



The Media Equation

The screen shows a bag of popcorn on a table. Several kernels have fallen from the bag. An adult, watching with preschool children, asks a question about the picture: "What will happen to the popcorn if I pick up the television set and turn it upside down?" A lot of kids say that the rest of the popcorn will spill out of the bag.

While interesting, this mistake is not really surprising. Children often take television literally. They'll soon be older, and will be wise enough to know better.

A novice computer user, excited about getting a new machine, sits down to try some new software. When the application is started, an animated dog on the screen greets the user with a smile and a big "Hello," then asks, "What would you like to do today?" When the dog appears, the user grins and says "Hello" back.

The smile isn't really so surprising, either. Computer novices like special treatment. Experts, however, don't seem to need or respond to on-screen social niceties.

In a televised presidential debate, members of a studio audience asked the candidates questions. One of the candidates began a response by leaving his place at the podium and walking toward the person who asked the question. On television, the candidate appeared to narrow the distance between himself and each viewer watching at home. Many

viewers responded quite favorably to his comments, not because of what he said, but because he seemed to move physically closer to the viewers.

Surprising? Not really. Distracted by the task of following a debate, viewers might forget that a screen image isn't a real person. If they thought about it for a moment, it seems unlikely that they would be influenced by the candidate's apparent entrance into their personal space.

A research team from an artificial intelligence lab demonstrated a new robot to a group of computer specialists. They described a one-legged hopping machine that had memory, could learn new instructions, was bred for intelligence, could focus attention, could accept punishment and reward, and even had a personality. No one at the meeting noticed that they were using the vocabulary of human psychology to describe the abilities of a collection of wire, silicon, mechanical joints, and computer code.

The language used to describe the robot is also not surprising. The researchers thought they were just using the metaphor of the human mind as a useful tool for analyzing intelligent systems. No one at the meeting seemed to think that the robot was a real person or had human attributes; they just found it convenient to talk about the robot as human.

How Different Are Media and Real Life?

These stories show that there are good reasons why some people might confuse media and real life. Children do it because they're young, and novices because they don't have enough experience. People engaged in a difficult task may be confused because they are too distracted to notice otherwise and smart people may do it because it's a useful metaphor. All of these dismissals have one other thing in common, however. They all suggest that the confusion of mediated life and real life is rare and inconsequential, and it can be corrected with age, education, or thought.

We have collected a great deal of evidence that shows this conclusion

The Media Equation is not true. Equating mediated and real life is neither rare nor unreasonable. It is very common, it is easy to foster, it does not depend on fancy media equipment, and thinking will not make it go away. The media equation—*media equal real life*—applies to everyone, it applies often, and it is highly consequential. And this is surprising.

The media equation comes from a research project that we call Social Responses to Communication Technologies. In short, we have found that individuals' interactions with computers, television, and new media are *fundamentally social and natural*, just like interactions in real life. The key word is "fundamentally." Everyone expects media to obey a wide range of social and natural rules. All these rules come from the world of interpersonal interaction, and from studies about how people interact with the real world. But all of them apply equally well to media.

Testing the Media Equation

Consider this research example, a prototype of our work on the media equation. If either of us called you on the phone to ask how well you liked this book, you would likely be polite and say that it was fine. But if *someone else* were to make the same phone call, the evaluation might be less positive. This is an example of a simple *social* rule: People are polite to those who ask questions about themselves.

We found that this same rule applies to media. When asked to evaluate the performance of a *computer*, people gave more positive answers about how well a computer performed if it asked questions about itself than if a different computer asked the same questions. People are polite to computers, too.

The same equation—*media = real life*—works for pictures as well. For example, in the real world, motion demands attention, especially if what moves is coming toward us. This is a simple natural rule. The presence of motion determines how attention is allocated, and ultimately, what is remembered about an experience.

We found that this rule applies to media as well. Motion in *pictures*, especially motion that appears directed at the viewer, stimulates

physical activation in the brain as if the moving objects were actually present. Pictures, too, are natural experience.

Evidence for the media equation now includes thirty-five studies that have recreated a broad range of social and natural experiences—but with media taking the place of real people and places. All of the studies go far beyond pleading with a computer or yelling at a television. We have looked at everything from interfaces that flatter to the personalities of cartoon characters, from talking computers to movies on home theaters, from pictures of objects thrown at the viewer's face to faces that cover a wall, from super-high-resolution images to fifth-generation home video, and from questionnaires that ask people what they think to brain-wave measures that tell us what people could not otherwise describe.

The consistency of responses in all of these studies led us to the media equation: Media equal real life. But that's not where we started.

Intuitions to the Contrary

The media equation is not intuitive. In fact, the equation competes with ideas about media that seem much more obvious. What seems most obvious is that media are *tools*, pieces of hardware, not players in social life. Like all other tools, it seems that media simply help people accomplish tasks, learn new information, or entertain themselves. People don't have social relationships with tools.

Likewise, pictures *seem* to be merely symbols that represent people and places that are not actually present. People seem to attend to media with detachment and they seem to focus on *inequalities* between images and the real world. Everyone seems to believe that media are special, and everyone has pet theories, personal experiences, stories, and complaints about media that bolster the belief that media are unique.

Many of these ideas characterized our own thoughts about media when we began our research. We believed that people might occasionally confuse media and real life, but the confusion wasn't pervasive, and most important, it was curable. Along with many participants in our research, we were tempted to think that the results were curious but

The Media Equation
personally irrelevant. *I* know that computers don't have feelings. And *I* can certainly distinguish between life on the screen and the real thing.

We now think our intuitions were wrong, however, and this led us to a second important conclusion: People respond socially and naturally to media even though they believe it is not reasonable to do so, and even though they don't think that these responses characterize themselves. This conclusion is, by definition, counterintuitive, and it significantly complicates research. Participants in our experiments were not aware that they equated media with real people and places. Therefore, attempts to verify the media equation can't rely solely on talking to people, listening to their stories, or asking them questions on a survey. Social and natural responses to media are not conscious, and as a consequence, people are not able to confirm the media equation, even if they'd like to help. This means that our research story is also about how to observe what people cannot themselves describe.

The media equation tempts many qualifications, but few are needed. One tempting qualification is that people might be fooled, but only by the rificest of new media—virtual reality systems, IMAX theaters, or full-motion holography of the future. Our research suggests otherwise. Even the simplest of media are close enough to the real people, places, and things they depict to activate rich social and natural responses. Many of our studies generate these responses with rather pathetic representations of real life: simple textual and pictorial material shown on garden-variety technology. The equation still holds, however. Mediated life equals real life.

Moreover, the media equation is not limited to a particular type of person. While it might be tempting to assign the confusion between media and real life to problems of age, knowledge, distraction, or convenience, our research shows that social and natural responses are remarkably common, and true for every group we have tested, including children, college sophomores, people in business, and technology experts. *All* people automatically and unconsciously respond socially and naturally to media.

Are there times when the media equation doesn't apply? Of course. People are quite capable of thinking their way around it. People can

treat media as tools or as images that only represent the real world but are not real themselves. But these responses, all highly thoughtful, share many surprising characteristics: They require a lot of effort, they tend not to occur when people are tired or when other things compete for their attention, and they are always difficult to sustain. The automatic response is to accept what seems to be real as in fact real.

People have done some amazing things in our labs. They have taken great care not to make a computer feel bad, they've felt physically threatened by mere pictures, and they've attributed to an animated line drawing a personality as rich as that of their best friend. It eventually occurred to us that people were not doing these things because they were childish, inexperienced, distracted or because they needed a metaphor. We had to acknowledge that these responses were fundamental to human, and we had to acknowledge that they were important.

Benefits of Considering the Media Equation

An understanding of the media equation accomplishes four things, all of them important, and each difficult to obtain otherwise.

Improving the Design of Media

The first benefit of the media equation is practical. If human-media interactions are social and natural, then there are a number of unexpected ways to improve the *design* of media. Many of these applications stem from one important conclusion: Humans are experts on social relationships, and they are experts on how the physical world works. Rules for using media as tools, on the contrary, are often arbitrary and must be learned. When media conform to social and natural rules, however, no instruction is necessary. People will automatically become experts in how computers, television, interfaces, and new media work.

