1

Introduction to Implicit Assessment

Whether you are opening your mail, answering your phone, or browsing the Internet, the odds that you will be asked to report your attitudes toward a product, a politician, or a social issue are high. Given the plethora of opinion surveys confronting the average citizen on a daily basis, she or he might reasonably presume that measuring attitudes is a snap. Simply provide clear, precise questions and a scale to respond with (often ranging from 1 = *strongly disagree* to 7 = *strongly agree*), then crunch the numbers. Easy, right? Nothing could be further from the truth.

The problem is getting at the truth. Unlike geologists, attitude researchers cannot whip out a measuring tape and wrap it around a rock. Rocks have the enviable property of not shifting around when you measure them. Attitudes are mental constructs, not tangible things, so measuring them is always an inferential endeavor. You cannot peer inside people's heads to "see" how they evaluate something. Even if you could, attitudes are not stored away in a mental drawer (like a pair of socks), to be taken out when researchers ask how you feel about X. They are slippery and they shape-shift, depending on context. "Context is king" when it comes to attitudes, meaning they can be altered by systematic factors, such as how the questions are framed and what order they come in, as well as random factors, such as people's moods, the weather, and current events (Eagly & Chaiken, 1993). All of which can make attitude assessment agonizingly hard to achieve. The difficulty mounts when you consider that, until recently, researchers have had to take it on faith that what people report on a questionnaire reflects their true attitudes. But when people have control over their responses (by, say, circling a number), two immediate concerns arise, dubbed the "willing and able" problem. First, people may not be willing to report their honest opinion; and second, they may not be able to introspect adequately to surmise what their attitude is (Nisbett & Wilson, 1977; Wilson & Dunn, 2004).

The fact that people can edit (or distort) their explicit attitudes has long made attitude researchers wary of taking people's self-reports at face value, particularly when the topics being considered impinge on people's morality (Crowne & Marlowe, 1960; Gaertner & Dovidio, 1986; Paulhus, 1984; Dovidio & Fazio, 1992; Thurstone, 1928). Attitudes toward behaviors that are illegal (e.g., stealing

and drug use) or immoral (e.g., lying and cheating) are prominent examples, as are attitudes toward anything having to do with sex or religion. Because we are social creatures, it is human nature to present oneself in a manner that will be viewed favorably by others.

Similarly, topics such as prejudice and discrimination (e.g., attitudes and behaviors toward minority groups) have become moral issues. In the United States, legislative changes during the 1960s and 1970s outlawed discrimination based on race, gender, age, ethnicity, and religious orientation. It became illegal, as well as immoral, to discriminate against people based on their group membership. As a result, scores on explicit (i.e., self-report) measures of prejudice have steadily decreased (Judd, Park, Ryan, Brauer, & Kraus, 1995; Schuman, Steeh, Bobo, & Krysan, 1997), while normative pressures to be egalitarian have increased (Dunton & Fazio, 1997; Plant & Devine, 1998). In fact, many people sincerely believe they are not biased (Pronin, 2007). At the same time, Americans are inundated with cultural messages that, for example, people of color are relatively poor, uneducated, and more likely to be in trouble with the law. These messages are likely to permeate our mental apparatus even when we refuse to endorse them (Devine, 1989).

To circumvent these problems, social psychologists have devised innovative techniques to measure *implicit attitudes*, that is, attitudes that people may not be aware of, or that they are unwilling to report. The most advanced techniques rely on response latencies (i.e., reaction times) when people perform various tasks, rather than deliberate responses. The researcher does not ask people what they think or feel. Instead, people's attention is focused not on the attitude object, but on performing an objective task; attitudes are then inferred from systematic variations in task performance (Cook & Selltiz, 1964). Collectively known as *implicit measures*, response latency methods solve the willing and able problem because (1) people are less able to control their responses and (2) they can reveal attitudes that people may not even know they possess.

