Implicit and Explicit Self-Esteem: What Are We Measuring?

ROMIN W. TAFARODI
CAROLINE HO
University of Toronto

Abstract

For nearly 60 years, researchers and practitioners have struggled toward agreement on the definition and measurement of self-esteem. Both consensus and precision have proven elusive, and debate about what we are or should be measuring with our instruments continues today. In this article, we offer a clarifying account of the nature of self-esteem as a key aspect of personal identity and examine its legitimacy as a hypothetical construct. The distinction between implicit and explicit self-esteem is discussed in this context, raising critical questions about the theoretical status of the former.

Any discussion of the validity of a test presupposes a definite quality to be measured. An easy separation of measurement from definition, however, is possible only when the quality reduces to purely operational terms that are rarely contested (e.g., distance, conductivity, blood pressure). Moreover, the "routine measurement" of even such qualities is intimately tied to supporting networks of theoretical and empirical relations that ultimately blur the distinction between defining the quality and measuring it (Cliff, 1982). This interdependence becomes all the more clear for psychological qualities. Here, the question of whether a test is working as intended usually cannot be answered by referring to a single operationally defined criterion. Rather, one is forced to take up the more fundamental task of trying to determine just what the test is measuring and whether this quality or qualities are consistent with the claims of the designer. This is accomplished by building around the test a "nomological net" of confirmed relations to pertinent constructs and related measures (Cronbach & Meehl, 1955). The validity of the test is reflected in the extent to which

the network is logically consistent with the working definition of the quality that is the target of measurement and the inferences that follow from this working definition. Thus, "construct validity" supports both the definition of the quality and one's success in measuring it. A test that fails to demonstrate a crucial relation posited in its nomological net may reflect either poor measurement or misunderstanding of the quality being measured. This presents a dilemma for the researcher who is confronted with a critical disconfirmation. How does one proceed? Should the test be improved? The quality redefined? Both? The measurement of self-esteem over the past half-century has been marked by this programmatic dilemma and the dynamics of uncertainty and change are no less prominent today. In the hope of providing a useful starting point for making sense of recent developments in this continuing evolution – especially the current emphasis on implicit versus explicit measurement – we offer a brief conceptual analysis.

Self-Esteem as a Hypothetical Variable

We begin by specifying the kind of quality that self-esteem represents. A long-standing distinction proposed by MacCorquordale and Meehl (1948) can be applied here. On one hand, there are psychological variables that amount to nothing more than the quantitative or qualitative relations of more primary observable entities. They contain no surplus content and are merely formal or informal abstractions from empirical observations or relations. These "intervening variables" are exemplified by dispositional concepts such as *plasticity, dominance,* and *reactivity.* Clearly, these are not entities, processes, or events of any kind. They are summary abstractions of behavioural patterns. On the other hand, there are psychological variables that entail the existence of entities, processes, or events that are not directly observable but are posited to explain that which is observable. Often these "hypothetical variables" begin as metaphors needed to fill theoretical gaps. Over time, many become reified into self-standing concepts with existential content. Historical examples in psychology are *engram, libido,* and *iconic store.* The key distinction, then, lies in the ontic implications of the variables. Intervening or "abstractive" variables do not require that there *be* something above and beyond the constituent empiri-

cal variables that are abstracted, combined, or related. In many cases, intervening variables can be thought of as a form of "shorthand." Hypothetical variables, in contrast, point to the existence of something that cannot be fully reduced, decomposed, or redefined in terms of other variables. They pick out something new, or at least distinct, in the field of what is there to be described. This applies irrespective of what the hypothetical variable refers to, whether entity, process, or event. Which type of variable, then, is self-esteem?