Because people have a strong positive bias toward social relationships and predictable environments, the more a media technology is consistent with social and physical rules, the more enjoyable the technology will be to use. Conforming to human expectations means that

there is instant expertise just because we're human, and there are positive responses, including feelings of accomplishment, competence, and empowerment.

All of the chapters to follow include design rules that are based on our experiments, and on the larger social science literature in the area being discussed. These rules include simple ways to give personality to a computer: when to use multiple voices in an interface; how movement can guide visual attention; the trade-offs between visual and audio fidelity in video-conferencing; how to make effective advertisements; and general principles for making interactive multimedia presentations more polite, friendly, or arousing.

Evaluating Media

The second application of the research concerns judgments about media and especially worries about new media. If media can do harm, then the explanation for how that process works will depend on the *psychology* of the people who use them, not just the technologies per se, nor the industries that make the appliances and produce the content. Psychology is at least as important in the evaluation of media as economics, policies, laws, and the media industries themselves. And it is from the perspective of psychology—the study of how people think and feel—that the media equation comes.

Media have evolved to capitalize on fundamental human responses to them. This is true of everything from negative political advertising, to the development of chatty "happy-talk" news, to discussions about computer agents. By trial and error, people who design media are gradually discovering the intricacies of how media work. The media equation helps to organize these principles and to form a basis for a critique of media.

Useful discussions of how to change what we don't like about media need to consider that not all responses to media are obvious; some are constrained by unconscious thoughts and feelings. In the chapters to follow, we discuss implications of the research for communication policy, ethics, activism, and even personal media habits. These discussions

The Media Equation

include ideas about how media affect political cynicism, worship of media personalities, violent behavior, shyness, and even deviant sexual behavior.

New Methods for Research

The third benefit of the research concerns the *methods* used to find out new things about media. Our experiments use the same methods that psychologists use to determine how people respond to each other and to the physical world. The biggest advantage of borrowing these methods is that they differ substantially from other techniques that are currently used to study media. Many popular methods, especially focus group techniques, rely heavily on the assumption that people can be introspective about media experiences. This is an assumption that we think is frequently wrong.

Many of the procedures in the studies we have conducted are uncommon, and many of them require elaborate or unusual ways to record responses. We have tried to indicate the advantages of each method in relation to more traditional techniques for studying media, and we have tried to provide enough detail about the research methods and results to enable readers to evaluate evidence for the media equation.

Big Issues

The fourth benefit of the media equation is the opportunity to reconsider big issues related to media. These are ideas, most often discussed in universities, that have often assumed too much about how people interact, watch, listen, and read. The first is the idea that media are predominantly tools, and that people consciously use media tools to serve their needs. Rather, we think our research shows that media are perceived as real people and places, and that human responses to media are determined by the rules that apply to social relationships and navigating the world. Responses to media are not primarily governed by rules about how to use appliances more akin to a hammer or car.

Another common assumption about technology is that when people treat media as human, they are guilty of anthropomorphism, a mistaken belief that inanimate objects are human. Anthropomorphism,

The Media Equation

however, is rare and is not the basis for the media equation. Social responses are commonplace, even when people know that the responses are inappropriate.

There is also a strong tendency in discussions about media to consider the *real* similarities between media machines that think and the thinking done by humans. The consequence of this confusion is a bias toward ideas about "hardware" similarities between media and people, for example, the similarity between neurons in the brain and neural networks in a computer. This view, however, underestimates a critical human tendency to confuse what is real with what only *seems* to be real. For example, people are not only deemed intelligent because they think well; they are also intelligent because they *appear* to think well, quite apart from the real merits of their mental capacity. Because of this, we should not worry as much about whether media are the same as their real-world counterparts, and instead think more about whether media are *perceived* as identical. Perceptions alone can alter a broad range of human behavior.

The media equation also challenges the cherished assumption that words and pictures in media are *symbolic representations* of things that are not actually present. When information is mediated, we often assume that people only think about *who* sent the information and *why*, and most importantly, what the information *means*. Our research shows that this view can assume too much. When a picture threatens, we don't think about who made it, what they intended, and what it all means; we often think only about what we should *do*.

Why Do People Respond Socially and Naturally to Media?

In a recent U.S. congressional hearing, entertainer Shari Lewis answered questions from senators about television violence and children. She brought her hand-puppet, Lamb Chop, with her. At one point in the hearing, the story is told, Lamb Chop the puppet made a statement (through the projected voice of Shari Lewis) about television violence. A senator then said, "Miss Lewis, do you agree with Lamb Chop?"

If it took you a moment to laugh at this story, you are in good

company. None of the senators laughed right away either. Everyone sitting in the hearing room seemed to think that it was perfectly reasonable for a puppet to give testimony, and it was perfectly reasonable to ask the puppeteer if she had a *different* opinion than the puppet. A puppet is a natural conversational partner. The hard part, once engaged, is to think of it as anything else.

How is this possible? The answer is that people are not evolved to twentieth-century technology. The human brain evolved in a world in which *only* humans exhibited rich social behaviors, and a world in which *all* perceived objects were real physical objects. Anything that *seemed* to be a real person or place *was* real.

During nearly all of the 200,000 years in which *Homo sapiens* have existed, anything that acted socially really was a person, and anything that appeared to move toward us was in fact doing just that. Because these were absolute truths through virtually all of human evolution, the social and physical world encouraged automatic responses that were, and still are, the present-day bases for negotiating life. Acceptance of what only *seems* to be real, even though at times inappropriate, is automatic.

Modern media now engage old brains. People can't always overcome the powerful assumption that mediated presentations are actual people and objects. There is no switch in the brain that can be thrown to distinguish the real and mediated worlds. People respond to simulations of social actors and natural objects as if they were in fact social, and in fact natural. A puppet, while obviously different from a human when people *think* about it, is close enough to a human to fool our old brains, especially if we are otherwise engaged or running on automatic. So, did Shari agree with Lamb Chop or not? That's the question.

People often don't scrutinize their actions or their environment. People frequently live life mindlessly and with little introspection. When our brains automatically respond socially and naturally because of the characteristics of media or the situations in which they are used, there is often little to remind us that the experience is unreal. Absent a significant warning that we've been fooled, our old brains hold sway and we accept media as real people and places.

With effort, people can perhaps think their way out of primitive, automatic responses. This is effective, for example, at a scary movie. A mental reminder that "it's only a movie, it's only a movie," does offer an escape. But this strategy makes it hard to follow the plot, and is not typical or usual. The default is to automatically and unconsciously ignore fabrication and expect reality, as if the technology itself were invisible. The fact that the movie scared us in the first place is good evidence that media are real first, and false only after we think about it. There are vestiges of old brains in modern thinking.

How We Did the Research

If media and real life are similar, then knowledge of how people respond to *real* people and places should reveal a lot about how people respond to media. The place in the library for this information is the social science section, especially psychology, sociology, and communication.

Social science is not an obvious place to start. A reference librarian might suggest a hundred other places first. An interest in media machines or interfaces gets you to electrical engineering, computer science, or to a media laboratory; an interest in media content and production might get you to the fine arts section or to a film school; and an interest in media jobs could get you to journalism, broadcasting, or to a newspaper.

An interest in how people work, however, is the province of social science, and it is this area that has the best information with which to leverage the media equation. What social scientists know about how people respond to each other can be used to explain how people respond to *media*. But there is even more reason to land in the social sciences: Not only do social scientists know about social and natural responses, they are also good at measuring them, especially when the responses cannot be known through intuition, casual observation, or introspection.

The social science section of the library contains results of a century of research about how people respond to each other, about how they respond to their natural environment, and about how to measure both.

Our strategy was to borrow, as outlined with simplifying overstatement, using the following steps.

Step 1: Pick a social science finding about how people respond to each other or to the natural environment. This might be a finding about how people express politeness, how they respond to personalities different from their own, how people negotiate interpersonal space, or how they stereotype by gender. On the natural side, this could be a finding about how people respond to motion, how they regulate emotional states, or how what people hear influences what they see.

Step 2: Find the place in the report where a social or natural rule is summarized. For example: "People like to be praised by other people, even if the praise is undeserved"; "people like other people who have personalities that are like their own"; or "the onset of motion elicits intense focus."

