3.

Indirect and Implicit Measures of Religiosity

JONATHAN JONG, BONNIE POON ZAHL, AND
CARISSA A. SHARP

The Problem of Psychological Measurement

The fundamental problem in psychology as a scientific discipline is that of measurement. How can we access (and assess) what people are feeling and thinking? The most commonsensible answer to this question is, of course, to *ask* them. We ask people questions and make inferences about their mental states from their verbal responses, either spoken or written. We conduct opinion polls, structured interviews, and focus groups; we construct psychometric scales and other kinds of questionnaire-based measures. It is no exaggeration to say that most of what we know—*or think we know*—in the social sciences comes from self-report data, either taken at face value or interpreted through some theoretical or hermeneutical lens. The social scientific study of religion is no exception to this generalization. It is, and has been for over thirty years now, a truism in the psychology of religion that there is a preponderance of self-report measures of religiosity (Gorsuch 1984). To this point, Hill and Hood (1999) managed to compile a list of 126 self-report measures; and many more have since been published (Hill and Pargament, this volume).

Well-formed questionnaires are very useful tools in a psychologist's repertoire; at the same time, however, researchers have also long acknowledged the limitations of such *direct* or *overt* or *explicit* measures. We begin this chapter with a brief overview of the limitations of self-report measures of religiosity. We then describe and review a number of *indirect* or *implicit* measures of religiosity that have been designed specifically to circumvent these limitations. Finally, we conclude with recommendations for how these indirect or implicit measures may be

adapted for use by social scientists to supplement more traditional measures of religiosity.

Self-Report Measures and Their Discontents

There are, broadly speaking, two classes of problems with self-report measures. The first is the "strategic responding" (Wittenbrink and Schwarz 2007:2) problem: participants might not always be honest when asked directly about their feelings, beliefs, and desires. Indeed, people also often misreport their behavior: for example, sociologists have consistently found that American Christians overreport their religious activity, such as religious service attendance (Brenner 2011; Brenner, this volume; Hadaway and Marler 2005). The standard interpretation of such trends is that people bias their responses in *socially desirable* ways, either because they want to manage others' impressions of them or because they are engaging in self-enhancing self-deception, or both. Indeed, researchers have linked social desirability tendencies to religiosity, finding that religious people—particularly intrinsically religious people (Allport and Ross 1967)—have a greater tendency toward socially desirable responding (Gillings and Joseph 1996; Leak and Fish 1989; Sedikides and Gebauer 2010; Trimble 1997; see also Brenner, this volume). Furthermore, this relationship between religiosity and socially desirable responding is particularly strong in contexts that place a higher value on religiosity (e.g., USA > UK; Christian universities > secular universities). To complicate matters, explicit self-report measures might also serve as *demand characteristics*: cues that lead participants to respond on the basis of beliefs that they form about researchers' expectations (Orne and Whitehouse 2000). When participants think (rightly or wrongly) that they know the purpose of a study, they may change their responses in order to affect the outcome of the study. This is one reason why psychologists are often reluctant to inform participants of the true purposes of any study.

The second class of problems is that of the limits of introspective access: even if we can guarantee honest responses from participants about their *explicit* attitudes, there might be *implicit* attitudes, or aspects of attitudes, of which participants are not aware, and therefore *unable* to report.

Over the last two decades, the view that there are aspects of human psychology that are not fully accessible to introspection has established itself as orthodoxy among social and cognitive psychologists. The literature is now replete with dual-process models of cognition, which variously distinguish between the implicit and explicit (e.g., Nosek 2007), or the automatic and controlled (e.g., Bargh and Chartrand 1999), or the unconscious and conscious (e.g., Dijksterhuis and Nordgren 2006), or the heuristic and systematic (e.g., Chen and Chaiken 1999), or the associative and the propositional (Gawronski and Bodenhausen 2006). This theoretical interest in dual processing has come with methodological developments: the past twenty years have seen remarkable growth in the number of *indirect* and/or *implicit* measures, as well as studies using such measures (see Sherman, Gawronski, and Trope 2014; Chaiken and Trope 1999; Wittenbrink and Schwarz 2007). Explicit and indirect (or implicit) measures of attitudes have been shown to independently predict behavior (Greenwald, Smith, Sriram, Bar-Anan, and Nosek 2009), to predict different kinds of behaviors (e.g., verbal v. nonverbal; Dovidio, Kawakami, and Gaertner 2002; Dovidio, Kawakami, Johnson, Johnson, and Howard 1997; spontaneous v. deliberative; Perugini 2005), and to be influenced by different causal factors (Gawronski and Strack 2004; Rydell, McConnell, Mackie, and Strain 2006). In the domain of religiosity, researchers have investigated how believers behave differently when asked directly about their propositional beliefs compared to when their beliefs are assessed via tasks that require more intuitive responses (Barrett and Keil 1996; Barrett 1998, 1999). In spite of significant unresolved disagreements about the precise nature of implicit cognitions, it is clear that our traditional self-report measures only provide a partial picture of what is going on in people's minds.

What Are Indirect and Implicit Measures?

We have already implied that the terms "indirect" and "implicit" are not synonymous; before we proceed to consider various examples of such measures, it is important to clarify the distinction. Indirectness refers to a property of the measurement procedure: a measure is indirect insofar as it does not rely on participants' self-assessment, but involves inference about participants' mental states and traits from some other behavior.

For example, a direct measure of racial prejudice might involve asking someone about his or her attitudes toward racial outgroups, whereas as indirect measure might involve observing the person's body language when interacting with racial outgroup members. All the measures described in this chapter are, to different extents, indirect in this way. Implicitness, on the other hand, refers to some feature of the psychological process that produces the responses in any given measurement procedure, whether indirect or otherwise. More specifically, a measure may be said to be implicit insofar as responses to it are made unconsciously, unintentionally, uncontrollably, and/or resource efficiently (Gawronski and De Houwer 2014; Moors, Spruyt, and De Houwer 2010).

Consciousness is notoriously difficult to define, but it will suffice for our purposes to say that a response is conscious to the extent that it involves the respondents' awareness and, concomitantly, unconscious to the extent that the respondent is unaware of some element of his or her response. For example, measures that involve *subliminal priming*— presenting stimuli in such a way that the participant does not realize that he or she has been exposed to them—can be said to be unconscious; however, participants' responses to the task (e.g., keypresses) might still be conscious. Intentionality and controllability are related properties that apply to the relationship between goals and behavioral outcomes. To intend an outcome is to have a goal to produce that outcome; to control an outcome is to successfully produce the intended outcome. It is important to note that unintended outcomes are not just random *accidents*: for an unintended outcome to be psychologically meaningful, it has to be related to a person in some stable way. An accidental fall on a wet floor does not tell us as much about a person as repeated unintentional falls under different circumstances. Similarly, a one-off slip of the tongue does not count as a measure of implicit attitudes unless it is replicable in some reliable way. Finally, resource efficiency pertains to the extent to which a response requires attention: thus, measures that are immune to distraction are implicit in this sense.