The ability to measure attitudes and beliefs in ways that bypass deliberate, and often distorted, responses has afforded remarkable new insights into the human mind and spawned a new discipline: implicit social cognition. Because of their advantages, implicit measures have been widely heralded, and popularly used. Searching PsycINFO for the two most widely used implicit measures (evaluative priming and the Implicit Association Test) in the title, abstract, or keywords revealed over 890 results.[1]Because they represent a state-of-the-art assessment tool, they are an important topic for behavioral scientists to learn about.

Goals of the Implicit Measures Volume

The primary objective of this volume in the series is to teach nonexperts how to use implicit measures in their own research. To do this, I will take an approach

that is more practical than theoretical, with the aim of answering such basic questions as: how do you design and validate such a measure? What are the best practices to avoid common errors? How do you interpret and report the results? How have other researchers effectively used implicit measures? The goal is that after reading this volume, you will be able to build and administer your own implicit measures. You should also be able to use this volume as a reference as your research progresses.

In this volume, I will focus on the two most prominent implicit measures: evaluative priming and the Implicit Association Test (IAT). Although there are many differences between them, each employs reaction time tasks that measure people's attitudes indirectly. There are many other types of implicit measures, but evaluative priming and the IAT have received the lion's share of research attention and both have shown the ability to predict behavior (i.e., they yield *predictive utility*). Because behavioral scientists are interested in accurately predicting human behavior, predictive utility is the "gold standard" by which any new assessment technique is evaluated. However, it is not the only kind of validity, and evaluative priming and the IAT have also shown substantial *known groups* validity (i.e., they distinguish well between groups that are "known to differ"). For several types of behaviors, particularly those that impinge on morality, evaluative priming and the IAT have shown better predictive utility and known groups validity, compared with self-reports (for reviews, see Fazio & Olson, 2003; Nosek, Greenwald, & Banaji, 2007; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Finally, the underlying processes that drive their effects are likely to be similar, albeit not identical. For these reasons, they were chosen as the best candidates for this volume.

Basic Terminology and Assumptions

To begin, a brief discussion of basic terminology and assumptions is needed to provide some background. First, an *attitude* is a psychological tendency to evaluate a given object with some degree of favor or disfavor (Eagly & Chaiken, 1993). Second, an *attitude object* is a broad term that encompasses physical objects but also anything that can be evaluated. The self, others, specific people, groups of people, social issues, situations, and goals are just a few examples. Even attitudes can serve as attitude objects (e.g., attitudes toward prejudice). Third, if you ask people how they feel about X, you are using *explicit* measures (a.k.a. self-reports, surveys, and questionnaires). By contrast, if you do not ask people directly how they feel, but instead infer their attitudes on the basis of how they behave or perform a task, you are using an *indirect* measure (Cook & Selltiz, 1964). A classic behavioral example is measuring how far away people choose to sit when they are told they are going to interact with someone (e.g., of a different race: Bogardus, 1927; Goff, Steele, & Davies, 2008). Fourth, if you use an indirect technique that

involves measuring response latencies (the speed with which a task is performed) in a manner that cannot be easily controlled, you are using *implicit* measures, the topic of this volume. Attitudes that are measured using response latency techniques are called *implicit* (or automatic) attitudes. By extension, any other construct that is measured using response latencies is referred to as implicit (e.g., implicit stereotypes, self-concept, and self-esteem).

Defining Implicit Attitudes

Implicit attitudes can be defined as associations in memory between objects and evaluation that are routinized to the point of being automatically accessed in the presence of the attitude object (Fazio, 1990). This definition applies equally to many explicit attitudes, if they are sufficiently strong, and it captures the assumption that attitudes in general are learned through experience, either directly (by encounters with the object) or indirectly (by exposure to information about the object). The key to this definition is that once learned, the attitude is spontaneously activated when the object comes into view, or simply by thinking of the object. However, there is a gap between people's attitudes and how they are expressed that prevents researchers from perfectly assessing either implicit or explicit evaluations. This gap may be wider for explicit attitudes because people can easily edit themselves when they report their attitudes, whereas they cannot edit their automatic associations. People can also second-guess how they "really feel" about something or someone on self-reports. They might also genuinely endorse different attitudes than their automatic associations would reveal. However, even when people are truthful, self-reports can *only* reflect what people believe about their attitudes, whereas implicit measures bypass this limitation. Although this analysis implies that implicit attitudes are more valid, this is far from the case. It is likely to be true when explicit attitudes are deliberately distorted, or when people are unable to accurately access their implicit attitudes in order to report them. However, all measurement strategies are subject to error and context effects, and this is certainly true of response latency methods as well as self-reports.