Step 3: Cross out the words "person" or "environment" in the studies, and substitute media. The substitution could be a media machine, some exemplary content, or an interface. For example, if the interest is computers, the rule that people like other people who flatter them becomes this: People will like *computers* that flatter them. If the rule is that people feel strong emotions when another person comes close to them, then people will feel strong emotions when a *picture of a person* comes close to them. If people orient visually to movement in the natural environment, then people will orient to movement on a *screen*. The result in each of these cases is a prediction, complete with accompanying explanation, about human interaction with media.

Step 4: Find the part of the report that describes how the rule was tested. For the flattery rule, for example, a study was done that observed students who were praised by a tutor. In the personality study, two people talked across a table, and tried to solve a puzzle. In the motion study, orienting responses were measured by recording brain electrical activity.

Step 5: Cross out, again, one of the people or the description of the environment, and substitute media. Instead of a human tutor in the flattery example, you now have a *computer* tutor. Instead of two people solving a puzzle, you would have a person and a *computer* solving a

puzzle. Instead of measuring brain activity in relation to actual movement, brain activity is measured relative to moving *pictures*. This substitution provides a method for testing the prediction.

Step 6: Run the experiment. This is important. To arrive at the media equation, we made no *assumptions* about whether the results would be the same for real people and places, and their mediated counterparts. Rather, people were asked to visit a laboratory, and to experience media. We gathered data that were summarized, statistically analyzed, and interpreted using the same standards that were used in the original experiments. When an experiment is successful, the results of the media study match the original study.

Step 7: Draw implications. We are interested in how media should be discussed, and this includes the ways the experiments can inform designers, producers, directors, critics, regulators, owners, and consumers. A desire to apply social science in these areas is one of the reasons that we started the project in the first place. We also draw implications for the theory of human-media interaction as well as implications for the social sciences more generally. But always, we try to apply the theory.

A lot of steps, but nothing terribly complicated. In the chapters that follow, we will show how this method, and the media equation that resulted from its use, can frame and organize some surprising things that people do with media.

Politeness

The three-term mayor of New York City, Ed Koch, used to ask a simple question: "How am I doing?" His question closed speeches, greeted crowds, and made a great sound bite. Imagine that the mayor, exuding enthusiasm and smiling brightly, turned to you with his question, "How am I doing?" Is your first impulse a critical evaluation of his political agenda? Likely not. "You're doing swell, Mr. Mayor!"

Imagine that in the evening, a pollster from the *New York Times* calls with the same question. "How is the mayor doing?" Without hesitation, your answer might be more truthful: "Not so well."

What explains your two different responses? Was the first a lie? Not really. Instead, it was *polite*, a virtue not a vice, even in New York City. Trying to make people happy is the norm, and it generally works well. When the mayor asks about how he's doing, he's implicitly telling listeners what would make him happy—receiving a positive answer. When someone *else* asks the same question, however, the mayor's feelings are not at stake; honesty prevails.

We don't all carry an etiquette handbook, but everyone seems to know good manners. Although violations exist, most people are polite most of the time. The level of conformity is striking, a fact blurred by vivid memory for occasional lapses. Politeness is ubiquitous, and it's practiced automatically. Communities encourage it, and the rules are a centerpiece of childhood socialization. Politeness, even more than early mastery of letters and numbers, is a genuine mark of an educated child, as any kindergarten teacher can attest.

The example of the mayor's question helps identify one important politeness rule. When people ask about themselves, they will usually receive more positive responses than when an independent person asks the same question. That is, if the person in the question also asks it, politeness reigns.

There is still some chance, however, that a response could be negative. After all, the mayor just asked a question; he didn't give explicit instructions about what the response should be, nor did he mention the response that he would like to hear. The desired response must still be inferred. This is likely, but something could still go awry, especially if the respondent were tempted to think too much about politics. If you were the mayor, is there a way to gain even greater control?

Yes, there is. Imagine the response if the mayor looks you in the eye and says, "I think I'm doing great! How do you think I'm doing?" Now the stakes are even higher. The mayor isn't relying on an assumption that you will be polite. He's giving you an explicit prompt about just what you should say. If you don't answer positively, you are at odds with his expressed self-opinion. That's really impolite.

The subtleties of politeness rules, however, are not fully exposed yet. Consider again the pollster from the *Times*. She is not only going to hear more negative responses about the mayor, she will also hear more varied responses. The mayor himself, since people are likely to be polite, is going to hear pretty much the same thing from everyone: "Things are fine, Mr. Mayor." But the pollster, since she is more likely to hear the truth, will hear a range of responses: some good, some bad, some in between. This suggests another rule: When people ask about themselves, the answers will be more homogeneous than when someone else asks the same questions.

Politeness has received a great deal of attention from psychologists. Their results lend credence to the rules that the mayoral example summarizes. While the scientific experiments confirm an important set of rules for social behavior, they are not surprising. Most people could easily imagine responding in similar ways. Our question was whether the same rules might apply to media. Are people polite to computers?

Politeness Rules for Humans vs. Politeness Rules for Computers

Millions of dollars are spent each year trying to make computers friendly and polite. Sometimes it works, sometimes it doesn't, but in any case, it is a useful exercise because there are humans at the other end who will recognize and appreciate the effort. Designers want people to like media machines, and politeness is one way to ensure this.

But what about manners in the other direction? Do computer users actually reciprocate? Do people try to be polite to machines? If the answer is yes, we could make the following predictions, substituting computers for the mayor and the *New York Times*:

Rule 1: When a computer asks a user about itself, the user will give more positive responses than when a different computer asks the same questions.

Rule 2: Because people are less honest when a computer asks about itself, the answers will be more homogeneous than when a different computer asks the same questions.

What "social scene" do these predictions suggest? Imagine that you have just used a computer to do some work. Afterward, the machine, using simple text, praises its abilities. The same machine then asks for an evaluation. Are you polite (as most were to the mayor) instead of truthful? Would you be more likely to tell the truth if you wheeled your chair over to a second computer that asked for the same evaluation? And would the collection of responses to the second computer, because the truth was more likely told, represent a greater variety of opinions?

Many people would answer no to these questions. First of all, computers do *not* have feelings, so it would be foolish for users to be polite to them. Whose feelings would be spared, anyway? Second, it would be quite unusual to think of two different computers as two independent people. What is it that they are independent of? Certainly not each other. Finally, people are trained to be honest to computers, not to humor them. There is good evidence that when people are interviewed about sensitive topics, they are more likely to tell the truth to a computer than to another person.

The two politeness rules may sound silly when you think about their application to media. Thinking about it, however, is exactly what people may not do. We predicted that people would still be polite, media notwithstanding.

Our rationale was this: Computers, in the way that they communicate, instruct, and take turns interacting, are *close enough* to human that they encourage *social* responses. The encouragement necessary for such a reaction need not be much. As long as there are some behaviors that suggest a social presence, people will respond accordingly. When it comes to being social, people are built to make the conservative error: When in doubt, treat it as human. Consequently, any medium that is close enough will get human treatment, even though people know it's foolish and even though they likely will deny it afterward.

Observing Polite Interactions with Computers

To determine whether our predictions were accurate, we conducted several laboratory experiments. In all respects possible, they were similar to the experiments done by psychologists who study politeness in human-human interaction. The only difference was that participants in our studies worked with a *computer* rather than a person.

Here is how the first experiment worked. We invited twenty-two people to a laboratory and told them they would be working with a computer to learn about various topics. We told them that at the end of the work session we would ask them to evaluate the computer that they used. They would have to tell us how they felt about the computer and how well they thought the computer had performed during the session.

One person at a time sat down in front of the computer to be tutored about various statistical profiles of Americans. The computer, a black NeXT computer with a 21-inch black-and-white monitor, was placed on an office desk in front of each person. The only thing that the computer displayed was text and graphical buttons: no pictures, no voices, not even an icon.

Twenty facts were presented in each session. Here is an example: "According to a Harris Poll, 30% of all American teenagers kiss on the first date." After the presentation of each fact, the computer asked the users if they knew anything about the fact they had just read. Using a mouse for input to the computer, participants indicated whether they knew "a great deal," "somewhat," or "very little." Participants were told that the computer would provide some additional facts based on how much they said they already knew. In reality, however, everyone received the same information presented in the same order.

After participants finished hearing the facts, the computer gave them a test and then told the participants which answers were right and which were wrong. Then the computer told each user what it thought of its own performance; in all cases, the computer said that it had done a great job.