These four properties—dubbed the Four Horsemen of *automaticity* by social psychologists (Bargh 1994)—are related but distinct; furthermore, psychological measures, including indirect measures, vary to the extent that they involve each of them. In other words, not all indirect measures are implicit measures, and not all implicit measures are

implicit in the same way. Both of these principles apply to the methods we describe in the remainder of this chapter, and we will revisit them when we discuss best practices in the use of indirect and implicit measures. In the next two sections, we will introduce a series of indirect and implicit measurement paradigms that have appeared in the research literature before concluding with some guidelines on how to adapt these paradigms for use in social scientific research.

"Low-Tech" Options

In their attempt to measure religiosity indirectly (and/or implicitly), some psychologists have adopted and adapted techniques that allow them to infer participants' psychological traits and states through measuring and interpreting participants' responses to certain (and sometimes seemingly unrelated) cues. We call these "low-tech" measures because their administration does not require any advanced technology like computers or special equipment. Two good examples of these are partially structured measures and assimilation bias measures.

Partially Structured Measures

Traditional projective measures, such as the Rorschach test (Rorschach 1927) and the Thematic Apperception Test (Murray 1943) require participants to respond to images—ambiguous visual patterns in the former case, and drawings of events in the latter—and testers to interpret participants' responses to infer personality characteristics and other psychological traits and states. In this particular form, results are heavily reliant on the tester's subjective (and potentially idiosyncratic) interpretations of participants' responses, and often reveal more about the tester than about the participant. Recently, however, Vargas, von Hippel, and Petty (2004) attempted to revive and reform partially structured measures, designing a narrative-based (as opposed to image-based) version to assess participants' personal religiosity. Consider the following vignette:

> Mary didn't go to church once the whole time she was in college but she claimed that she was still a very religious person. She said that she prayed

occasionally and that she believed in Christian ideals. Sometimes she watched religious programs on TV like the 700 Club or the Billy Graham Crusade. (Vargas et al. 2004:197)

In Vargas et al.'s (2004, Study 4) study, participants were presented with twenty such vignettes describing a character's religiosity, and were asked in each case, "How religious was the behavior [Mary] performed?" and "How religious do you think [Mary] is in general?" Participants' judgments about Mary's religiosity depend in part on their *own* religiosity: specifically, participants who are very religious will perceive Mary to be much less religious than those who are themselves relatively nonreligious. On the basis of this idea, Vargas et al. (2004) asked participants to respond to both questions on an eleven-point scale anchored at "not at all religious" and "extremely religious." Given that religious participants are expected to rate Mary as relatively *non*religious, their scores were reversed-coded so that higher scores implied that participants were very religious: the forty ratings were averaged together to form a single religiosity score. As predicted, participants' scores on this measure predicted their self-reported religious behavior; they were also correlated with participants' self-reported religious attitudes. Crucially, they predicted unique variance in religious behavior beyond what was accounted for by the self-reported attitude measure. Thus, this partially structured measure has incremental validity; it does not merely serve as an alternative to its self-report counterpart, but can also supplement it.

Partially structured measures have also been used to assess particular theological beliefs. Barrett and Keil (1996) designed a series of vignettes to assess the extent to which participants held *anthropomorphic* representations of God. They presented eight vignettes, such as the following:

A boy was swimming alone in a swift and rocky river. The boy got his left leg caught between two large, gray rocks and couldn't get out. Branches of trees kept bumping into him as they hurried past. He thought he was going to drown and so he began to struggle and pray. Though God was answering another prayer in another part of the world when the boy started praying, before long God responded by pushing one of the rocks so the boy could get his leg out. The boy struggled to the river bank and fell over exhausted. (Barrett and Keil 1996:224)

These vignettes were designed to be ambiguous. In the case above, the vignette—or so Barrett and Keil (1996) argue—neither states nor necessarily implies that God is at one particular place at any given time, or that God moves at any point, or that God answered the two prayers sequentially. The measure of theological anthropomorphism relies on this ambiguity. Unlike in Vargas et al.'s (2004) measure, Barrett and Keil's (1996) participants were not asked to make any particular judgments, but instead to recall information from these vignettes. Participants' recollections were then coded for the extent to which they featured anthropomorphic interpretations of the vignettes. To the extent that participants (mis)remembered that God moves from one place to another or answers prayers sequentially, Barrett and Keil (1996; Barrett 1998) inferred that they displayed anthropomorphic beliefs about God (e.g., being limited by space and time). Indeed, their study showed that religious believers held these "theologically incorrect" beliefs even while denying them in a direct self-reported measure. Admittedly, however, Barrett and Keil's (1996) partially structured measure has not been assessed for its psychometric properties; their studies demonstrated a phenomenon that could be exploited to construct an indirect measure, but we do not yet have a fully developed measurement tool here.

In summary, partially structured measures seem like promising tools, adaptable for measuring different aspects of religiosity: Vargas et al. (2004) assessed individuals' levels of religious commitment, and Barrett and Keil (1996) assessed individuals' anthropomorphic representations of God. However—as with other indirect and implicit measures—they remain controversial as it is still unclear whether participants' responses reliably indicate underlying attitudes, or reflect some extraneous artifact of the task or the context of measurement. It is also unclear whether these measures count as *implicit* measures, as defined earlier. Vargas et al. (2004) and Barrett (1998; Barrett and Keil 1996) do claim that these measures assess implicit attitudes, but they do not clearly fulfill the criteria for automaticity (Bargh 1994; Gawronski and De Houwer 2014). More psychometric research needs to be done for these understudied and underutilized measures.

Assimilation Bias Measures

The *assimilation bias*, sometimes known as the *confirmation bias*, is the tendency of people to search for and evaluate evidence in ways that maintain their current attitudes. Social psychologists have exploited this bias to infer individuals' attitudes in an indirect way. Saucier and Miller (2003), for example, designed their Racial Argument Scale (RAS) in this way. The RAS consists of a series of sixteen short paragraphs, some of which presented arguments with pro-Black conclusions and others of which presented arguments with anti-Black conclusions. The participants were asked to evaluate each paragraph by rating how well each argument supported its conclusion on a five-point scale ranging from "not at all" to "very much." Participants were not asked to rate how much they agreed with the conclusion, only the extent to which the conclusion followed from the premises in the argument. Saucier and Miller (2003) found that test-retest reliability for the RAS was high ($r = .81$); RAS scores were moderately correlated with various explicit measure of racial prejudice ($r = .42$ to $.57$) and uncorrelated with social desirability measures. Furthermore, RAS scores predicted respondents' willingness to provide their contact details to a Black student organization for the purposes of completing a phone survey, and also their willingness to provide negative feedback about Black authors' work beyond the contribution of explicit measures. As with the partially structured measures in the previous section, this demonstrates incremental validity.

There is, as yet, no religiosity measure that exploits the assimilation bias in this way. The closest approximation is Norenzayan and Hansen's (2006, Experiment 2) measures of people's beliefs in supernatural agency. They presented participants with (what was ostensibly) a *New York Times* article describing an experiment on the efficacy of prayer, which purportedly showed that women who were prayed for were far more likely to successfully conceive a child than those who were not prayed for. Participants were then asked to rate their agreement with a series of statements. Some of these statements were directly about participants' supernatural beliefs (e.g., "God/a higher power exists"), but others were about the article itself (e.g., "The study was scientifically rigorous"). Norenzayan and Hansen (2006) did not analyze responses to the latter statements, but Saucier and Miller's (2003) work strongly

suggests that such items might serve as good indirect measures of supernatural belief (or, more specifically in this case, belief in the causal efficacy of intercessory prayer).

