Towards Usage of Avatar Interviewers in Web Surveys

Lawrence A Malakhoff^{*}, Matt Jans[†]

Tags: survey practice

[DOI: 10.29115/SP-2011-0015](https://doi.org/10.29115/SP-2011-0015)

Survey Practice

Vol. 4, Issue 3, 2011

Towards Usage of Avatar Interviewers in Web Surveys

AN INVITATION

Imagine one morning in the near future a fellow named Dave arrives at his office and begins reading his email to prioritize tasks for the work day ahead. One message contains an embedded video. He clicks on the play button and sees an animated human head speaking with a human voice that asks him to click on a link to a survey. Normally, he would move on to other messages and return to do the survey later, but today he decides to complete the survey immediately because he is curious about the animated character (“SitePal,” n.d.).

After Dave clicks on the link, he notices the same character he saw in the email is now visible from the waist up and is in a seated position. Text, appearing as a text bubble in a comic strip, states that her name is Pat, and she asks the employee to turn on his webcam and speakers if he has them. His computer is equipped with a webcam, microphone, and speakers, so he switches them on. Pat asks Dave to say or click a button labeled “begin the survey.” Usually, Dave does not speak to his computer, but no one else is in the office yet so he decides to continue by saying “begin the survey.”

* **Institution:** U.S. Census Bureau

† **Institution:** U.S. Census Bureau

Figure 1 “Pat” asks a potential survey participant to respond to a Web survey.

THE SURVEY INTERVIEW

Dave views the screen and it seems Pat is looking right back. Pat blinks occasionally and her head and torso move when she adjusts her position. Dave notices that she is not a live or pre-recorded human being, but rather an animated digital character. He realizes this is going to be like no video conference he’s participated in before. Pat turns slightly to face Dave and begins asking questions. A blank area for dialogue and response is displayed and updated with each new spoken question. As Dave speaks his answer, it is displayed. Pat asks for a street address, and the employee provides it, but he also includes the name of his city and zip code. Pat blinks several times, and Dave thinks he has made a mistake by giving extra information, but sees those fields filled in on the screen and Pat moves on to the next question. Pat then asks Dave about his age. He hears a colleague walking nearby, so he chooses to type in his latest answer for privacy.

The survey interview continues with Dave answering Pat’s questions just as if she were a live survey interviewer. Pat raises her eyebrows in expectation after asking each question; she also nods slightly or says “Thank you” when Dave provides an answer to acknowledge the response. As the interview progresses, the questions become more sensitive. Pat asks a question about Dave’s income before taxes last year. He hesitates, looks down at the floor for a moment, and then sees Pat place her hand on her chin to consider what to say next. Pat then leans forward and asks a series of questions to bracket the range of Dave’s income because bracketing questions has been shown to reduce income non-response (Heeringa, Hill, and Howell 1995).

Dave chooses an income bracket. After he provides the last response to the income questions, Pat leans back, nods, and smiles. She thanks Dave for his time and closes the survey website.

After the final question, Pat thanks the respondent for participation in the

survey and closes the video connection.

HOW DID PAT “CONDUCT” THE SURVEY INTERVIEW?

Digital characters that communicate with users are known as virtual or animated agents, or avatars. “Pat” is a special type of avatar known as an Embodied Conversational Agent (ECA) (Cassell et al. 2000). These ECAs differ from avatars and agents because they can see and hear the respondent through a webcam and react to body posture and spoken language. Using a webcam and microphone with speakers provided two-way communication between Pat and the respondent.

Pat detected hints about Dave’s psychological state by looking at the change in position of the head and body through the webcam. A slumped position can indicate boredom or lack of attention (Scherer 1980). When asked about income, Dave paused longer before answering the question than he usually did, and he also avoided making “eye contact” with Pat. As a result, Pat changed the question to wording that was less invasive and easier to answer. Pat also used body language such as leaning forward to emphasize a point and nodding her head to show she was listening to the respondent’s input. She also showed facial expressions by smiling and nodding to provide feedback a response was given. The use of nonverbal behavior to moderate the conversation creates a data collection mode that may feel more like an interviewer-administered mode than a self-administered mode, blurring the lines between the two (Cassell and Miller 2008).