It seems obvious to us that self-esteem is more than an intervening variable. As a conceptual aspect of identity, it exists only in and through its "aspectual shape" as an "intrinsic intentional state" (Searle, 1992). This is to say that the qualities of self-esteem are given by its appearance as a conscious experience. Even if one were to argue for "unconscious" self-esteem, it would presumably be defined it terms of its intentional content as a *possible* conscious state (e.g., "He *hates* himself unconsciously"). The point to recognize here is that self-esteem does not refer to a *representation* in the mind of something assumed to exist beyond it, such as a rock or a butterfly. It is defined completely by its subjective form. Furthermore, the conceptual address of self-esteem is that of one's worth as a person with a particular identity. Such an encompassing moral understanding of oneself requires a linguistic framework through which the interpretive synthesis and elaboration of narrative identity construction can occur. In this regard, self-esteem is the result of creative symbolic activity upon traces of the past. This activity produces a distinct evaluative belief about one's own personhood, one that is accompanied by strong feelings. As a product of self-interpretation, self-esteem refers to an identifiable intentional state; it is not merely a scientific or folk abstraction of conscious or physical states. It exists in its own right. For this reason, self-esteem can be considered a hypothetical variable. We expand on this understanding below.

Most researchers and laypersons begin with the understanding that however else it might be interpreted, self-esteem is first and foremost a reflexive phenomenon whereby the individual apprehends his or her own value.[1] This apprehension appears to be critical to knowing who one is in a shared world of symbol-

ic meaning. Personal identity entails understanding how where one presently stands is defined by past positions – remembered intentions, feelings, actions, perceptions, hopes, relationships, physical characteristics, etc. To "know oneself" in this fundamental, prosaic sense requires knowing where one has come from and where one is going, or would like to go. This is the essence of personal identity, which is inherently narrative and even "mythic" (McAdams, 1993) in form. The loss of self that follows the loss of narrative memory is tragically illustrated by Sacks's (1970) "lost" or confabulating Korsakoff patients. They simply do not know who they are.

To understand what the narrative nature of identity means for self-esteem, imagine reading an autobiography. As you read, you inevitably form a critical moral understanding of the author that takes into account both failings and virtues. If someone were to ask you afterward what you "thought of" the author, you would be considered odd in any culture or language to respond, "I think the author was born in Michigan, was a devout Catholic, suffered measles as a child, worked as a paperboy, briefly dated the high school principal's daughter, failed a university calculus course, became an architect, married late, fathered three devoted daughters, and died from stroke at the age of 82." Rather, you would more typically offer in your response an evaluation of the author's character, one that goes beyond a descriptive account of his or her actions and shifting social roles. The evaluation would be relational in that it describes not only qualities inferred from the author's thoughts, actions, and roles, but also the form of interest you took in the author as a result (praise, condemnation, respect, pity, curiosity, lust, disdain, etc.). For example, if you were to describe the author as generous, you would not merely be describing generosity, but also admiring it. Furthermore, you would not be admiring the *generosity*, but the *author* for being generous. Spontaneous, everyday valuation of persons is relational in that it serves to position oneself against another within a common moral space. Self-esteem is no different in this respect. The valuation is of oneself by oneself. That is, the *I* of self-consciousness "thinks about" its own *me* narrative and stands in moral relation to the subject of the narrative as if toward another (Mead, 1934; Ricoeur, 1992). This process is analogous to an autobiographer coming to know himself/herself through his/her own writing. His/her new understanding transforms the meaning of the narrative, perhaps even prompting the author to change what is already written. So too with one's own narrative, where self-valuation is continuous and recursive within a hermeneutic circle of reflection on the pre-

---

1 One of us has argued elsewhere (Tafarodi & Milne, 2002; Tafarodi & Swann, 2001) that the experience of one's own value includes both an amoral, immediate sense of oneself as *effective* or not and a more considered *moral* sense of where one stands in relation to the good. In the present conceptual analysis, we restrict our attention to the latter dimension, "self-liking," which has been the main focus of self-esteem research in recent decades.

sent in terms of the past and the past in terms of the present. To summarize, knowing oneself is inherently valuative in that it involves a moral stance to one's own narrative identity. The moral stance one assumes is both a reflection on narrative identity and the source of its revision. Such reflection is symbolically mediated, as only the concepts of language could provide the moral framework that makes the valuation of character possible. Cats, dogs, and horses do not and cannot understand themselves as moral agents.