Response Latency–Based Measures

In contrast to the "low-tech" measures described above, by far the most commonly used indirect measures are those that use computers or other equipment to record participants' response latencies or reaction times in some task and, from these latencies, infer participants' attitudes and other psychological states and traits. These response latency–based measures are also generally considered to be *implicit*, in the sense that they involve automaticity to varying degrees. For the remainder of this chapter, we will consider three classes of measures that rely on participants' response latencies. The basis of each of these is the *simple choice reaction time* task, which requires participants to respond to stimuli by categorizing them into one of two options as quickly as they can. An example of this is the *lexical decision task*, which requires participants to categorize strings of letters as either words or nonwords: "key" is a word, whereas its anagram "yek" is not. Another common example involves categorizing words as either positive or negative: "delightful" is positive, whereas "awful" is negative. The stimuli in these tasks are not limited to words, though they are typically presented visually: thus, images of particular objects or types of objects may also feature. In each case, psychologists are interested in how *quickly* (and accurately) individuals perform these tasks: the speed (and accuracy) of making these classifications is taken as an indication of how information is cognitively stored or processed. The simple choice reaction time task forms the basis for all the other implicit measures of religiosity that we will consider below.

Property Verification Measures

A *property verification task* measures the links or associations, in participants' minds, between some stimuli and some property of the target objects to which the stimuli refer. If, for example, the researcher is interested in participants' positive or negative evaluations of Christianity, she might choose target stimuli like "church," "priest," and "cross," and

property words like "good" and "pleasant" on one hand and "bad" and "unpleasant" on the other. The way in which participants categorize the targets—as positive or negative—is likely to closely reflect their *explicit* attitudes, but the *speed* at which they perform the task may tell us more: speedier responses indicate stronger attitudes. This task therefore potentially includes both an explicit and an implicit measure at the same time. For any given task of this kind, researchers have to determine what the targets are as well as what the properties are: this depends on the psychological construct of interest.

Let us turn to actual examples of property verification tasks applied to religiosity. Cohen, Shariff, and Hill (2008) presented participants with a series of nouns, which participants simply had to categorize as quickly as possible as either "real" or "imaginary." Some of these words referred to things that were uncontroversially real (e.g., car, water); others referred to things that were uncontroversially imaginary (e.g., Darth Vader, Superman); others still referred to objects of religious faith (e.g., God, Devil); and yet others referred to objects of secular faith (e.g., black hole, Socrates). There were eight stimuli for each category, and each stimulus was presented twice: stimuli were presented one at a time in random order, with a five-hundred-millisecond gap between each trial. Clearly, the category of interest was one referring to the objects of religious faith: furthermore, given the properties—real and imaginary— Cohen et al. (2008) were not evaluating how positively or negatively people felt about these objects, but the extent to which people believed in them. Cohen et al. (2008) found that individuals' classification of religious items as real or imaginary was highly reliable ($\alpha = .95$), and response latencies for religious items showed moderately good test-retest reliability (after seven days), even controlling for overall response speeds ($\beta = .24$). Furthermore, only response latencies for religious items were correlated with self-reported religiosity: indeed, there was a *curvilinear* relationship between them. That is, very religious people categorized objects like "God" and "Devil" as real faster than did more nominally religious people; similarly, very nonreligious people categorize those objects as imaginary more quickly than did more nominally nonreligious people.

More recently, Jong, Halberstadt, and Bluemke (2012, Study 3) ran a modified and simplified version of this measure, and found evidence for

Table 3.1. Stimuli for Jong et al. (2012) simple choice reaction time measure

Category	Stimuli
Supernatural	angel, demon, devil, god, heaven, hell, miracle, prophet, soul, spirit
Real	eagle, helicopter, otter, Puerto Rico, turtle
Imaginary	batmobile, fairy, genie, mermaid, Narnia

discriminant validity relative to a self-report measure of religious belief, while studying the relationship between religiosity and death anxiety. In their version of the task, they had twenty stimuli: ten referred to objects of religious belief, five referred to patently real objects, and five referred to patently imaginary objects (see table 3.1). Each stimulus was presented thrice, and the order of presentation was randomized for each participant. Prior to each trial, a fixation cross was presented for 750 milliseconds; the fixation cross reappeared immediately after participants made their response about whether each object was real or imaginary. Jong et al. (2012) found that thinking about death reduced nonreligious participants' scores on a self-report measure of religious belief, but *increased* their scores on this implicit measure: the experimental manipulation had diverging effects on explicit and implicit religious belief, thus showing the added value of assessing both types of belief.

Property verification tasks have also been used to measure other aspects of religiosity besides belief in the supernatural. Gibson (2006), for example, has constructed a yes/no task in which participants are asked to determine whether or not certain adjectives describe God; Sharp, Rentfrow, and Gibson (2015) then adapted this task to compare Christians' representations of the three "persons" of the Trinity. Other researchers have also extended this work to examine individuals' positive or negative evaluations of God (Yarborough 2009; Zahl 2013).

There is a sense in which property verification tasks are not indirect measures at all, given that they *directly* ask participants to make judgments about religious concepts. However, they are indirect in the sense that strength of religious belief is inferred not from participants' categorization responses, but from participants' response latencies on the critical target stimuli. Similarly, these tasks do not unambiguously fulfill the four criteria of implicitness or automaticity. The categorization task itself is performed intentionally, consciously, controllably, and

not particularly resource efficiently, though Järnefeldt, Canfield, and Kelemen (2015) provide some evidence that speeded responses are less than perfectly amenable to conscious control. Using a yes/no task, they asked professed atheists whether a series of natural phenomena were "purposefully made by some being" or not. Approximately half the participants were told to answer "as quickly as possible" (with a maximum time of 865 milliseconds per trial), while the other half were given as much time as they needed. They found that those in the speeded condition categorized more phenomena as having been created than those in the unspeeded condition. Furthermore, while responses to property verification tasks are not necessarily automatic in principle, participants are usually unaware that their response latencies are being measured, and are thus also unlikely to attempt to intentionally control their speed.

Sequential Priming Measures

A sequential priming task essentially involves a simple choice reaction time task, with additional stimuli that participants are meant to ignore ("primes"); the primes may be presented either above the threshold of conscious awareness (supraliminally) or below (subliminally). Furthermore, the primes may be presented prior to or simultaneously with the target stimuli, which participants have to categorize. As a measure of attitudes, this task exploits the *congruency effect*: participants tend to respond more slowly to targets when they are paired with *incongruent* primes, and more quickly when paired with *congruent* primes. Broadly speaking, pairs of stimuli can be *semantically* or *evaluatively* congruent or incongruent, though there is some overlap between the two.