Pat “heard” the respondent and processed spoken responses using Natural Language Processing (NLP) (Bobrow 1964). NLP uses speech recognition technology to recognize all words in a spoken phrase and compares them to key words expected as a response for a particular question. NLP technology can also be used in cases where respondents answer more than one question at a time, such as when Dave provided the city name and zip code along with the street address requested by the ECA. When a respondent answers this way, the ECA can process this data and move to the next question. The ECA did not need to ask for the city name and zip code again once the respondent had volunteered it.

In many interviewer-administered survey designs, the interviewer asks questions in a predetermined order and controls the flow of the conversation. Yet natural communication rarely happens this way. When the respondent answers questions found later in a survey along with the current question, however, this type of interaction is known as a mixed initiative dialogue because control of the conversation changes from the ECA interviewer to the respondent if even briefly (Komatani and Kawahara 2000). This happens in survey interviews all the time (whether by design or not), and could be programmed into an ECA. The mixed initiative dialogue design may take longer to program and requires greater computing resources than a traditional

interviewer initiative dialogue design, but may be beneficial for surveys where the respondent plays an active role such as event history calendars (Sayles, Bell i, and Serrano 2010) or conversational interviewing (Conrad and Schober 2000).

WHY USE AN ECA VERSUS A HUMAN INTERVIEWER?

Data collection modes are packages of social and psychological dimensions. One dimension particularly relevant to interviewer-administered modes is called social presence, which is the degree to which the communicative partners are both in the same physical and social space. A face-to-face (FTF) interview conducted by a live human being has the highest social presence of all modes because the interviewer and respondent interact in the same physical space for the duration of the interview. Self-administration, on the other hand, (e.g., paper-and-pencil questionnaires, web questions, or other text-based communication) has the lowest social presence because no interaction with a live interviewer occurs (Short, Williams, and Christie 1976). Telephone interviews would be somewhere between these two extremes, more closely resembling face-to-face interviews. The location of ECA-based modes on this continuum is still a topic of research and discussion. Face-to-face interviews enable high rapport because the interviewer and respondent can engage in “small talk” (traffic, weather, etc.) at different times in the interview, which can assist a skilled interviewer in gauging the psychological state of the respondent. An FTF interview tends to permit smoother turn-taking while asking and answering questions than modes like telephone interviews, because all communication channels are present (e.g., visual, aural). In a face-to-face interview both people can see and are aware of each other, so talking over each other is minimized. Smooth communicative transitions may also aid in rapport, and may also affect other survey errors. A fully intelligent ECA like Pat closely mimics true face-to-face interaction in the way it uses its own nonverbal behavior, as well as perceiving and acting on the behavior of the respondent.

However, high social presence can have a downside. Research has revealed respondents give more honest answers to sensitive questions when asked in a self-administered questionnaire than by a human interviewer (Tourangeau and Smith 1996). Numerous interviewer effects may exist with a FTF and telephone interviewing, depending on the survey’s content. If the interviewer’s race is different from the respondent, it can affect data gathered on questions with content related to race and ethnicity (Schuman and Converse 1971). If there is an age and gender difference between the interviewer and respondent, there is a risk for an effect on data quality with certain survey items (Ehrlich and Riesman 1961). If the clothing an interviewer wears shows a difference in social class or religious affiliation with the respondent, data quality can be impacted (Katz 1942).

CREATING THE BEST INTERVIEWER FOR THE JOB

Using ECAs as interviewers offers researchers the ability to customize an

interview by manipulating the interviewer's appearance, including clothing, facial features, voice qualities, body size and shape, inflection of words spoken, and physical movements. This is a large design task, the technologies for which have not all been fully developed yet. However, the potential benefits of specialized ECAs are two-fold, offering data collection tools that may be better than live interviewers in some respects and better than self-administration or low-fidelity avatars in other respects.