The participants were divided into two groups to evaluate the computer's performance. Half were assigned to answer the evaluation questions on the same computer that had just praised itself. The other half answered the identical questions on a different computer located on the other side of the room.

In the evaluation, participants were asked how well different adjectives described their session with the computer. The adjectives were chosen to capture how well each person thought the computer performed, as well as how much they liked the interaction. Twenty-two adjectives were used to evaluate the computer, including *accurate*, *analytical*, *competent*, *fair*, *friendly*, and *helpful*.

People Are Polite to Computers, Too

If machines don't deserve our positive regard, there should be no differences in evaluations of the computers based on which one asked the questions. If computers are social actors, however, then participants who responded to the *same* computer that taught them should be polite, and uniformly so, just as if the machine were a real person with real feelings.

What happened? As predicted, the participants who answered questions on the same computer gave significantly more positive responses than did participants who answered on a different computer. The computers got the same treatment that people would get. The respondents who interacted with the same computer throughout the experiment rated it more positively on twenty of the twenty-two adjectives presented. Based on statistical tests, we can be confident that these results did not occur by chance.

The variance in responses also conformed to the prediction. Evaluations made on the same computer had a significantly smaller range of responses than did evaluations made on the other computer. Participants felt freer to be honest when an independent computer asked the questions, and this increased the variance in evaluations of the computer's performance.

What did the participants themselves think about these results? When we told them what we predicted (after the experiment was over, of course), all of them said confidently that they did not, and never would, change their evaluations just to be polite to a computer. From these comments, we concluded that social responses to media were unconscious and automatic.

When research results are first discussed, colleagues often ask tough questions. The questions in this case were about alternative explanations for these social responses. One issue was the definition of "other." In the first experiment, the "other" asking the questions was a different computer, set up on the other side of the room from the first computer. Could there have been something about the particular placement of the computers that caused the differences, rather than a perception on the part of users that a computer was a social entity that warranted polite treatment?

To test this, we decided to see if the same results would occur if we made the "interviewer" something other than a computer. We did the experiment again, and this time we made the "other" a paper-and-pencil questionnaire instead of a computer.

The participants who used a paper-and-pencil questionnaire to evaluate the tutoring sessions were less favorable in their evaluations of the

computer than the participants who completed the same questionnaire on the computer that had just instructed them. The questionnaire elicited significantly more varied responses as well, as had the second computer in the first study. The conclusion: The paper-and-pencil questionnaire, like a different computer, was perceived as an "other" that did not require a polite response.

Both studies used computers that showed only text. We wondered what would happen if the social presence were more explicit. What if the computers were fancier and even more suggestive of human presence? To answer this question, we decided to repeat the experiment using voices rather than text. We wanted to find out whether voices would accentuate politeness when compared with text.

In the voice experiment, the facts and other information were presented with human speech coming from a small speaker attached to the computer. All participants heard a single voice on a single computer that tutored and then praised itself. One group of participants heard the same voice on the same computer ask for an evaluation. A second group heard a different voice on a different computer request the evaluation. A third group gave their evaluation with pencil and paper.

We found the exact same differences with voices as we did in the earlier studies that used only text. When a voice on a computer asks about itself, people are more positive and less honest than when they are questioned by a different voice on a different computer or when they give their responses on a questionnaire. The conclusion: Users are polite to computers whether they use text or voices.

One final result tells something about how impressed we should be with the newer capabilities of media. The result is actually a *lack* of differences. Voices did not make the interaction any more social than text. The presence of voices was apparently no big deal, at least as far as creating a social presence. It doesn't take virtual reality to create the sense that another person is present; people don't need much of a cue to respond socially.

How Should One Think About Media?

People were polite to computers. Not only were the computers in these experiments tools for learning new information, they were social actors that people reacted to with the same polite treatment that they would give to another human. This certainly adds a new dimension to an understanding of human-media relationships.

Before radically altering how we think about media, however, a reasonable question is whether there is a chance that this conclusion is wrong. Several questions about the research could be raised, and our answers to the questions are particularly important because they apply not just to the present studies, but to virtually all of the research in the book. Here are some things to keep in mind.

First, it's important to remember that all participants did exactly the same things in the lab. Everyone received the same facts, the same test, and the same evaluation, and they used identical interfaces on identical machines. The only difference was which computer asked the questions. So the results must be attributable to that one difference.

In laboratory experiments, there is also a danger that participants will figure out what is being studied and then try to help by telling the experimenters what they want to hear rather than what they really think. This didn't happen in the computer experiments, however. When asked, none of the respondents guessed that the number of computers had anything to do with the experiment, and no one guessed that the experiment had anything to do with politeness. Everyone believed that the study was about how people use computers to learn.

What if participants believed that computers *really did* have human capabilities? To make sure this wasn't true, we selected subjects who would be least likely to hold this opinion—everyone in the experiments had extensive experience with computers. They all were daily users, and many even did their own programming. If anyone should have known that computers don't have feelings, they should have. Old brains, however, have not yet caught up with new media.

Another possibility is that people were merely impulsive. Many people occasionally yell at a newscaster or quarterback on television, or plead with a computer to give back a disk. These responses, however,

are instantaneous, and they are rarely sustained. In our experiments, the social responses lasted much longer than an instant—they characterized an entire learning session. Polite responses were related to the entire experience. Hence, social responses to media are more than impulses that punctuate more thoughtful moments.

Another question about the experiments is familiar in the humanities. Maybe people *willingly suspend disbelief* when they encounter media. Perhaps people make a conscious decision to "make believe," in this case, pretending that a computer is a person. In exchange, a user might be better able to understand a presentation (or in the case of entertainment, have a better experience). However, no one said that they were making believe that the computer was real just for fun or because it was helpful. So if there was a suspension of disbelief, it certainly wasn't *willing* or conscious. Indeed, it is *belief*, not *disbelief*, that is automatic.

One of the most interesting responses to the politeness studies is this: "I'd be polite to a computer, but I'm not thinking of the computer as a person, I'm merely responding to the person who wrote the computer program. And that person *is* real!" Perhaps the use of social rules is reasonable because the technologies are created by humans, and hence, they warrant human treatment.

This explanation can be ruled out for two reasons. First, no one in the experiments said that they were using social rules for *any* reason. Moreover, when asked specifically about whether they had considered the programmer when they made their evaluations, not one person said they had.

Is it then possible that people were *subconsciously* thinking about the programmer? This is not likely the case either. If there were a subconscious orientation to a programmer, then the people who used two computers would have had to think about two *different* programmers, one for each machine. However, when the people who used two computers were questioned about whether they thought the machines were programmed by the same person or different people, they all said that they assumed there was a single programmer (and they were right).

What Do Polite Responses to Media Mean?

We think that these experiments demonstrate, against the intuition of many scholars, and counter to the verbal reports of the participants, that social rules apply to media. In this case, the medium was a computer, but not one capable of virtual reality or any other obvious display of social presence. The computer showed plain text on a plain black-and-white screen. It is not necessary to have artificial intelligence or full-motion video to be social. The nerdiest of media, a computer that looks like it came from NASA control, is close enough to being human to trigger rich scripts for social interaction. Computers are social actors.

Social responses to media are not obvious, however, to those who exhibit them. The participants in the experiments denied that they had been intentionally polite to a computer, and we believed them. Instead, the responses occurred without conscious awareness. They were automatic and mindless. A significant part of the human brain works on unconscious responses, and that work is often completed without the results being available for analysis by those parts of the brain responsible for thoughtfulness and introspection.

At the broadest level, these studies demonstrate the viability of the media equation. Findings and experimental methods from the social sciences can be applied directly to human-media interaction. It is possible to take a psychology research paper about how people respond to other people, replace the word "human" with the word "computer," and get the same results.

Designing Polite Media

Initially, it may seem that these studies have more implications for what humans do with machines than for what machines should do with humans. Actually, the two are quite related. The reason is that polite behavior, as well as many other social behaviors, is part of an *interaction*. Social behaviors are not accomplished in isolation from the responses to them—social means reciprocal. This is pivotal for the design of interactive media, because the biggest reason for making machines that are

polite to people is that people are polite to machines. Everyone expects reciprocity, and everyone will be disappointed if it's absent.