The classic example of semantic priming is that, in a lexical decision task, the stimulus "doctor" is more easily and quickly recognized as a word when it is paired with "nurse" than when it is paired with "bread." Conversely, "butter" is more easily and quickly recognized as a word when it is paired with "bread" than when it is compared with "nurse" (Meyer and Schvaneveldt 1971). From this, we infer that the cognitive associations between the concepts *doctor* and *nurse* are stronger or closer to each other than they are to *bread* and *butter*, which are conversely more strongly associated to each other than they are to the previous pair. An evaluative priming task (sometimes known as an affective priming

task) is an extension of the semantic priming phenomenon, but relies on a valence congruency effect (Fazio 2001; Fazio, Sanbonmatsu, Powell, and Kardes 1986), and concomitantly requires participants to categorize target stimuli as either positive or negative. Thus, in their seminal studies, Fazio et al. (1986) showed that positive primes facilitate the categorization of positive stimuli, whereas negative primes facilitate the categorization of negative stimuli. Sequential priming tasks—both semantic and evaluative—have long been used as measures of attitudes; there is therefore considerable evidence of their reliability and validity. For example, Cameron, Brown-Iannuzzi, and Payne (2012) ran a meta-analysis of 167 studies involving sequential priming measures and found that they were both correlated with explicit attitude measures ($r = .2$), as well as with behavioral measures ($r = .28$); implicit measure scores continued to predict behavior even when controlling for explicit measure scores. Evaluative priming measures thus show incremental validity.

There are, as yet, no validated implicit measures of religiosity based on the evaluative priming task. However, Wenger (2004) has designed a task that is potentially useful as such a measure. In this task, participants were primed with the words "Christian" (religious prime), "student" (neutral prime), and "housetop" (nonhuman neutral prime); they then categorized a series of phrases as describing behaviors that are either "possible" or "impossible" to perform. The stimuli consisted of sixteen phrases, four describing religious behaviors (e.g., worship God), four, academic behaviors (e.g., take tests), and eight, nonsensical nonactions (e.g., open sand). Wenger (2004) theorized that the extent to which the "Christian" prime facilitated responses to the religious actions indicated close correspondence between religious ideas and religious practice; indeed, he found that performance on this task predicted *intrinsic* religiosity but not *extrinsic* religiosity (see Hill and Pargament, this volume). Wenger (2004) interpreted this task as "indicating a kind of internalization of beliefs for intrinsically oriented individuals" (p. 12), but with different stimuli, the evaluative priming task may be adapted to measure other religiosity constructs. With measures of religious belief like Cohen et al.'s (2008) and Jong et al.'s (2012) property verification tasks, for example, researchers may use the word "god" (and other words related to religion) as the prime, and positive (e.g., genuine, actual) and negative (e.g., bogus, fictional) existential concepts as targets, to be categorized

either as positive/negative existential concepts (i.e., real/imaginary) or as words/nonwords in a lexical decision task.

So far, we have been describing congruence—whether semantic or evaluative—as if it were an objective property of the pairs of stimuli, but whether two concepts are cognitively associated is bound to vary somewhat from person to person, and even from situation to situation. For some stimuli, there is so much stable intersubjective agreement that we can assume, for example, that "bread" and "butter" go together, and that words like "delightful" are positive and words like "awful" are negative. But when we are using a sequential priming task as a measure of attitudes, we are trying to *assess* rather than assume the extent to which stimulus pairs are congruent or incongruent for any given individual in any given situation. And given that personal religiosity is a highly subjective topic, the degree of variation between people in their individual perception of the degree of congruence between religious primes and specific targets will result in meaningful differences in performance on these tasks.

Implicit Association Test and Its Variants

The Implicit Association Test (IAT; Greenwald, McGhee, and Schwartz 1998) and its descendants are among the most commonly used implicit measures of psychological constructs, including prejudice (Rudman, Greenwald, Mellott, and Schwartz 1999), self-esteem (Greenwald and Farnham 2000), anxiety (Egloff and Schmukle 2002), aggression (Lemmer, Gollwitzer, and Banse 2015), consumer attitudes (Maison, Greenwald, and Bruin 2004), and attitudes toward alcohol (Ostafin and Palfai 2006). As the name suggests, IATs measure the strength of cognitive *associations*; in particular, they compare two target-attribute pairs on the basis of relative response times.

IATs consist of two critical categorization tasks, implemented across five phases. Phases 1 and 2 are both practice phases, in which participants learn to respond to the *target* and *attribute* stimuli as quickly and accurately as possible by pressing two keys. For example, participants might first be asked to categorize target images of Pepsi-Cola and Coca-Cola by pressing the "Z" and "M" keys of a keyboard, respectively. Then, in Phase 2, they would be asked to categorize positive (e.g., "delightful") and negative (e.g., "awful") attribute words using those same keys. Phase 3

is the first critical categorization task, in which both kinds of stimuli—targets and attributes—are presented, randomly interspersed, and participants have to categorize each of them accordingly ("Z" for Pepsi OR positive; "M" for Coke OR negative). Given the valence congruence effect described earlier, we would expect people who prefer Pepsi-Cola over Coca-Cola to find this task easier (and therefore to respond faster) than those who prefer Coca-Cola over Pepsi-Cola. Phase 4 is another practice round like Phase 1, but now the key presses are reversed ("Z" for Coke, "M" for Pepsi). Finally, Phase 5 is identical to Phase 3, except that the target responses correspond to Phase 4 ("Z" for Coke OR positive, "M" for Pepsi OR negative). For this phase, we might expect Coke fans to outperform Pepsi fans. The critical phases are 3 and 5, when participants have to categorize both targets and attributes. The response latencies observed in Phases 3 and 5 allow us to compute the *relative* strength of one pair of associations ("Coke good, Pepsi bad") with another pair of associations ("Pepsi good, Coke bad"). There are multiple ways to compute an implicit association score from raw response latencies, but the basic idea is to compare between Phases 3 and 5 (see Greenwald, Nosek, and Banaji 2003 for detailed algorithm), thereby producing a score that assesses the implicit attitude towards one object in relation to another object. Greenwald, Poehlman, Ullmann, and Banjo's (2009) meta-analysis of 184 independent samples found that IAT scores were correlated with explicit measures ($r = .36$) and behavioral measures ($r = .27$), and also provided evidence for incremental validity as we have seen with other implicit measures.

An example of an IAT in the domain of religiosity is the Religiousness-Spirituality IAT (RS-IAT; Labouff, Rowatt, Johnson, Thedford, and Tsang 2010). This IAT measures the extent to which people consider themselves religious relative to others. Participants were asked to categorize words as belonging to "self" ("I," "me," "my," "mine," "self") or "other" ("they," "them," "their," "it," "other"), and to "religious-spiritual" ("religious," "spiritual," "faithful," "theistic," "believer") or "not religious–not spiritual" ("nonreligious," "nonspiritual," "faithless," "atheistic," "agnostic"). Higher scores on the RS-IAT indicate that people are faster at categorizing self/religious-spiritual and other/not religious–not spiritual than at categorizing self/not religious–not spiritual and other/religious-spiritual. Scores on the RS-IAT were found to be positively correlated

with self-report measures of religiosity, and to account for unique variability in predicting people's attitudes towards gay men and lesbian women.