One benefit is that an ECA, compared to a low-fidelity avatar, will conduct a more natural interview if it can converse naturally with the respondent, including reading and using verbal and non-verbal cues that are present in human interaction, (see "Ask Anna!" at [IKEA's](#) website for an example of a low-fidelity avatar). Conversation includes spoken language, facial expressions, body postures, and hand gestures. If any of these components are missing, the ECA will not be realistic, may not be believable, and worst of all may lead to break-off as a result (Cassell and Bickmore 2000).

Another benefit is that an ECA designer can create an interviewer that "looks neutral." When viewing a videotape of a person speaking questions, the voice naturally matches the facial movements of the speaker and meets the expectations of the viewer. That is not guaranteed with a programmed ECA. The design challenge for ECAs is then to pair voice with appearance. Naturally, respondents will expect speech coming from the ECA to match the ECA's facial movements, and will expect the voice to match the general appearance of the ECA (e.g., a female voice coming from a female ECA). If mouth and face movements do not match the speech, the ECA will not be believable. Respondents may find higher-pitched speech from an ECA that appears to be a large man hard to understand. The inconsistency between the voice and appearance of the ECA can change the perception of the speech and may cause it to be misunderstood (McGurk and MacDonald 1976).

Other features and characteristics of the ECA need to be explicitly addressed as well. In some cases, research with human interviewers can inform design decisions. Research on human interviewers reveals that an interviewer's race can impact response (Schuman and Converse 1971). Race effects are difficult to test in FTF interviews because of the cost and complexity of matching respondents and interviewers of different races. However, in ECA-based research skin color can be varied independently (and experimentally) with other features, like voice, sex, and clothing. Controlling such characteristics can be difficult or impossible in a FTF interview, but usage of an ECA allows precise experimental control of one or more features of an ECA.

Other features of the interviewer, like body size, can also impact on response. Researchers at RTI International have shown this in research conducted in the on-line virtual environment Second Life ("Introduction to Second Life," n.d.). In Second Life, users interact with each other through avatars. Avatars in Second Life are not yet capable of the more sophisticated capabilities of ECAs,

such as reading body language and Natural Language Processing. However, the respondent (interacting through their own avatar) sees the interviewer's avatar and vice versa, so the appearance of the interviewing avatar might be expected to have effects on responses. While voice communication is possible in Second Life, only text messaging was used in the RTI study. The study found that respondents were more likely to report a higher body mass index to a heavier-looking avatar than a thinner-looking avatar (Dean et al. 2009). These findings, along with other research on interviewer effects suggest that if information about a respondent's appearance or other demographics is known, a survey designer could potentially create an ECA with the best physical characteristics to minimize error related to interviewer effects.

IT'S NOT WHAT YOU SAY, BUT HOW YOU SAY IT

An advanced ECA could potentially monitor a respondent's psychological state by observing his or her facial expressions. Facial expressions change continuously and are affected by the content of the conversation. A respondent might wrinkle their nose, forehead, or eyebrows when discussing something unpleasant or showing worry. Just as in human to human conversation, these expressions could replace spoken words or accompany them (Ekman 1979). However, the present state of the art of facial expression recognition only permits computer processing of facial expression data in real time at five to fifteen video frames per second, depending on the computing power. Video output is normally 30 frames per second, so it is possible to miss some behaviors ("Noldus Face Reader Software," n.d.). With faster computers, live data processing at 30 video frames per second may be available in the future.