When media violate social norms, such as by being *impolite*, the media are not viewed as technologically deficient, a problem to be resolved with a better central processing unit. Rather, when a technology (or a person) violates a politeness rule, the violation is viewed as social incompetence and it is offensive. This is why we think that the most important implication of the politeness studies is that media themselves need to be polite. It's not just a matter of being nice; it's a matter of social survival.

Grice's Maxims for Politeness

How can designers ensure that computers are polite? Again, our answer is to borrow from the researchers who study politeness—the social scientists. Perhaps the most general and powerful politeness rules that media could obey are *Grice's Maxims*. H. Paul Grice, a philosopher and psychologist, viewed conversation as an exercise in which people try to be helpful. Grice argued that *all* people feel that conversations should be guided by four basic principles that constitute the rules for polite interaction: quality, quantity, relevance, and clarity.

Quality. Speakers should say things that are true. This is the one Gricean maxim that computers obey pretty well. They may be insensitive in delivery or too quick to disappoint, but at least they tell the truth. It is important to remember, however, that accuracy is a shield and not a sword. Accuracy can breed frustration because of a perceived lack of cooperation. If someone stops his or her car to ask, "Where am I?" the answer "in a car" is accurate but quite impolite. In Grice's terms, the driver is annoyed not because the answer was inaccurate, but because the comment wasn't cooperative.

Quantity. Each speaker in an interaction should contribute only what the conversation demands, not too much or too little. This rule is frequently violated by interactive media. For example, most menu systems present a single word or at most two words for each option, and this is true no matter how complex the action. The result is that users

often feel that the program is not cooperating. Why is that computer not giving me the whole story? The outcome is frustration.

Can icons, the favorite exemplar of brevity, solve the problem of quantity? Not all the time. A single icon that represents a complex task can be just as frustrating as a single word. The use of plain English (full sentences or at least multiple words in logical phrases) would make an enormous difference in understanding and satisfaction. The success may be traded against time, but almost by definition, politeness takes time. In real life, most of us would choose politeness over brevity, even at work, and even in our most productive moments.

Another way to solve the problem of quantity is to use people's ability to elaborate abbreviated messages with information that they already have. Messages are often too much or too little for *someone*, but over time, people can learn that a short message stands for a larger response. As two people get to know each other, there are times when politeness can be abbreviated (e.g., "Hi" substitutes for "Hello, how are you today?" once people get acquainted). Familiarity can also bring opportunities to elaborate (e.g., "I've been meaning to tell you..." substitutes for a stifled "Fine"). A polite system will give information at a level of detail that matches the user's social expectations.

In the same vein, providing users with technical abbreviations (e.g., "Drive Error: Abort, Retry, Fail?") or pages and pages of detail violates the quantity rule. Much better are systems that allow users to set a level of sophistication that determines the amount of information that they would like. With a tracking system on the computer (i.e., a function that counts various occurrences), a computer could know how often a particular message has been delivered, for example, and adjust the quantity of information accordingly.

Relevance. What people (and media) say should clearly relate to the purpose of the conversation. A good example of this rule is the disabling of menu options, depending on context. An interface shouldn't say anything about things it can't do at the moment. Icons that represent possible actions could be highlighted, for example, and the icons for impossible actions dimmed or removed from the screen.

One aspect of relevance that is ignored in interfaces is response to user goals. The early days of television provide an excellent example of the consequences of this mistake. In early television, producers assumed that people watched the news to gather information. They thought that viewers were civic-minded, and consequently, the news was presented as seriously and efficiently as possible. However, research began to show what now seems obvious: People turn on the news for all kinds of reasons. These reasons include the desire to be entertained or merely to feel socially connected. The recognition that viewers had several different goals initiated several new concerns: the attractiveness of anchors, design of the news set, and enough "happy talk" to maintain interest.

The same thing is likely true for computers. Someone writing a letter, even on the most sophisticated word processing package, is likely to have multiple goals. One goal may be to complete a task; that is, write the letter. But it's hard to imagine writing a letter and not also doing some combination of the following: blowing off steam, clarifying feelings, impressing the boss, avoiding boredom, and so on. Why shouldn't computers modify interactions in relation to these goals? Interfaces that provide a *single* way of presenting information, without taking into account multiple goals of users, risk violating the rule of relevance. Anger and frustration could be the result.

Clarity. Contributions to an interaction should not be obscure. Designers often remove ambiguity so that a message can have only one meaning. This is desirable, but it comes at a price. To avoid ambiguity, highly technical language is often necessary, and much of that language is obthorship of the following sentence: "The coercive, mimetic, and normative forces in the institutional environment homogenize the garbage-can decision processes." One might argue that several paragraphs would have been required to achieve the same precision with more commonplace words. But that's silly. There wasn't a single reader who wouldn't have gladly traded precision for simplicity.

The upshot of this rule is that it would be better to have a statement

with even three meanings than to have one that is precise but unknowable. This is especially true if the ambiguity can be resolved in later exchanges. It would be worrisome if the user consistently resolved the ambiguity incorrectly, because the computer would then seem incompetent. User testing, however, can determine the most common way that ambiguity is understood; if most users resolve it incorrectly, it can certainly be rewritten. Furthermore, highly technical language, even if precise, can actually lead to *more* guesswork for users. Interactive media should not be obscure or ambiguous; but too often, interfaces have opted for the former.

A key point about Grice's maxims is that people will assume that violations have social meaning. If a speaker violates any of the rules, the listener will assume that the speaker is not paying attention, or is being sarcastic, or is being intentionally unpleasant. All of these conclusions lead to negative consequences for media, because people will ascribe meaning to failure, whether the entity that fails is a person or a machine.

It's Impolite to Reject

Media increasingly provide the ability to change how they look and work during an interaction. When you change something, however, you reject one option in favor of another. If the rejection is aimed at socially meaningless features, such as changing the color of the computer desktop, this is no problem. But when you change features that are more obviously social (e.g., a voice or a picture that represents a helper), the rejection is also more social. In the social world, rejection is significant; it is impolite.

Although it seems a bit weird, people can feel the same inhibitions when rejecting social representations on a screen as they would rejecting a real person. For example, imagine that you have been working with software that uses a character to help complete tasks on the computer. After a while, you become tired of the character, and you're aware that the software offers an option to change characters. Our research suggests that it is difficult to simply replace the old character

with a new one. People don't want to be impolite by making the current character feel bad.

What to do? Use the tried and true solution that works in real life. A polite invitation for change might go like this: "It's been really fun working with you, but some people like to change characters on occasion. Would you like to do that?" There are three advantages of this statement, and the accompanying question: (1) The statement legitimates the change, (2) it makes the decision impersonal, thereby limiting the need for a polite response, and (3) it doesn't reveal the feelings of the character asking the question.

Rules of Etiquette

There are many more popular sources for politeness rules than the psychology literature we have discussed so far. We recommend them all highly. We have both laughed mightily thinking about a bunch of computers dutifully taking notes while listening to Dale Carnegie or reading Emily Post and Miss Manners. But humor aside, that is exactly our prescription. If mediated and real-life conversations are more on a par than previously imagined, then media should be judged by their *social* as well as *technical* sophistication. Consider the following simple rules, even though they seem more suited to a handbook of etiquette than to a Computer Science 101 course.

It's Polite to Say Hello and Good-bye

How do you enter or leave a social situation? In any face-to-face conversation, people don't turn around and leave. First, they indicate intent and then ask permission to leave, at least implicitly. The opportunity to break this rule in media is legendary. In a famous interface project, a character suddenly disappeared from the screen due to a bug in the program. Users became disturbed, the designers noted, because they felt that the character was angry and had left as a result. Users did not view the disappearance as a problem with the technology. Characters that leave the screen should always take leave by saying

"good-bye" or at least making a sound or gesture. They shouldn't evaporate into the digital ether.

It's Polite to Look at People When Speaking

Humans, as well as many other animals, are sticklers for eye contact. Eyes are the number one place to look to size up a partner, understand his or her feelings, or predict what will come next. When people look at faces, half their time is spent watching eyes. When we can't see someone's eyes, we get worried, and this is likely why it is impolite not to show your face. The same is true for faces on a screen. Dan Rogers and Mr. Rogers understand this well. They never turn their backs, and instead stare right into the camera. Their counterparts in computing, however, often do not make eye contact.