Despite being a tried-and-true measure of implicit attitudes, one common criticism of the classic IAT is that it requires two contrasting targets (e.g., Pepsi v. Coke), and therefore that it is a measure of *relative* attitudes: the classic IAT does not tell us how someone feels about Pepsi, so much as about how he or she feels about Pepsi compared to Coke. This limitation is particularly problematic for assessing some religiosity variables, such as religious beliefs, because it is not clear what the contrasting target would be, given that researchers are not necessarily interested in individuals' religious beliefs relative to their secular beliefs. Fortunately, there is a variant of the IAT, the Single Target IAT (ST-IAT, Wigboldus, Holland, and van Knippenberg 2006; also Karpinski and Steinman 2006; Penke, Eichstaedt, and Asendorpf 2006; and see Nosek and Banaji 2001 on the Go/No Go Task). It is largely similar to the classic IAT, but assesses only one target. For example, if we were to assess implicit attitudes towards Coca-Cola, the first critical categorization task might involve classifying images of Coca-Cola OR negative words with the left key, and positive words with the right key. The second critical categorization task would involve classifying negative words with the left key, and positive words OR images of Coca-Cola with the right key. Just as with the IAT, the implicit association score is the difference between the two critical test phases (see Bluemke and Friese 2008 for more information on computing ST-IAT scores, and for a psychometric evaluation of the ST-IAT).

Together with the property verification task described earlier, Jong et al. (2012, Study 2) also designed a Religious Belief ST-IAT to assess strength of implicit religious belief. In Phase 1, participants categorized attribute words connoting "real" and "imaginary" by pressing the "Z" and "/" keys on their keyboard, respectively (see table 3.2 for stimuli). Each stimulus was presented thrice, in random order, preceded by a five-hundred-millisecond fixation cross. In Phase 2, words referring to supernatural entities were also presented, interspersed with the "real" and "imaginary" words; participants responded to "real" OR "supernatural" words (presented thrice each) by pressing "Z," and to "imaginary" words (presented six time each) by pressing the "/" key. In Phase 3,

Table 3.2. Stimuli for Jong et al. (2012) ST-IAT Measure

Category	Stimuli
Supernatural	god, demon, devil, angel, heaven, hell, soul
Real	real, genuine, existent, actual, true, valid, factual
Imaginary	imaginary, fake, false, fictional, bogus, untrue, illusory

the key for "supernatural" words was switched such that participants responded to "real" words by pressing "Z" and "imaginary" OR "supernatural" words by pressing "/".

As mentioned earlier in our discussion of the property verification task, Jong et al. (2012) found that thinking about death decreased non-religious participants' self-reported religious beliefs, but *increased* their implicit religious beliefs. This finding was replicated using the Religious Belief ST-IAT: thinking about death increased both religious and non-religious participants' ST-IAT scores. This is not to say that explicit and implicit beliefs *always* diverge. For example, Shariff, Cohen, and Norenzayan (2008) also employed an ST-IAT, and found that reading essays that criticized religion reduced both explicit and implicit religious belief. This shows that the relationship between explicit and implicit beliefs depends at least in part on the experimental manipulation. This, along with the fact that explicit and implicit measures of religious belief are positively correlated ($r = .27$, Jong et al. 2012; .31, Shariff et al. 2008), is also consistent with the view that explicit and implicit cognitions are not totally dissociated from one another, but are related albeit nonidentical constructs.

Constructing and Using Indirect and Implicit Measures

There is one obvious difference between self-report measures of religiosity and indirect or implicit measures of religiosity: whereas there are dozens upon dozens of more or less *fixed* versions of the former that have been psychometrically evaluated, this is not the case for the latter. Rather, we have various *paradigms*—partially structured tasks, tasks involving the assimilation bias, simple choice reaction time tasks, sequential priming tasks, Implicit Association Tests, and so forth—that form the basis for constructing specific implicit measures. Indeed,

researchers are much more likely to construct their own implicit measures of religiosity than to reuse one that has been used before; this is in stark contrast to the general tendency to prefer established self-report scales over ad hoc ones. Accordingly, in this final section, we provide some basic guidelines for how to construct, use, and evaluate implicit measures of religiosity from the sort of paradigms described above.

Why (and Why Not) Indirect and Implicit Measures?

Before considering the question of which indirect/implicit paradigm should be used, researchers should first ask a more basic question: why use indirect/implicit measures at all? As we have discussed earlier, indirect and implicit measures were developed to overcome self-report biases and the limits of human introspection. We therefore recommend using implicit measures under the following conditions:

1. there is good reason to believe that self-report measures cannot yield any meaningful data (e.g., because participants have limited introspective access on the trait in question); and/or
2. the researcher is interested in the effects of religiosity on some outcome variable, but does not want participants to know that their religiosity is being assessed (e.g., because awareness is likely to affect their behavior in undesirable ways); and/or
3. there is good reason to expect a divergence between explicit and implicit religiosity.

Insofar as there is good theoretical reason to assume that an implicit measure can illuminate some aspect of religiosity that would meaningfully contribute to the research question, then the paradigms we have presented can be valuable: they should not supplant but support (and perhaps even be meaningfully contrasted with) self-report measures or observational studies of religiosity. There is no question that the richness of religious experience and the depth of conviction that people hold about their religious beliefs could never be adequately reflected in response latencies or performance on a comprehension task with fictitious narratives. Nevertheless, the measures we have introduced are based on empirically grounded theories about human cognition, and

we have every reason to think that these theories and paradigms will enhance our ability to study religiosity scientifically.

Our enthusiasm for indirect and implicit measures should not be confused with the naïve belief that they are panaceas that swiftly and straightforwardly cure us of our methodological weaknesses. First, it would be a gross oversimplification to say that indirct and implicit measures are completely immune to those self-presentational effects that they seek to avoid. It is not impossible to "fake" responses to implicit measures (e.g., Fiedler and Bluemke 2005; Röhner, Schröder-Abé, and Schütz 2013; Steffens 2004), though they are much harder to fake than traditional self-report measures. It is similarly unjustified to assert that implicit measures provide privileged and unequivocal access to "the unconscious," as if this were a mysterious mental realm far beyond the reach of our introspective powers, nor is it correct to value implicit attitudes as somehow *more* true or more genuine than explicit attitudes. Against the view of radical independence is the fact that the correlations between explicit and implicit measures tend to be nonzero and positive, albeit only moderately so (e.g., .24 in a meta-analysis by Hofmann, Gawronski, Gschwendner, Le, and Schmitt 2005 on the Implicit Association Test). Furthermore, the fact that different implicit measures can and sometimes do diverge from one another indicates not only that they vary in quality but also potentially that different implicit measures are assessing different psychological constructs, or different aspects thereof (e.g., Bosson, Swann, and Pennebaker 2000; Fazio and Olson 2003). A more reasonable attitude toward implicit measures is perhaps to say that, under certain circumstances, they help us to access and assess aspects of human cognition that are related to but nevertheless distinct from those aspects that are more easily measured via self-report measures.

Having decided to use indirect or implicit measures, the researchers must choose which paradigm to use. This choice will largely depend on both theoretical and practical considerations.