Even though real-time recognition of facial expressions at 30 frames per second is not currently feasible, the ECA can attempt to display appropriate facial expressions during an interview using the Facial Action Coding System (FACS) developed by Paul Ekman (Ekman and Friesen 1978). The FACS is based on contractions or relaxations of facial muscle groups and is comprised of action units that signify a particular emotion. Computer graphical face modeling systems, like CANDIDE ("Candide Facial Expression Generator Software," n.d.) use a FACS to set action units to generate expressions.

Head and face movements provide information about the respondent's psychological state and play an important role in conversation. These non-verbal cues fall into three functional groups. *Syntactic functions* include movements that accompany a word or accented syllable that is emphasized. Emphasis could be accomplished by both raising the pitch of the voice and blinking when pausing or on a key word or syllable. *Semantic functions* are movements which may also reinforce what is being said by replacing a word or by smiling when happy or wrinkling the nose when disgusted about something. *Dialogic functions* are movements that control speech, such that when two people are conversing they are looking at each other and do not interrupt when the other person is talking. Any of these head and face

movements can be modified by other verbal and nonverbal behavior: a speaker who is lying may avoid eye contact, while one who is telling the truth may maintain eye contact. The ECA may also signal it is listening and paying attention by using backchanneling (e.g., speaking short phrases like “I see”, “mhm”) (Scherer 1980).

Figure 2 This CANDIDE-3 face model uses the points where lines intersect to animate facial expressions.

Hand gestures form another component of a face-to-face conversation. Four basic types of gestures occur while speaking. *Iconics* represent a feature of the spoken words that can be emphasized, such as holding up and creating “air quotes” by curling the index and third finger of both hands repeatedly or showing approval with an upraised thumb. *Metaphorics* represent an abstract thought accompanied by movement, such as referencing a person or place by indicating a point in space. An example is pointing to the floor, and asking “You live HERE?” *Beats* are small formless waves of the hand which can appear as a chopping motion and occur with emphasized words (e.g., pointing at space with a finger while saying “Cases One, TWO, and THREE show€”). Hand gestures and speech complement each other because the hand movement may show the way an action is carried out when the meaning is not apparent with speech (McNeill 1992).

APPEARANCE OF THE ECA AND THE UNCANNY VALLEY

Designing a good ECA isn’t as easy as just making choices about all the design features listed so far. The specific combination of choices can make the difference between an ECA that works and one that doesn’t work. The concept of the Uncanny Valley shown in Figure 3 must be kept in mind when selecting the appearance of an ECA for a survey. This graph was developed by Masohiro Mori based on his research and theory about robots (Mori 1970). It is based on a continuum of representations of the human form (on the x-axis) ranging from low to high human likeness. It shows that as the human likeness of a robot (or ECA) gets more realistic, the familiarity (or comfort) of a human interacting with the robot will experience increases up to a point. When the

robot begins to look so much like a human that a human observer isn't quite sure whether it is human or nonhuman, familiarity plunges. If a robot or ECA is so perfectly human that a user cannot tell that it's a robot, familiarity would be as high as with another real person. This drop and rise in familiarity is known as the Uncanny Valley. Mori hypothesized that the effect would be more extreme for robots that move than those that are still.

Figure 3 The Uncanny Valley effect is greater when the avatar is moving.

MacDorman and Ishiguro have shown the Uncanny Valley concept applies to digital avatars, of which ECAs are a special type. Their research also investigated why people felt revulsion at seeing human-like characters. The authors assert that avatars may trigger fear that people can be replaced, since avatars frequently are copies of actual people. Movements of the avatar may be jerky, which can be unsettling because it generates a fear of losing bodily control. They also found avatars may also cause people to feel the same revulsion they would feel as seeing a corpse or a visibly diseased person. Therefore, if the avatar is too realistic, it may be seen as a human doing a terrible job at acting like a normal person (Karl 2006).