Reflection on one's own value is conscious and reflexive, although not always voluntary or deliberate. A person is no more able to engage in *unconscious* consideration of her/his own moral significance as outlined here than to unconsciously read and reflect on the autobiography of another. Reflexive consciousness refers to awareness of one's own intentional states, not the processes and elements that underlie and influence these states (Jaynes, 1976). (Just as one can be fully aware that one is reading an autobiography and taking a position toward its author yet remain wholly ignorant of the visual and cognitive operations that make all this possible.) Self-esteem, defined here as a reflective moral stance taken toward oneself, is therefore a necessarily conscious phenomenon. Its exists as a mental phenomenon *through* being experienced; it has what Brentano (1874/1995) referred to as "intentional inexistence" – existence *within* consciousness as a direction or intention of the mind toward itself. We have argued that the nonreducible intentional content of self-esteem identifies it as a hypothetical rather than intervening variable. This conclusion has implications for measurement.

Explicit Measurement

The first implication of the above analysis is that any *direct* measurement of self-esteem as a private moral stance must capture the individual's conscious act of self-judgment, episodic memory for this act, or semantic memory for its propositional derivatives. Anything short of this is either an *indirect* measure (remote indicator or correlate) of self-esteem and/or a direct measure of some other construct. To illustrate, imagine that you are asked whether you like yourself. There are at least three ways to answer the question appropriately. The first is to engage in a brief episode of moral self-examination and report on the resulting judgment. The second is to recall past instances of such examination and report the average outcome. Finally, you could recall your long-standing belief that you like yourself, though you may be unable to recall a single instance of how and in what context this propositional knowledge was inferred. We can assume, however, that it was inferred through

reflection, not pulled from thin air. This is because the consideration of one's value as a person is not an optional activity in the interpretive construction of narrative identity, but is integral to it. There can be no amoral narratives and no amoral characters "emplotted" within them. Knowing where one stands in moral terms is of perennial and constitutive importance for personal identity (Taylor, 1989). It is therefore safe to assume that no one engages with his/her self-esteem for the first time upon direct questioning. Every respondent has an extended history of such reflection to draw upon. This history has given rise to propositional self-knowledge – *beliefs* about one's worth that can be readily accessed, acknowledged, and considered. Therefore, researchers most often rely on the third strategy described above, which assumes respondents' ability to bring to mind and report on preexisting beliefs about their personal worth or value.

Direct or explicit questioning, as illustrated above, has been the most common approach to measuring self-esteem since the publication 58 years ago of the first self-esteem instrument (Raimy, 1948). Direct measures require respondents to indicate the extent to which they agree with first- or second-person statements taken as defining or implying a positive or negative valuative stance. Such statements may capture an overall stance (e.g., *I feel good about myself; I am unworthy*) or some more specific belief, attitude, tendency, or experience presumed to support or justify the overall stance (e.g., *I feel good about my appearance; I am an unworthy father*). Graded or categorical agreement is summated across statements to form a single score or several domain-specific subscores (often hierarchically arranged), which are then used as quantitative indices of self-esteem. Such scores are rarely if ever confirmed as anything more than ordinal in metric form, but are nonetheless treated as interval scales for purposes of statistical analysis. Popular examples include Rosenberg's (1965) Self-Esteem Scale, Coopersmith's (1967) Self-Esteem Inventory, and Janis and Field's (1959) Feelings of Inadequacy Scale.