Which Paradigm? Theoretical Considerations

The first theoretical decision to be made concerns the aspect of religiosity that is of theoretical relevance or interest. As the theoretical and methodological research on the multidimensionality of religiosity has

amply shown, it is simply too vague to say that we are interested in measuring "religiosity." Not only are there important differences between religious *identity* and *belief* and *behavior*, but there are different kinds and levels of each. "Religious beliefs" may include attitudes toward God, which may be entirely independent from one's attitude towards organized religions, beliefs about angels and demons, and beliefs about the afterlife. It is therefore important for the researcher to clearly identify a conceptual framework for what precisely he or she wishes to study (including how implicit cognition comes into play in the particular framework), before choosing the most appropriate methodology. Fortunately, the paradigms we have introduced in this chapter are flexible with respect to content, and can therefore be adapted to measure a wide variety of religious cognitions.

The second theoretical decision pertains to the measurement problem at hand. The tasks we have described vary on whether they were designed to solve the strategic-responding problem, the limitations-of-introspection problem, or both. In other words, they vary on the extent, as well as the sense, in which they involve automaticity. Recall that a measure is automatic insofar as responses do not require intention, can occur without conscious awareness, are difficult or impossible to control, and do not require much attentional resource. These four criteria of automaticity are neither binary properties, nor do they always co-occur. The fact that the criteria are dissociable means that different measurement techniques might capture psychological constructs that are automatic in different ways. It also means that a single task can be an implicit measure in one sense, but not in another. Thus, a valid interpretation of what any particular implicit measure tells us requires additional information about the precise sense in and extent to which a measure is implicit. For example, as we have already discussed, property verification tasks are implicit to the extent that they manage to bypass participants' conscious, intentional control; however, given that participants may, with some ease, respond in such a way as to "fake" their responses, it is fairer to conclude that property verification tasks are only minimally implicit on these aspects of automaticity. In contrast, studies have shown that while IATs are not immune to faking, they are much more difficult to fake than traditional questionnaire measures, particularly for novice participants. Furthermore, participants completing IATs

are less consciously aware of what is being measured than participants completing property verification tasks: the latter might not realize that their response latencies are being used as measures, but the former might not even know that the task measures attitudes at all. That said, the categorization tasks involved in IATs might arouse participants' suspicions about the true purpose of the IAT; there is indeed some evidence to that effect (De Houwer and Moors 2012), showing also that IATs for some attitudes (e.g., racial prejudice) arouse more suspicion than IATs for others (e.g., political attitudes). Sequential priming techniques seem even less amenable to faking, though it is again not impossible (Degner 2009; Klauer and Teige-Mocigemba 2007; Teige-Mocigemba and Klauer 2008). To reduce fakeability—that is, intentionality and controllability—researchers can reduce the time between the presentation of the prime and the presentation of the stimulus (even to the point of subliminality; e.g., Wittenbrink, Judd, and Park 1997), and also impose a deadline by which a response must be made in order to be counted as valid (e.g., six hundred milliseconds; Degner 2009). Furthermore, while there is no direct evidence that sequential priming techniques can access cognitions about which participants are themselves unaware, the fact that subliminal priming is effective in this context means that sequential priming can access cognitions without participants' awareness that this is being done at all.

There is still insufficient evidence regarding indirect measures in general, let alone measures of religiosity more specifically, for us to make firm conclusions, but sequential priming techniques seem to most completely fulfill the automaticity criteria, relative to other paradigms (cf. Moors, Spruyt, and De Houwer 2010; for a more detailed comparison, see De Houwer, Teige-Mocigemba, Spruyt, and Moors 2009). Thus, if *implicitness* is crucial, all else being equal, researchers should use sequential priming tasks (see Wentura and Degner 2010 for more practical recommendations on this task). However, there may be other reasons for preferring other techniques, such as the IAT. For example, given that sequential priming tasks do not require participants to attend to the category of the stimuli—this is a feature that increases its implicitness—it is, under some circumstances, more likely to be affected by idiosyncratic features of individual stimuli: that is, sequential priming measures might tell us more about participants' attitudes toward the particular *stimuli* in the task than about their attitudes toward the *categories* they

are intended to represent (Livingstone and Brewer 2002; Olson and Fazio 2003). As with the choice between paper- and computer-based tasks, practical considerations like the feasibility of subliminal priming or the familiarity of potential participants with any given technique may also influence a researcher's decision about which paradigm to employ.

Which Paradigm? Practical Considerations

If the researcher determines that indirect or implicit measures would indeed produce meaningful and valuable data to answer specific research questions, then the second set of considerations concerns the practical administration of indirect/implicit measures. In this chapter, we have introduced five measurement paradigms, two of which are "low-tech" options that can be done on paper, and three of which are based on partic-ipants' response latencies. The former category can be easily administered in a range of contexts, though some measures may need to be adminis-tered in a quiet space that is free of distraction. For example, a task that measures a participant's memory of certain types of stimuli as an indirect measure of religious beliefs may need to be administered in a quiet room that is free of distraction in order ensure that the data are valid. Response latency–based tasks can, in principle, be run on paper, but are more com-monly administered via computers (see Bassett et al. 2005 for more on paper-and-pencil religiosity IATs). One factor that might determine which paradigm is most appropriate is therefore the extent to which it is possible or desirable to run a paper-and-pencil measure rather than a computer-based measure. It may not, for example, be possible to run computer-based measures in certain situations outside the research laboratory (e.g., dur-ing fieldwork): if so, a researcher might have to use a "low-tech" indirect measure or a paper-and-pencil IAT instead. Practical considerations aside, we recommend that the choice of measurement technique be primarily dictated by one's theoretical interests, as discussed above.

Indirect Measures and the Contemporary Social Scientific Study of Religion

The best scientific tools help us not only to answer our current ques-tions but also to ask those questions in better ways and even to provoke

us to ask *new* questions. At the beginning of this chapter, we introduced indirect and implicit measures as means by which to solve longstanding measurement problems. Very quickly, however, we saw how methodological advances in implicit *measures* opened up new avenues of research into implicit—that is, automatic—*cognitions* and *processes*. We have already learned the lesson of the multidimensionality of religiosity from methodological advances in self-report measures, and the new tools covered in this chapter extend this multidimensionality in a different direction, toward exploring aspects of religiosity that may be *unconscious*, for example, and even seemingly at odds with explicitly held attitudes. Indeed, these new methods have already begun to bear theoretical and empirical fruit: as it has done in other social scientific domains, the notion of implicit cognitions now features prominently in developmental and evolutionary theories of religion (e.g., Atran 2002; Boyer 2001; Kelemen 2004; Tremlin 2006; Uhlmann, Poehlman, and Bargh 2008). Therefore, insofar as social scientists are interested in the correlates, causes, and consequences of religiosity, we should welcome indirect and implicit measures into our arsenal of measurement techniques, and celebrate the new opportunities they afford us to explore this most intriguing of human phenomena.

Glossary

AUTOMATICITY: In cognition, this refers to the influence of external stimuli on a person's psychological processes without his or her knowledge or awareness of such influence. Automaticity can refer to any one or more of the following features: conscious awareness, intentionality, controllability, and efficiency.