SitePal permits web designers to upload a photograph to create a photo-realistic (Avatar 1 in Figure 4) or select a low-fidelity cartoonish avatar (Avatar 2 in Figure 4). Specifically, Avatar 1 appears human because of the variation of skin tones and shadows. The still image resembles an airbrushed photograph, but when Avatar 1 is speaking, the animation in the mouth area is coarse and

jumpy, which would reduce the familiarity a human observer feels compared to viewing a video of a person speaking. Avatar 2 is evenly illuminated with no shadows. There is much less detail in the mouth (the tongue is not always shown) when Avatar 2 is speaking but the animation seems more natural, which is why we suggest that Avatar 2 appears to be more familiar than Avatar 1. Both of these avatars can be viewed at the SitePal site. The Uncanny Valley effect is more apparent with motion as seen in this video clip of various robots: <http://www.youtube.com/watch?v=CNdAIPoh8a4&feature=related>

Figure 4 Two types of avatars: photo-realistic and low-fidelity.

Another feature of low-fidelity avatars that web survey designers should consider is the appearance of the eyes. The white part of the eye, in addition to gross head movement, provides cues to humans where someone is looking (Than 2006). Avatar 2 has eyes with a larger white area in comparison to the pupil than Avatar 1. Since the white area is more apparent in Avatar 2, it is not difficult for human observers to assess where the low-fidelity avatar is looking, as they often do in a conversation with another human. The designer could use this feature to create the illusion of eye contact, perhaps to signal expectation of a response. Programming the avatar to appear to be looking down at a piece of paper could signal times during which the avatar is “thinking” and the respondent should wait.

HOW TO PUT AN ECA TO WORK TODAY

Running an ECA like “Pat”, which engages in natural multi-modal responses with survey respondents, requires multiple software programs and significant processing power in turn. The numerous software sub-programs that run Pat’s

natural language processing, head movements, hand gestures, body position, facial expression and expression recognition functions would be very difficult for any individual research team to assemble and integrate. Further, software for these functions is frequently written for isolated research projects, and is never intended to work together. Hung-Hsuan Huang and researchers in Japan and Croatia collaborated to address these issues by creation of a generic ECA (GECA) framework (Huang et al. 2008). This framework will enable rapid prototyping of an ECA and sharing of results among researchers. The goals of this collaboration are create an ECA environment that can recognize verbal and non-verbal inputs from the respondent or user, interpret them to formulate a response, and have the ECA display appropriate verbal and non-verbal behaviors as a response through a computer generated animation player. An early usage of this type of ECA was created by Justine Cassell and her colleagues (Cassell and Miller 2008; Cassell and Bickmore 2000). REA (Real Estate Agent) was programmed to converse with a human customer to sell them a house or rent them an apartment. While intelligent in that area, she was limited to that topic only. The GECA framework will enable more general types of conversation than REA. Huang and colleagues plan to make the GECA framework publicly available with a reference ECA toolkit for rapid development of ECA prototypes.

ECA's can be used for experimental research to study their effects on survey data, as there are many unanswered questions about the use of ECA's in data collection. Conrad et al. (2008) recently studied respondent's reactions to an ECA named Derek. This study was conducted with a "Wizard-of-Oz" technique, making Derek appear intelligent to the respondent even though he was controlled by an experimenter. Depending on the experimental condition, Derek would provide clarification about a question's meaning or key terms (with high dialogue capability), or only provide neutral probes (e.g., "let me repeat the question," or, "whatever it means to you") when the respondent asked for help or the experimenter thought they needed it. Also depending on the experimental condition, Derek was displayed as a disembodied head with choppy mouth movements and no blinking, or with shoulders and more refined mouth movements and blinking (i.e., low versus high visual realism). There was no effect of dialogue capability on respondents' ratings of how personal Derek seemed when the high visual realism Derek was displayed, and the respondents smiled more at the visually realistic Derek than the unrealistic one. However, when Derek was shown as a disembodied head (low visual realism), raters called the high dialogue version of Derek more personal, possibly due to his ability to communicate and answer questions. Derek's highest ratings on the impersonal-personal scale were seen when he was less visually realistic, but able to communicate. The study did not compare Derek to a human interviewer, so we still don't know if the high dialogue capability and visually realistic version of Derek would be as effective as a human interviewer in improving understanding of questions and concepts and increase response accuracy. While Conrad et al, Huang, and the research by

RTI all show the potential for ECAs in survey work, none of their applications are yet easily implemented. True ECAs with hand and body gestures and facial expressions and the capability to recognize human expressions, gestures, and speech may come about through further research to create a generic ECA framework as Huang as suggested.