A media character should never turn its back without an announcement, especially during an interaction. For example, in a prototype for a children's multimedia product, a character turned away from the user whenever it wanted to "see" something on the screen. This seemed reasonable to the designers, who intended the character to appear to look *with* the user at something important. But it was also impolite. To allay the user's discomfort, the character could have said, "Let's see what else we can find" before turning. Admittedly, this is another action in a domain where the fewest lines of code carry the day, but here as always, politeness does have some cost. The gain is more important: The user didn't have to wonder what the character was doing.

It's Polite to Match Modality: Answer a Letter with a Letter

It is polite to respond to friends using the same method that they used to contact you. A letter gets a letter in return, a phone call gets another call, and so on. We suspect that this applies to human-computer interaction as well. It is the rule of "matched modality."

There is asymmetry when a user receives information in one medium but answers in another, but in computing there often is no way to circumvent this mismatch. For example, some computer products ask questions verbally but accept only text or mouse input. This can be uncomfortable because the computer actually forces an impolite response.

What to do? If an interface accepts only text input, perhaps it should produce only text output. If the user can respond with voice, then a voice-based interface might work better. In any case, the criterion for choosing an appropriate way to respond should not merely be the most sophisticated mode available; it should be the one that allows for politeness between user and machine.

Politeness and Product Testing: Eliminating Positive Bias

The politeness studies also apply to product testing. One implication is that the same computer should not present products and then ask for evaluations. People who give opinions to the same computer that just demonstrated a product will likely react more positively than they really feel. It would be better either to use a paper-and-pencil questionnaire or simply to have another computer ask the questions.

A second implication extends the results of the politeness studies to real people associated with a product during testing. If an interviewer helps a person use a media product, the user will want to be polite to the interviewer as well. There is a good solution, however. First, two products can be presented, and the interview can focus on the differences between them. Not only is this a good measurement idea, since people are great at comparing things, but when focusing on the comparison, respondents are not thinking about what the polite response should be. This should encourage truthfulness.

A final point about product testing is that we should be suspicious of verbal responses. Many of the most important reactions and responses of users are those that are not conscious, and hence not available for verbalization. The people in our experiments assure us that they are not being polite to computers—but our data say otherwise. Subjects often do not know how they really feel or how they really will behave in a given situation. If what people said they wanted was what they actually liked, all of network television might look like public broadcasting. This ignorance is not necessarily a human deficiency; it is simply a human fact.

The Media Equation

Politeness across Cultures: Differences and Similarities

Most guides to good manners caution readers about cross-cultural differences in politeness, and international consultants earn a good living offering insurance against major *faux pas*. Cultural differences are certainly no secret to the companies that have a multinational business. They know that language translation is not the whole story. People also have to negotiate interpersonal space, wait the appropriate amount of time before talking, address each person in a proper manner—and the rules vary from country to country and from culture to culture. It is interesting, therefore, that the translation of language is often the *only* consideration for internationalizing media products. Mistakes in manners are a frequent result.

This point, while important, is not novel. What is not well understood, however, is how much importance we should attach to cultural differences. We think they are a bit overrated (which *doesn't* mean we think they are irrelevant). The differences are overrated mostly because they focus attention away from what is common to all human beings: *Everyone* is polite. This certainly does not mean that the specific behaviors that constitute politeness are exactly the same in every culture; they are not. But it does mean that everyone recognizes politeness, everyone tries to obey politeness rules, and everyone feels bad when they are broken. In a rush to celebrate cultural differences, we are often too quick to concentrate exclusively on those differences.

As children are taught by adults, being polite costs very little, and the benefits are enormous. Should we ask less of media and their makers?



Interpersonal Distance

When people interact, space matters. Spatial arrangements determine what people say, how they say it, and even whether it's necessary to say anything at all. The space between two people can determine the duration of an interaction and its emotional tenor. Everyone uses distance to tell other people what he or she thinks about them and what he or she thinks about the current situation. Not only is this true in all cultures, it's true for most species.

There are rules for interpersonal space. The parameters of these rules can vary across cultures (e.g., close proximity in conversation is more acceptable in Latin America than in the United States), but everyone knows what distance means. A polite response delivered at close range means more than the same from afar. In fact, it is hard to think of any social behavior that doesn't change in meaning, at least a bit, when the distance between people changes.

Social scientists have shown that violations of distance rules can cause all kinds of problems. This is especially true when strangers are the culprits. Unwanted advances induce anxiety, negative characterizations of the invader, and even hostility. If you don't like someone at a distance, it only gets worse up close. On the positive side, incursions into the personal space of another, when invited or desirable, can invoke everything from curiosity to the most extreme excitement that people can experience.

Interpersonal distance dictates the *intensity* of responses. Standing

The Media Equation

close turns up the volume knob and heightens concerns about personal pleasure and pain. Close is arousing. When arousal increases, people are more focused on the cause of the excitement, they are more attentive, and they remember more. There are also physiological preparations for possible action. None of these responses, however, are triggered by conscious control. Rather, people are built to attend automatically to things that arouse.

When people attend to a person who is close, they focus on what the person intends. The quickest way to make that judgment is to look at the face. By observing the eyes, mouth, and eyebrows of another person, people can assess the person's plans. When arousal is high, as when a person is physically close, people accent the processing of facial features. As a result, responses to strangers when they first approach often depend more on whether their faces appear threatening or friendly than on the words they speak.

What happens when people encounter mediated representations rather than real people? Does interpersonal distance still matter? Admittedly, this sounds far-fetched. We can all safely say that real people don't reside inside of media or on a screen, so it shouldn't matter whether *images* of people make them appear close or far away. If a pictured person appears too close, it is obviously because of the position of the camera, the size of the screen, or the position of the viewer. None of the pictorial cues that represent distance are cause for any action on our part. After all, it's only a picture.

Representing Interpersonal Distance in Pictures

Our prediction was that the media equation extends to expressions of manners even more primitive than politeness—it extends to the apparent physical distance between viewers and the screen. If people recognize polite and impolite computers, then the same automatic responses could cause people to bestow the same equivalence to pictures. Changes in the spatial arrangements between people and the pictures they view could be as social as the communication of manners through text and voices.

Interpersonal Distance

Replacing "a person" with "a picture of a person," we made the following predictions:

Rule 1: When viewers see a picture of a person who appears close rather than far, their evaluations of the person in the picture will be more intense.

Rule 2: Viewers will pay more attention to pictures of people who appear close rather than far.

Rule 3: Pictures of people who appear close will be remembered better than will pictures of people who appear far away.

Invading Personal Space with Pictures

To test the predictions, we conducted an experiment that resembled psychological studies done with real people and real spaces. In those studies, researchers vary the actual physical distance between subjects when they interact. This can be accomplished by having someone on the research team stand either close to or far from the participants, and then having someone else ask participants to describe their feelings about the interaction.

This is a tough simulation with pictures. Media can't actually walk up to someone and start talking. There are several ways, however, that a mediated person can *appear* to be close or far. The key to distance cues in pictures is visual field or the extent to which any image fills the space the eyes are capable of seeing. When something takes up a large portion of the visual field, one of two things must be true: The object must be close or the object must be large. Since all faces are approximately the same size, people will conclude that faces that take up a large portion of vision must be close.

A picture can take up space in the visual field in three ways. Viewing distance is the simplest. People on a screen could be perceived as near or far (i.e., fill more or less of the visual field) because of the physical space between the viewer and the screen. Just as one person can stand near or far from another, a person can view a screen image from near or far.

The Media Equation

In the case of television, viewing distance is determined by room architecture (distance constrained by walls) or personal preference (where you place your favorite chair). In the case of computers, the variance in viewing distance is more limited because most computing is done within arm's length.

A second way to vary perceived distance is the size of the screen that shows a picture. The same images on a large screen may appear closer than those on a small screen because the latter takes less space in the visual field. Modern media permit the exact same picture to appear on everything from a hand-held portable screen 2 inches high, to a wall-sized home theater with a picture from floor to ceiling, to an IMAX screen 40 feet tall. People can't easily correct for the size of a screen, because throughout evolution, small and large things were just that. Size was absolute. Only recently have humans developed ways to make big and little pictures that represent things substantially different from their real-life size.

A third way to vary perceived distance is the type of camera shot used to compose a picture. Faces appear closer when a camera frames a face so that the entire height of the screen includes *just* the face, as opposed to the entire body. Even when the distance to the screen and the size of the screen are the same, the camera has closed the distance to the subject in the picture. Close shots are compelling because more attention is focused on the eyes, mouth, and eyebrows—the most significant features in an interpersonal encounter.

