Computer Assisted Survey Information Collection Methods. In M.P. Couper *et al.* (eds.), *Computer Assisted Survey Information Collection.* New York: John Wiley & Sons, Inc.

de Leeuw, E. and W.L. Nicholls (1996). Technological Innovations in Data Collection: Acceptance, Data Quality and Costs. *Sociological Research Online 1*(4).

Dumas, J.S., and J. Redish (1994). *A Practical Guide to Usability Testing.* Norwood, N.J: Ablex Publishing Corp.

Fleiss, J.L. (1981). *Statistical Methods for Rates and Proportions.* Second edition. New York: John Wiley & Sons, Inc.

Fowler, F.J. and T.W. Mangione (1990). *Standardized Survey Interviewing: Minimizing Interviewer-Related Error.* Newbury Park, CA: Sage Publications.

Greatbatch, D., *et al.* (1993). Interpersonal Communication and Human-Computer Interaction: An Examination of the Use of Computers in Medical Examinations. *Interacting with Computers 5*(2):193-216.

Groves, R.M., M. Berry, and N. Mathiowetz (1980). Some Impacts of Computer Assisted Telephone Interviewing on Survey Methods. Pp. 519-524 in *Proceedings of the Section on Survey Research Methods.* Alexandria, VA: American Statistical Association.

Groves, R.M. and N. Mathiowetz (1984). Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents. *Public Opinion Quarterly 48*:356-69.

Jenkins, C.R. and D.A. Dillman (1997). Towards a Theory of Self-Administered Questionnaire Design. In L. Lyberg *et al.* (eds.), *Survey Measurement and Process Quality.* New York: John Wiley & Sons, Inc.

Lepkowski, J.M., *et al.* (l998). Instrument Evaluation, Behavior Coding, Trace Files, and Usability Testing. Paper presented at the Annual Meeting of the American Association of Public Opinion Research, St. Louis, May.

Sanchez, M.E. (1992). Effects of Questionnaire Design on the Quality of Survey Data. *Public Opinion Quarterly 56*:206-17.

Schaeffer, N.C. (1991). Conversation With a Purpose--or Conversation? Interaction in the Standardized Interview. In P. Biemer *et al.* (eds), *Measurement Errors in Surveys.* New York: John Wiley & Sons, Inc.

Schaeffer, N.C. and D.W. Maynard (1996). From Paradigm to Prototype and Back Again: Interactive Aspects of Cognitive Processing in Standardized Survey Interviews. In N. Schwarz and S. Sudman (eds.), *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research.* San Francisco: Jossey-Bass Publishers.

Shneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* Second edition. Reading, Massachusetts: Addison-Wesley Publishing Company.

Sperry, S., B. Edwards, R. Dulaney, and D. E. B. Potter (1998). Evaluating Interviewer Use of CAPI Navigation Features. In Mick P. Couper *et al.* (eds.), *Computer Assisted Survey Information Collection.* New York: John Wiley & Sons, Inc.

Suchman, L. and B. Jordan (1990). Interactional Troubles in Face-to-Face Interviews. *Journal of the American Statistical Association 85*(405):232-41.

Sudman, S.., N.M. Bradburn, and N. Schwarz (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology.* San Francisco: Jossey-Bass Publishers.

Tullis, T.S. (1997). Screen Design. In M.G. Helander, *et al.* (eds.), *Handbook of Human-Computer Interaction.* Second edition. New York: Elsevier.