If the goal is to simply design an avatar for survey data collection, there are numerous websites that provide tools to create avatars. The U.S. Census Bureau does not and cannot endorse any particular company, but SitePal offers the capability for web survey developers to incorporate avatars into their survey. These avatars are visible from the shoulders up and do not use their hands, or actually see or hear a potential respondent, like Pat in the opening example, and thus are not technically considered to be ECAs.

Second Life provides another avenue for survey researchers. Web survey practitioners can conduct interviews in Second Life, such as those done by RTI. Although interviewers and respondents could potentially speak through microphones and listen to each other through speakers, the RTI study had the interviewer and respondent interact through their avatars by text messaging only. This had the benefit of keeping a written record of the conversation. As information security is always a concern with interview data, interested survey researchers may wish to explore the possibility of creating secure spaces (islands) in Second Life where access to the interview site is controlled by a user name and password. A standard has been developed for Federal Government virtual worlds known as V-GOV which addresses collaboration and training courses in a secure environment (“Federal Consortium on Virtual Worlds,” n.d.).

Given enough time and research, an ECA can be developed that can closely mimic human interviewer behavior. As a result, the issues survey practitioners encounter now with human interviewer effects may become less of a problem. On the other hand, ignoring what we know about interviewer effects could lead to programming of interviewers that actually create more interviewer effects than they remove. The research field is wide open for exploring these issues. Conrad and Schober’s work indicates the dialogue spoken by the ECA when interacting with the respondent may be more important than the appearance. It is possible that all the artificial intelligence (AI) problems of automating the ECA to move, speak, and display expression and recognize and correctly react to these behaviors in human respondents through the GECA framework as Huang suggests will be solved in the shorter term. However, what the ECA says in the interaction with the respondent is a more difficult and longer term problem. Through AI techniques and robust dialogue design, ECAs can become another resource for web survey developers in the near future. Whatever the future brings for the use of ECAs in survey research, it seems to be a bright and interesting one.

Disclaimer: This report is released to inform interested parties of ongoing

research and to encourage discussion of work in progress. The views expressed are those of the authors and not necessarily of the U.S. Census Bureau.