Meeting People in Pictures Who Appear Near and Far

The goal of our media experiment was to allow people to experience each of the three methods for making people on a screen seem near or far. If pictures are like natural experience, then viewers should respond differently to faces on a screen depending on how close the people in the pictures seem. Faces that seem near should cause more intense evaluations, they should receive more attention, and they should be better remembered. These responses should be true even though the

Interpersonal Distance
faces are nothing more than fluorescent light rendered on a piece of glass.

We invited thirty-two men and women to participate in the experiment. One at a time, each person was asked to watch several faces in several combinations of screen size, shot size, and viewing distance. The distances and sizes in the experiment were selected to resemble, as much as possible, variations in pictures that people experience in the real world. The close viewing distance was 4 feet from the screen; the far distance, 10 feet. The large screen size was 41 inches; the small, 15. The close-up shot framed people from chin to top of head; the long shot showed people from head to toe.

Participants had a chance to see pictures in all combinations of the three parameters. To make sure everyone got a chance to experience each combination fully, we showed five faces in each of the eight combinations of distance, size, and shot-type. The people who appeared in the pictures were from 20 to 40 years old; half of them were men and half were women; and they were shown on the screen for 10 seconds as they talked about their favorite restaurant or their last vacation.

Measuring Responses to People in Pictures

We measured evaluation of the pictures as well as attention and memory. Evaluation was the most straightforward to measure. We merely asked participants to rate the people in the pictures using adjectives that described liking and personality. The questions began with "Does this person seem ...". Some of the items were *calm/anxious*, *pleasant/unpleasant*, *competitive/cooperative*, and *violent/gentle*.

The questionnaire wasn't appropriate for studying attention and memory, however. For these outcomes, we borrowed techniques from the researchers who study responses to real people.

Attention

There is not much to observe in humans that indicates the intensity of attention. Unlike rabbits, people don't perk up their ears when

The Media Equation

engaged. Asking people how much attention they gave to a message, especially when the question comes some time after the attention is given, doesn't work well at all. Self-reports about mental effort are tough even if people are willing to help. It's possible to know where people are looking by tracking their eyes, but beyond this assessment of selective attention, it is quite difficult to know how much *effort* they invest after the selection.

Psychologists have developed techniques for measuring attention that help overcome these problems, but the techniques are not obvious. One method is called the secondary reaction-time task. This measure assesses mental engagement without requiring people to talk about their behavior after the fact. To study mediated messages, we stole this method lock, stock, and barrel from psychology. Here's how it works.

When the mind is engaged in a primary task, it has trouble adding other tasks without compromising performance on the original one. For example, it's difficult to continue walking quickly after a friend poses a hard question, and sometimes, it's even hard to continue walking at all. We slow down or stop the primary task (walking) to perform the secondary task (answering the question). Conversely, when people are more intent on the primary task, it takes longer to address the secondary task. Compromises in performance are especially true when any one of these tasks becomes demanding. People can allocate their mental effort in many different ways, but they can't increase their total energy beyond some finite point. Mental capacity is *limited*.

The secondary-task method takes advantage of this limitation. Participants are invited into the lab, and they're asked to watch something on a screen. They are asked to pay close attention to the screen because they will be questioned about what they see after they view. This ensures that watching television is the primary task. We then tell participants to listen for audio tones during the presentation, and this becomes the secondary task. When viewers are highly engaged in the primary task (watching the screen), they have less capacity left over to process the audio tones (the secondary task).

Participants are told to press a button on a joystick as quickly as

Interpersonal Distance

possible when they hear one of the tones. The tones occur at random points in the television program, about one every 15 seconds. A computer then records the time, in milliseconds, that it takes to respond to the tone. The resulting data are reaction times to the secondary task, hence the name of the measure. When participants are paying attention to the video (the primary task), their reaction times to the tones are slower because they have less mental effort left over to process the tones. When reaction-time scores are slow, we conclude that attention to the screen was high.

This measure is not a standard one in research about media. One reason is that it's a bit of a nuisance. You have to hook up a reaction-time device to a computer. Also, there are some tricky statistical issues which complicate the analysis of data. Consequently, it's important to know the rationale for the extra effort.

The biggest advantage of the secondary-task method is that attention is measured with a system that does not require introspection. Viewers do not have to be *aware* that they are pressing the button slower or faster—and in fact, they are not aware. People don't think that their speed pressing the button is related to their attention to the screen. Even experienced researchers have little sense of when their own response times change, and they find it difficult to "beat" the system even though they know how it works. When something interesting is on the screen, reaction times reliably get slower—and this rule seems to apply to everyone.

Memory

Memory is a huge concept in psychology, and it has many definitions and measures. One of the most important distinctions in memory is the difference between recall and recognition. Recall refers to the ability to access information in our minds without help. When information is recalled, it is dredged up merely through thought. We recall the name of a friend (without any help); we recall the route to her house (without looking at a map); and we recall what she said last night (without any reminders).

Recognition memory is different. It is the ability to match something that a person sees, hears, or reads with a previous experience. Even if you can't recall a name, you'll *recognize* it when you hear it. You can't recall the directions to a house, but you'll recognize the street when you pass it. You can't recall what someone said, but if you hear a snippet, you can construct the rest of the conversation.

In this experiment, we focused solely on recognition memory. (Recall is too difficult to study for facial memory because there is no good way for a viewer to demonstrate recall short of a long written description or by drawing a picture.) After participants had finished viewing the faces, we measured recognition memory with a standard test. We quickly flashed pictures of several faces on the screen, and asked participants whether they had seen each face during the experiment. Of the thirty-two faces in the memory test, half of the faces had been seen earlier; the other half were new pictures of faces that viewers had never seen before. The score for each participant's recognition memory was the number of faces that he or she could identify accurately.

A Viewer's Personal Space Can Be Invaded

The first social rule that we thought participants would apply to pictures was that evaluations of the people in the pictures would be more intense when the faces seemed close. That is indeed what happened. The close faces (especially those framed with a close shot), were evaluated more intensely than the faces that seemed farther away. For both positive and negative evaluations, scores were more extreme for close images than for distant images.

Attention and memory were also enhanced when pictures seemed close. Participants paid more attention (had slower reaction times) to the close faces than the far ones. This was especially true for faces on the larger screen and for faces framed with a close shot. On average, viewers took about 10 percent longer to respond to the tones when the pictures were close, which is a significant difference with reaction times. Just as in human-human interaction, participants focused more attention on faces that seemed close.

Recognition memory was also better for the close faces, especially those that were viewed from a close distance and faces framed with a close shot. On average, participants accurately identified 12 percent more of the faces that seemed close.

Other Explanations

The results of this experiment show that mediated manners extend well beyond the manners that people are taught. People have a primitive or innate interest in the physical arrangements that define an interaction, and they respond automatically, even to pictures.

The responses do seem weird, however. These are, after all, only pictures. Is there any room for skepticism? It may seem so, but there are several points that should lessen doubts about these results.

First, the experimental design used in the study ensures that all participants have exactly the same experience when they come to the lab. The only difference in the presentations related exactly to what we were studying—the apparent distance between viewer and picture. We can also safely say that no one in the experiment guessed what the study was about and then changed his or her responses to please the experimenters. Even if people could guess, it would have been difficult for them to control their reaction times precisely.

Could media experts have done better or worse than media novices? Likely not. When it comes to viewing video, virtually everyone is an expert. All of the participants in the study had ample experience with television and film, and everyone knew that all of the faces shown on the screen were, in real life, approximately the same size. Social responses don't depend on instruction, nor can they be extinguished with experience.

One critique of studies like this one is that the results might be related to the *particular* content presented. For example, it is likely that some of the faces shown in the research might have been more arousing or memorable than others, regardless of how close they appeared in the experiment. This would be similar to liking a movie with one actor, but hating it if a different actor were substituted.

The Media Equation

How can we be sure that the characteristics of particular faces didn't account for the differences in the present study? We presented a lot of *different* faces to each participant in the study, and we mixed and matched faces with viewing distance, screen size, and shot type. In this study, it meant that we collected a total of forty different faces, and we randomly picked five different faces to show each person in each of the different conditions. Essentially, we created a unique video production for each participant. This is a burden, but it pays off. The results of the experiment are not attributable to the particular faces that were shown.