But let us imagine that you have reasonably valid instruments for measuring implicit and explicit attitudes toward the same object, and you discover they produce uncorrelated results. Which instrument should you trust? In many cases, it is entirely possible that both implicit and explicit attitudes are legitimate, but that they stem from different sources of information (Rudman, 2004). For example, people may have a set of beliefs that they sincerely endorse while simultaneously possessing vestiges of "old beliefs" that they may have initially learned (e.g., as a child, before they were able to challenge them) or that they have been routinely exposed to through their cultural milieu (Devine, 1989; Greenwald & Banaji, 1995; Wilson, Lindsey, & Schooler, 2000). This characterization has often been

applied to implicit racial stereotypes and attitudes, which are often weakly correlated with self-reports. But beyond race-related concepts, there is growing evidence that implicit and explicit attitudes are distinguishable by their sources, and not merely by the methodologies used to obtain them. For example, some implicit attitudes stem from developmental experiences that are emotional in nature, whereas explicit counterparts reflect more recent events (e.g., implicit but not explicit gender attitudes reflect maternal attachment, Rudman & Goodwin, 2004; and implicit but not explicit attitudes toward smoking reflect childhood experiences, Rudman, Phelan, & Heppen, 2007). Further, implicit attitudes toward groups are influenced by social hierarchies, such that members of high status groups automatically favor their ingroup more so than members of low status groups; the opposite pattern is more commonly observed using self-reports (Jost, Pelham, & Carvallo, 2002; Rudman, Feinberg, & Fairchild, 2002). There is also evidence that implicit attitudes are more difficult to change, whereas explicit attitudes are more readily updated (Gregg, Seibt, & Banaji, 2006; Rydell & McConnell, 2006; Smith & DeCoster, 2000; Wilson et al., 2000). It may also be the case that explicit attitudes are capable of being more objective and nonpartisan, compared with implicit attitudes (Gawronski & Bodenhausen, 2006; Rudman & Phelan, 2009; Strack & Deutsch, 2004). In other words, implicit attitudes may be more impulsive and affective in nature. From this point of view, explicit and implicit attitudes are equally legitimate, but they may reflect different learning experiences or different facets of an attitude object. Rather than replacing explicit attitudes, it is better to think of implicit attitudes as providing another level or aspect of evaluations that often conflict with explicit attitudes but nonetheless influence people's judgments and behavior (Banaji, Nosek, & Greenwald, 2004; Greenwald & Banaji, 1995; Wilson et al., 2000).

Are Implicit Attitudes Nonconscious?

Implicit attitudes are thought to be automatic not only because they are fast acting, but also because they can emerge (1) without intention (i.e., are involuntary and not readily controlled) and (2) outside of conscious awareness (Bargh, 1989; 1994). For this reason, implicit attitudes have also been described as nonconscious (e.g., Blair, 2001; Quillian, 2008).

Research supports viewing implicit attitudes as involuntary; efforts to motivate people (e.g., with cash incentives) to alter their scores on implicit attitude measures have largely been ineffective, suggesting responses cannot be easily faked (e.g., Banse, Seise, & Zerbes, 2001; Kim, 2003; Egloff & Schmukle, 2002). But whether implicit attitudes are nonconscious is a point of debate (Fazio & Olson, 2003; Gawronski, LeBel, & Peters, 2007). It is certainly the case that people are often surprised when their implicit attitudes deviate substantially from their explicit attitudes, suggesting they were not privy to their automatic preferences.