REFERENCES

- Bobrow, D.G. 1964. "Natural Language Input for a Computer Problem Solving System." 1964. <http://hdl.handle.net/1721.1/5922>.
- "Candide Facial Expression Generator Software." n.d. <http://www.icg.isy.liu.se/candide/>.
- Cassell, J., and T. Bickmore. 2000. "External Manifestations of Trustworthiness in the Interface." *Communications of the ACM*.
- Cassell, J., and P. Miller. 2008. "Is It Self-Administration If the Computer Gives You Encouraging Looks?" In *Envisioning the Survey Interview of the Future*, edited by F.G. Conrad and M.F. Schober, 161–78. John Wiley & Sons.
- Cassell, J., J. Sullivan, S. Prevost, and E. Churchill. 2000. *Embodied Conversational Agents*. Cambridge, MA: MIT Press.
- Conrad, F.G., and M. Schober. 2000. "Clarifying Question Meaning in a Household Telephone Survey." *Public Opinion Quarterly* 64: 1–28.
- Conrad, F.G., M.F. Schober, M. Jans, R. Orlowski, D. Nielsen, and R. Levenstein. 2008. *Features of Animacy in Virtual Interviewers: Dialogue, Capability and Visual Realism*. New Orleans: American Association for Public Opinion Research Annual conference.
- Dean, E., S. Cook, M. Keating, and J. Murphy. 2009. "Does This Avatar Make Me Look Fat? Obesity and Interviewing in Second Life 2(2): 3D Virtual Worlds for Health and Healthcare." *Journal of Virtual Worlds Research*.
- Ehrlich, J.S., and D. Riesman. 1961. "Age and Authority in the Interview." *Public Opinion Quarterly* 25: 39–56.
- Ekman, P. 1979. "About Brows-Emotional and Conversational Signals." In *Human Ethology*, edited by M. von Cranach, K. Foppa, W. Lepenies, and D. Ploog, 169–248. Cambridge: Cambridge University Press.
- Ekman, P., and W. Friesen. 1978. *Facial Action Coding System*. Consulting Psychologists Press, Inc.
- "Federal Consortium on Virtual Worlds." n.d. <http://www.ndu.edu/iCollege/fcvw/fcvw10/index.html>.
- Heeringa, S., D. Hill, and D. Howell. 1995. "Unfolding Brackets for Reducing Item Nonresponse in Economic Surveys." In *Health and Retirement Study Working Paper Series, Paper No. 94-029*. Survey Research Center, Institute for Social Research, University of Michigan.
- Huang, H., A. Cerekovic, I. Pandzic, Y. Nakano, and T. Nishida. 2008. "Scripting Human-Agent Interactions in a Generic Eca Framework." In *Applications and Innovations in Intelligent Systems XV: Proceedings of AI-2007, the Twenty-Seventh SGA International Conference on Innovative Techniques and Applications of Artificial Intelligence*, 103–15.
- "Introduction to Second Life." n.d. <http://www.youtube.com/watch?v=b72CvvMuD6Q>.
- Katz, D. 1942. "Do Interviewers Bias Poll Results?" *Public Opinion Quarterly* 6: 248–68.
- Komatani, K., and T. Kawahara. 2000. "Flexible Mixed-Initiative Dialogue Management Using Concept-Level Confidence Measures of Speech Recognizer Output." [Http://www.aclweb.org/anthology/C/C00/C00-1068.pdf](http://www.aclweb.org/anthology/C/C00/C00-1068.pdf).

- MacDorman, K.F., and H. Ishiguro. 2006. "The Uncanny Advantage of Using Androids in Cognitive and Social Science Research." *Interaction Studies* 7 (3): 297–337.
- McGurk, H., and J. MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264 (5588): 746–48.
- McNeill, D. 1992. *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago.
- Mori, M. 1970. "The Uncanny Valley." Translated by Karl F. MacDorman and Takashi Minato. *Energy* 7 (4): 33–35.
- "Noldus Face Reader Software." n.d. <http://www.noldus.com/human-behavior-research/products/facereader>.
- Sayles, H., R.F. Belli, and E. Serrano. 2010. "Interviewer Variance between Event History Calendar and Conventional Questionnaire Interviews." *Public Opinion Quarterly* 74 (1): 140–53.
- Scherer, K.R. 1980. "The Functions of Nonverbal Signs in Conversation." In *The Social and Psychological Contexts of Language*, edited by H. Giles and R. St. Clair, 225–44. Lawrence Erlbaum Associates.
- Schuman, H., and J.M. Converse. 1971. "The Effects of Black and White Interviewers on White Respondents in 1968." *Public Opinion Quarterly* 35: 44–68.
- Short, J.A., E. Williams, and B. Christie. 1976. *The Social Psychology of Telecommunications*. New York: John Wiley & Sons.
- "SitePal." n.d. <http://www.sitepal.com>.
- Than, K. 2006. "Why Eyes Are so Alluring." 2006. http://www.livescience.com/health/061107_human_eyes.html.
- Tourangeau, R., and T.W. Smith. 1996. "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Questions Context." *Public Opinion Quarterly* 60 (2): 275–304.