CONTROLLABLE/UNCONTROLLABLE: Refers to the extent to which an individual is motivated and able to counteract a particular cognitive process.

FIXATION CROSS: A symbol (typically a cross) presented on the screen, usually in the center, where the stimulus is expected to be presented.

IMPLICIT MEASURES: Measures to which responses are unconscious, unintentional, uncontrollable, and/or resource efficient.

INDIRECT MEASURES: Measures that do not rely on participants' self-assessment but involve inferring participants' mental states and traits from some other behavior or response.

IMPLICIT ASSOCIATION TASK: This task evaluates the relative strength of two related pairs of associations (e.g., "Coke is good, Pepsi is bad" and "Coke is bad, Pepsi is good") by comparing the response latencies of each pair of associations.

INTENTIONAL/UNINTENTIONAL: Intentionality refers to the extent to which the individual has control over the instigation of cognitive processes.

LEXICAL DECISION TASK: A task in which participants are asked to judge whether a string of letters is a word or a nonword. Both speed and accuracy of judgment are considered.

MASK (PRIMING): A symbol (typically ######) presented before and/or after a word-based prime that diminishes the visibility of primes.

PRIME: *Verb:* temporary and unobtrusive activation of one or more mentally represented concepts. *Noun:* the stimulus that is expected to activate a mentally represented concept. For example, words like "church" and "Jesus" might be primes for the concept of "Christianity."

PROPERTY VERIFICATION TASK: A task that requires participants to categorize stimuli on the basis of some property or attribute of the target objects to which the stimuli refer.

RESOURCE EFFICIENCY: Refers to the degree of mental resources that are required in the cognitive process. Automatic processing typically requires low mental resources.

RESPONSE LATENCY: A measurement of the time interval (typically in milliseconds) between presentation of some stimulus and an individual's response. The timing is typically measured using a computer. The speed is treated as an indicator of the strength of association between representations.

SEQUENTIAL PRIMING: A simple choice reaction time task that is preceded by presentation of primes. We describe two types of sequential priming tasks: (1) semantic priming, which is based on the observation that a response to a target is faster when it is preceded by a semantically related prime, and (2) evaluative priming, which is

based on the observation that a response to a target is faster when it is preceded by stimuli with congruent valence.

SIMPLE CHOICE REACTION TIME TASK: A task in which participants are asked to categorize stimuli into one of two options as quickly and accurately as they can.

SUBLIMINAL PRIMING: Presentation of priming stimuli below an individual's threshold of conscious awareness.

SUPRALIMINAL PRIMING: Presentation of priming stimuli just above an individual's threshold of conscious awareness, such that the individual is aware of the stimulus being presented without knowing what the stimulus is.

TARGET (OBJECT): The mentally represented concept that is the object of investigation.

REFERENCES

Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology*, 5, 432–43.

Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford, UK: Oxford University Press.

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. Wyer & T. Srull (eds.), *Handbook of Social Cognition* (pp. 1–40). New York: Erlbaum.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, 54, 462–79.

Barrett, J. L. (1998). Cognitive constraints on Hindu concepts of the divine. *Journal for the Scientific Study of Religion*, 37, 608–19.

Barrett, J. L. (1999). Theological correctness: Cognitive constraint and the study of religion. *Method & Theory in the Study of Religion, 11*, 325–39.

Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology*, 31, 219–47.

Bassett, R. L., Smith, A., Thrower, J., Tindall, M., Barclay, J., Tiuch, K., . . . & Monroe, J. (2005). One effort to measure implicit attitudes toward spirituality and religion. *Journal of Psychology & Christianity, 24*, 210–18.

Bluemke, M., & Friese, M. (2008). Reliability and validity of the Single-Target IAT (ST-IAT): Assessing automatic affect towards multiple attitude objects. *European Journal of Social Psychology, 38*, 977–97.

Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631–43.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

Brenner, P. S. (2011). Identity importance and the overreporting of religious service attendance: Multiple imputation of religious attendance using the American Time Use Study and the General Social Survey. *Journal for the Scientific Study of Religion, 50,* 103–15.

Cameron, C. D., Brown-Iannuzzi, J. L., & Payne, B. K. (2012). Sequential priming measures of implicit social cognition: A meta-analysis of associations with behavior and explicit attitudes. *Personality and Social Psychology Review, 16,* 330–50.

Chaiken, S., & Trope, Y., eds. (1999). *Dual-process theories in social psychology.* New York: Guilford Press.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (eds.), *Dual-process theories in social and cognitive psychology* (pp. 73–96). New York: Guilford Press.

Cohen, A. B., Shariff, A. F., & Hill, P. C. (2008). The accessibility of religious beliefs. *Journal of Research in Personality, 42,* 1408–17.

Degner, J. (2009). On the (un-)controllability of affective priming: Strategic manipulation is feasible but can possibly be prevented. *Cognition and Emotion, 23,* 327–54.

De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes? In R. Proctor & J. Capaldi (eds.), *Implicit and explicit processes in the psychology of science* (pp. 183–98). New York: Oxford University Press.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135,* 347–68.

Dijksterhuis, A., & Nordgren, L. (2006). A theory of unconscious thought. *Perspectives on Psychological Science*, 1, 95–109.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82, 62–68.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology, 33,* 510–40.

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology*, 83, 1441–55.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion*, 15, 115–41.

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and uses. *Annual Review of Psychology, 54,* 297–327.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50, 229–38.

Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the Implicit Association Tests. *Basic and Applied Social Psychology*, 27, 307–16.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132, 692–731.

Gawronski, B., & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (eds.), *Handbook of research methods in social and personality psychology*, 2nd ed. (pp. 283–310). New York: Cambridge University Press.

Gawronski, B., & Strack, F. (2004). On the propositional nature of cognitive consistency: Dissonance changes explicit, but not implicit attitudes. *Journal of Experimental Social Psychology*, *40*, 535–42.

Gibson, N. J. S. (2006). *The experimental investigation of religious cognition.* Unpublished doctoral dissertation, University of Cambridge, Cambridge, England.

Gillings, V., & Joseph, S. (1996). Religiosity and social desirability: Impression management and self-deceptive positivity. *Personality and Individual Differences, 21*, 1047–50.

Gorsuch, R. L. (1984). Measurement: The boon and bane of investigating religion. *American Psychologist, 39*, 228–36.

Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, *79*, 1022–38.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, *74*, 1464–80.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.

Greenwald, A. G., Smith, C. T., Sriram, N., Bar-Anan, Y., & Nosek, B. A. (2009). Implicit race attitudes predicted vote in the 2008 U.S. presidential election. *Analyses of Social Issues and Public Policy, 9*, 241–53.

Hadaway, C. K., & Marler, P. L. (2005). How many Americans attend worship each week? An alternative approach to measurement. *Journal for the Scientific Study of Religion, 44*, 307–22.

Hill, P. C., & Hood, R. W., Jr. (1999). *Measures of religiosity*. Birmingham, AL: Religious Education Press.

Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*, 1369–85.