Nonetheless, we cannot be sure that implicit attitudes are nonconscious because we cannot measure people's awareness independent of asking them (Greenwald, Banaji, Rudman, Farnham, Nosek, & Mellott, 2002). A means of investigating the issue involves testing theories based on nonconscious processes using both implicit and explicit measures. When implicit (but not explicit) responses support the theory, we can infer that implicit attitudes are nonconscious (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Hafer, 2000; Jost et al., 2002; Rudman et al., 2002). Indeed, when people are truly unable to access their automatic associations, only implicit measures can detect them. Nonetheless, because using the term "nonconscious" to describe implicit attitudes is rightly controversial, I will refer to them as "implicit" or "automatic" throughout this volume.

Although researchers cannot know if the contents of people's minds are consciously available to them (or not), there is another definition of implicit attitudes that focuses on their origins and how they function, as opposed to their contents. Greenwald and Banaji define implicit cognitions as "traces of past experience that affect some performance, even though the influential earlier experience is not remembered in the usual sense – that is, it is unavailable to self-report or introspection" (1995: 5). According to this view, implicit attitudes stem from forgotten experiences (their source is not consciously available) and they can *operate* nonconsciously, leaking into people's judgments and actions without their volition or awareness. This means that you may be able to recognize you have a positive or negative implicit attitude toward someone, but that at the time you were behaving, you thought you were acting objectively. For example, a professor may grade a paper from a student she is fond of more leniently than another student's paper and not realize her grading is biased (which is why it is best to use blinding procedures when we evaluate someone else's work).

A dramatic example of how implicit biases can operate is seen in orchestra auditions. After orchestras in the US began auditioning musicians behind a screen, masking the gender of the performer, there was a substantial increase in their hiring of female musicians (Goldin & Rouse, 2000). Because it is unlikely that female musicians suddenly improved their talent, or that decision makers were deliberately sexist, it appears performance evaluations were tainted by implicit gender bias. Greenwald and Banaji (1995) stress that what is nonconscious is how implicit attitudes operate, not necessarily the evaluation itself. The people who evaluated female musicians without the masking technique may have been aware that they believed men were better musicians, but they likely assumed their hiring decisions were not biased by such views. Most people believe that they treat others fairly, and that they are treated fairly in return.

Similarly, men exposed to sexist television ads subsequently rated a female interviewee as less competent and treated her as a sex object, relative to men in the control condition (Rudman & Borgida, 1995). During debriefings, the men in the sexist condition were adamant that the ads did not influence their behavior.

Are men unaware that women are often treated as sex objects by the media? Of course not, but they may be quite unaware of how such portrayals can influence their own behavior. The repeated linking of women with sexual gratification in our culture likely creates an automatic association in men's minds that can be activated (or *primed*) through recent exposure. In turn, priming this association can then affect how men behave toward a female job candidate without their knowledge or intention.

Finally, it should be noted that Greenwald and Banaji's (1995) definition of implicit attitudes relies on a classic distinction between using one's memory as an *object* or as a *tool* (Jacoby, Kelley, Brown, & Jasechko, 1989). For example, if you are asked, "What did you do last night?" your memory becomes the object of your attention. But if you are asked, "Did you have fun last night?" your memory becomes a tool to guide your response. In the second case, your attention is on whether you enjoyed yourself and your memory operates in the background. Similarly, if you are asked, "How do you feel about old people?" your attitude toward the elderly is the object of your attention. But if you are asked to evaluate an older person's job performance, your attitude toward the elderly could function in the background, as a tool to guide your opinion. For Greenwald and Banaji, implicit attitudes act like tools, whereas explicit attitudes are those we express when our attention is focused directly on them. One reason why implicit attitudes permeate our judgments even when we are trying to respond objectively is because we are seldom aware of their influence. By contrast, when people are aware that their implicit attitudes might bias their opinions or actions, they can work to overcome them.

Whether implicit attitudes themselves are nonconscious or typically operate nonconsciously is an issue for theoreticians to sort out. The good news for researchers is that they can be measured (Banaji, 2001) and that they often predict behavior better than explicit attitudes (Fazio & Olson, 2003; Greenwald et al., 2009). For the remainder of this volume, we will see how measuring implicit attitudes and beliefs is accomplished.

Note

1 Conducted on 6 April 2010. Evaluative priming was also searched for using the terms "affective priming" and "sequential priming" because these are often used synonymously.