The Media Equation Extends to Physical Space

What does all this mean for theories about media? Several things. Once again, human responses to media are social even though media supposedly offer only symbolic representations of people in the real world. The creation of space, even if accomplished with illusory tricks possible only with modern technology, is consequential in media as well as in real life. There is no "discount" for pictures.

This evidence also shows that social responses to media are not limited to *feelings* about the technologies themselves. When people were polite to computers, those responses, while quite social, were particular *attitudes* toward the machine. Changes in the perception of physical distance, however, determine very different psychological responses—from the mental energy required to pay attention to the ability to remember what had been seen. Social responses to media are part of a complex web of influence that extends from attitudes and evaluations to attention and memory.

This study also suggests that three different visual tricks, most often treated *separately* by professionals, are really quite similar. Viewing distance is controlled by the viewer (or constrained by the viewer's environment). Screen size is a characteristic of the appliance (and is determined when it is purchased). And shot type is controlled by the director (at least until we have TVs that allow us to zoom in on objects of interest). From the point of view of industry, then, the three distance

Interpersonal Distance
cues are radically different, and it would be surprising to find these three techniques grouped together in any discussion of media. But from a psychological perspective, as demonstrated in the present study, the three cues *should* be grouped, because they all elicit similar responses from viewers.

Implications of Distance Cues for Making and Evaluating Media

If the distance cues in pictures make us think that people are actually close or far, how should people think differently about how media are evaluated and created? The most important implication is that people presented in media are perceived as actually present. This perception is not only a primitive reaction; it is influential across a broad range of psychological life. Viewers evaluate faces on a screen, and they prepare to respond to the faces, in the same ways that they would for actual people. This is true even though the responses may seem improbable to the people doing the responding.

Implications for applying this part of the media equation to several important media issues are discussed in the next paragraphs.

Meeting People through Pictures versus Words

Pictures of faces, because they are processed as if they were people in actual interpersonal encounters, may be more different from words than previously imagined. It may be impossible for a person to appear in a picture without the baggage of primitive social cues, and with the good chance that those cues will influence how viewers respond to the pictures. People are not able to ignore pictorial social cues any more than they can ignore the same cues in real life. It's no wonder that it is difficult to concentrate only on what is said.

Consider this example of interpersonal distance at work on television. In a recent presidential debate, one of the candidates answered a question by saying that he would pledge to try and meet the needs of voters by emphasizing political issues, not images. The other candidate responded that he would also make the same pledge, and he added a

The Media Equation

personal comment: He was sick of having to start every day by figuring out how to defend himself.

Imagine reading a transcript of the exchange. It's possible to read spoken comments free of any influence from the physical situation in which the comments were generated (e.g., who stood where and how close they were to each other and to the audience). A video of the same interaction, however, adds a thousand words. The visuals contain a lot of information about the *social* arrangements between the speakers, and between the speakers and those watching at home.

The exchange in the debate started with relatively long shots of both candidates, from their waists to the top of their heads. The second speaker, however, didn't stand still when he talked. When he gave his personal comment, he moved toward the camera. Essentially, he moved directly toward every single person watching him on television. This led to a more tightly framed shot of the candidate for the remainder of his time on camera. The added sense of presence almost demanded that the viewer form some opinion about the intrusion, and since the candidate was smiling and speaking softly at the time, we can assume that many of the opinions were positive.

If we apply to the video of the debate the experimental results regarding attention and memory, we could also assume that people paid more attention to the speaker who advanced toward the audience (or at least more attention to his physical presence, if not to his words). We can also assume that viewers would better remember his picture. In contrast, the speaker who remained still, even though his picture certainly contained more information than could be inferred from reading the words he spoke, was visually much less present.

A minor point concerns persuasive strategies in a televised debate: To engage an audience, occasionally move toward it. Moving closer will work even if the audience is in front of a computer or TV screen. To stand still is to relinquish audience attention to other parts of the stage.

A more important implication is that pictures, especially pictures of people, add a lot to linguistic expression. Pictures focus attention on primitive social cues, potentially to the detriment of other information.

Interpersonal Distance

Pictures may help with memory, but at the expense of taking away time for people to decide what things mean. It's difficult to think about how a candidate stands on an issue if you have to simultaneously monitor his physical presence in your personal space. The effort needed to process his presence and his comments comes from a single and limited reservoir of mental energy.

Pictures, because they look like the things they represent, require less mental effort to translate between referent and reality. Pictures give information that is more familiar and easy to process, and this may promote more passive thinking. It is much easier for people to determine if they like a speaker based on how close the speaker stands than it is to actually listen to what the person has to say. Should it be a surprise, then, that any citizen, regardless of political sophistication, feels quite comfortable judging the personal qualities of a candidate once he or she has seen the candidate's picture?

The same process likely applies beyond politics. People feel quite comfortable thinking about media personalities as friends, confidants, and villains as well as doctors, detectives, lawyers, or other characters they portray. The basis for all of these relationships may be the primitive and real experience of another person's presence, even though it is pictorial. We can laugh at ourselves for the apparent gullibility, but our brains respond otherwise. Consider an extreme case: Would you feel comfortable naked in front of an anchor on the evening news? If not, why not?

Which Picture of Which Face?

Imagine that the experiment with faces and distance had been done with a single picture of one person. And imagine we found the same results. Could the study be easily criticized? The answer is yes. Our research, and most other studies that use pictures of people (or just about anything else), find *different* results for different pictures. There may be an overall trend, but the results for any one picture could easily be different from the rest.

This means that *casting* is critical. There are considerable differences

The Media Equation

between faces, differences that are determined by a hundred important qualities besides the distance of the face from a viewer (e.g., attractiveness, facial expressions, age, gender). If you produce media, however, you can present only one face at a time. Therefore, choose faces carefully.

Casting can obviously affect the success of a film or television show. Its influence is less obvious, however, for new media. When characters are used in computer interfaces (for example, as agents who help with the tasks of computing), the most likely first question is whether the character makes the experience better than the same experience with no character at all. This may not be the best question, however. Almost always, there will be *some* characters that are better than no character, and *some* characters that are worse. The better question is *which* character is best. The answer can come only from a screen test, no less appropriate for multimedia and computing than for film or television.

Scaling Physical Space

Directors in film and television usually have a good idea about how people and objects will be displayed to their audiences. That's useful, because media can be designed, even if unconsciously, to conform to the interpersonal arrangements of a given display. Movies are on big screens, television sets are much smaller, and there are unofficial rules about how to shoot and frame people within each medium. The biggest problem so far has been the need to create material that can be shown on a small television *and* on a large screen in a theater.

The problem of scalability is increasing. Moreover, production for newer technologies requires even greater thought about display formats, because size and shape can vary tremendously. Today, the same video data can be displayed in a two-inch window on a computer monitor or on the very large screen of a home theater. For the smaller picture, a full head shot might create a sense of intimacy. For the home theater, the camera might be better placed at a distance, with the people in the pictures framed head-to-toe. As one producer of IMAX films commented, Mick Jagger's face, 20 feet high, is frightening: Best to use a long shot when the screen is huge.

Interpersonal Distance

One solution to the problem of multiple displays might be to allow viewers the chance to frame their own shots. Viewers could change where they stand or sit relative to the people in the picture by manipulating the visual signals. If viewers wanted to get closer, they could zoom in on the picture or change the focal length of the lens that rendered it. If people wanted to withdraw, they could make a person appear smaller by reversing the process. This flexibility could enhance the feeling of lifelike interaction.

Face Size across Media

The power of approaching and withdrawing is not lost on film and television directors; they use it with great skill. Over the last thirty years, they have increasingly closed in on faces when the display of emotions is critical. When an interviewer has a con man on the ropes, for example, or when emotion is at its height, faces are framed to fill the screen.

Almost since its introduction, television has been steadily moving in on the personalities who deliver news. Close is intimate, intimate is engaging, and engaging means more viewers. An effective method in all of these examples is to exaggerate closeness by *changing* the distance between viewer and viewed—start with a long shot, and then slowly move closer.

People assume that a picture of a face, regardless of its size, is merely a symbol that stands for someone not actually present. But it is more. The size of a face can broadly influence psychological responses—from the mental energy required to attend, to the ability to remember what was seen, to thoughtful judgments of character.