Järnefelt, E., Canfield, C., & Kelemen, D. (2015). The divided mind of a disbeliever: Intuitive beliefs about nature as purposefully created among different groups of nonreligious adults. *Cognition, 140*, 72–88.

Jong, J., Halberstadt, J., & Bluemke, M. (2012). Foxhole atheism, revisited: The effects of mortality salience on explicit and implicit religious belief. *Journal of Experimental Social Psychology, 48*, 983–89.

Karpinski, A., & Steinman, R. B. (2006). The single-category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91,* 16–32.

Kelemen, D. (2004). Are children "intuitive theists"? Reasoning about purpose and design in nature. *Psychological Science, 15,* 295–301.

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of two-system theories. *Perspectives on Psychological Science, 4,* 533–50.

Klauer, K. C., & Teige-Mocigemba, S. (2007). Controllability and resource dependence in automatic evaluation. *Journal of Experimental Social Psychology, 43,* 648–55.

LaBouff, J. P., Rowatt, W. C., Johnson, M. K., Thedford, M., & Tsang, J. (2010). Development and initial validation of an implicit measure of religiousness-spirituality. *Journal for the Scientific Study of Religion, 49,* 439–55.

Leak, G. K., & Fish, S. (1989). Religious orientation, impression management, and self-deception: Toward a clarification of the link between religiosity and social desirability. *Journal for the Scientific Study of Religion, 28,* 355–59.

Lemmer, G., Gollwitzer, M., & Banse, R. (2015). On the psychometric properties of the aggressiveness-IAT for children and adolescents. *Aggressive Behavior, 41,* 84–95.

Livingstone, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82,* 5–18.

Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive validity of the Implicit Association Test in studies of brands, consumer attitudes, and behavior. *Journal of Consumer Psychology, 14,* 405–15.

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology, 90,* 227–34.

Moors, A., Spruyt A., & De Houwer, J. (2010). In search of a measure that qualifies as implicit: Recommendations based on a decompositional view of automaticity. In B. Gawronski & K. B. Payne (eds.), *Handbook of implicit social cognition: Measurement, theory, and application* (pp. 19–37). New York: Guilford Press.

Murray, H. A. (1943). *Thematic apperception test manual.* Cambridge, MA: Harvard University Press.

Norenzayan, A., & Hansen, I. G. (2006). Belief in supernatural agents in the face of death. *Personality and Social Psychology Bulletin, 32,* 174–87.

Nosek, B. A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16,* 65–69.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19,* 625–66.

Olson, M. A., & Fazio, R. H. (2003). Relations between implicit measures of prejudice: What are we measuring? *Psychological Science, 14,* 636–39.

Orne, M. T., & Whitehouse, W. G. (2000). Demand characteristics. In A. E. Kazdin (ed.), *Encyclopedia of psychology* (pp. 469–70). Oxford: Oxford University Press.

Ostafin, B. D., & Palfai, T. P. (2006). Compelled to consume: the Implicit Association Test and automatic alcohol motivation. *Psychology of Addictive Behaviors, 20*, 322–27.

Penke, L., Eichstaedt, J,. & Asendorpf, J. B. (2006). Single–attribute Implicit Association Test (SA–IAT) for the assessment of unipolar constructs. *Experimental Psychology, 53*, 283–91.

Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology, 44*, 29–45.

Röhner, J., Schröder-Abé, M., & Schütz, A. (2013). What do fakers actually do to fake the IAT? An investigation of faking strategies under different faking conditions. *Journal of Research in Personality, 47*, 330–38.

Rorschach, H. (1927). *Rorschach Test: Psychodiagnostic plates*. Cambridge, MA: Hogrefe.

Rudman, L. A., Greenwald, A. G., Mellott, D. S., & Schwartz, J. L. (1999). Measuring the automatic components of prejudice: Flexibility and generality of the Implicit Association Test. *Social Cognition, 17*, 437–65.

Rydell, R. J., McConnell, A. R., Mackie, D. M., & Strain, L. M. (2006). Of two minds: Forming and changing valence-inconsistent implicit and explicit attitudes. *Psychological Science, 17*, 954–58.

Saucier, D. A., & Miller, C. T. (2003). The persuasiveness of racial arguments as a subtle measure of racism. *Personality and Social Psychology Bulletin, 29*, 1303–15.

Sedikides, C., & Gebauer, J. E. (2010). Religiosity as self-enhancement: A meta-analysis of the relation between socially desirable responding and religiosity. *Personality and Social Psychology Review, 14,* 17–36.

Shariff, A. F., Cohen, A. B., & Norenzayan, A. (2008). The devil's advocate: Secular arguments diminish both implicit and explicit religious belief. *Journal of Cognition and Culture, 8*, 417–23.

Sharp, C. A., Rentfrow, P. J., & Gibson, N. J. S. (2015). One God but three concepts: Complexity in Christians' representations of God. *Psychology of Religion and Spirituality.* Advance online publication. dx.doi.org.

Sherman, J. W., Gawronski, B., & Trope, Y. (eds). (2014). *Dual process theories of the social mind.* New York: Guilford Press.

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*, 165–79.

Teige-Mocigemba, S., & Klauer, K. C. (2008). "Automatic" evaluation? Strategic effects on affective priming. *Journal of Experimental Social Psychology*, *44*, 1414–17.

Tremlin, T. (2006). *Minds and gods: The cognitive foundations of religion*. Oxford, UK: Oxford University Press.

Trimble, D. E. (1997). The Religious Orientation Scale: Review and meta-analysis of social desirability effects. *Educational and Psychological Measurement*, *57*, 970–86.

Uhlmann, E. L., Poehlman, T. A., & Bargh, J. A. (2008). Implicit theism. In R. M. Sorrentino & S. Yamaguchi (eds.), *Handbook of Motivation and Cognition across Cultures* (pp. 71–94). Oxford, UK: Elsevier.

Vargas, P., von Hippel, W., & Petty, R. E. (2004). Using "partially structured" attitude measures to enhance the attitude-behavior relationship. *Personality and Social Psychology Bulletin, 30*, 197–211.

Wenger, J. L. (2004). The automatic activation of religious concepts: Implications for religious orientations. *International Journal for the Psychology of Religion, 14*, 109–23.

Wentura, D., & Degner, J. (2010). A practical guide to sequential priming and related tasks. In B. Gawronski & B. K. Payne (eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 95–116). New York: Guilford Press.

Wigboldus, D. H. J., Holland, R., & van Knippenberg, A. (2006). *Single target implicit associations*. Unpublished manuscript, Radboud University, Nijmegen, The Netherlands.

Wittenbrink, B., Judd, C. M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, *72*, 262–74.

Wittenbrink, B., & Schwarz, N. (2007). Introduction. In B. Wittenbrink & N. Schwarz (eds.), *Implicit measures of attitudes* (pp. 1–13). New York: Guilford Press.

Yarborough, C. A. (2009). *Depression and the emotional experience of God.* Unpublished doctoral dissertation, Regent University, Virginia Beach, VA.

Zahl, B. P. (2013). *A social cognitive investigation of anger toward God.* Unpublished doctoral dissertation, University of Cambridge, UK.