The virtue of direct questioning is its immediacy. We gain insight into a person's self-esteem by asking her/him to reflect upon herself/himself. When such "questioning" is accomplished using statements in the present indefinite tense (*I feel/think/believe…*), the respondent tends to rely on memory for past episodes of moral self-reflection, and even more so, the self-defining beliefs that arose from those episodes. The obvious drawback of direct questioning is its susceptibility to misrepresentation by the respondent, as is at issue in any dialogical encounter. Given the West's increasingly emphatic valorization of self-satisfaction and feeling good about oneself (Giddens, 1991;

Hewitt, 1998; Kaminer, 1993), it is hardly surprising that direct self-reports of self-esteem converge to some extent with the tendency toward socially desirable responding (Paulhus, 1991, 2002). A number of distinctions deserve consideration here. Conscious or deliberate misrepresentation is fairly straightforward. A person responds insincerely to self-evaluative statements to avoid being perceived as unattractive, weak, insecure, troubled, or self-critical. She/he is keenly aware of private self-doubt, ambivalence, or perhaps even loathing, but does not wish to respond in a manner that would reveal this subjective reality. Less commonly, the respondent misrepresents downwards, reporting more negativity than is privately experienced. This might be occasioned by a deep concern with public modesty, a moral emphasis more prominent in non-Western cultural contexts.

A second form of misrepresentation is less straightforward. The conscious concealment of a reflexive moral stance can, in time, become habitual. The more practiced and habitual such behaviour becomes, the less self-awareness and deliberation it requires. Eventually, it may not be experienced by the respondent as misrepresentation at all. It becomes what Goffman (1959) characterized as "sincere performance." It must be noted, however, that over-learned, automatic, or "mindless" dissimulation of this sort does not constitute a second form of self-esteem alongside the one that it serves to conceal. At least not according to the present positioning of self-esteem as a hypothetical variable, one with distinct and irreducible intentional content derived from a lifetime of private moral reflection on one's narrative identity. However "natural" and effortless a dissimulation becomes, it remains a hollow performance of a social role, that of an actor communicating an acceptable level of self-esteem to a definite or indefinite audience that includes the actor. It does not rest upon any synthetic and historic process of self-interpretation.

The self-esteem of the *self-deceptive* narcissist is distinct from both forms of misrepresentation described above. We are referring here to the defensively skewed construction of, and reflection on, a personal narrative over time. Relevant is Fingarette's (1969) account of the self-deceiver as one who, for ego-defensive reasons, does not "spell out" the meaning of certain of his own transactions or "engagements" in the world and does not identify with ("avow") and take responsibility for these engagements. The narrative gaps produced by this "policy" of exclusion are filled by "sincere" fabrications, sincere because the stories "he tells us he also tells himself." Interpretive exclusion, disavowal, and fabrication all point to an acutely biased moral stance toward oneself. We can refer to this as

self-deceptive self-esteem. The full sincerity of its communication and its grounding in interpretive synthesis distinguish it from both conscious and mindless misrepresentation, as discussed above.

What happens to the unacknowledged engagements of the self-deceiver? Do they become the grist for an unconscious hermeneutic agent of moral reflection, which then generates a second, incompatible form of self-esteem? (This, of course, presupposes the unconscious weaving of a second, incompatible personal narrative.) Such reduplication strains credibility. Fingarette (1969) makes no such claim. For him, as for us, there can be no spelling out and moral interpretation outside of language and explicit consciousness. The positing of *two* parallel histories of moral reflection, one self-deceptive and the other "realistic," amounts to a schizoid characterization of intentionality. This is the very sort of confusion that Fingarette's account is aimed at avoiding. Equally problematic is the modeling of self-deception on other-deception, whereby the former is described as one mental subagency "deceiving" another. The homuncularist paradox of this conception can only be resolved through deflation of the deceiver into a dumb process – a "mental tropism" devoid of intentionality (see Johnston, 1991). This tactic clearly will not work in the present case, as the moral reflection required to create a realistic self-esteem beneath its conscious narcissistic counterpart entails a subagency complete with intentionality. Even the "self-deceiver," then, is limited to a single self-esteem to report, however unjustified that self-esteem may appear to an observer. There is no second, realistic self-esteem lying repressed beneath it. Insofar as the self-deceiver gains insight into his/her own deceptions, his/her self-esteem may well change. This would occur through the redactive incorporation within the personal narrative of previously unacknowledged engagements, including the practice of self-deception itself. Such change, however, is a reconstructive and reinterpretive process, not the replacement of one self-esteem by another.

The three cases described above, and their implications for measurement, can be summarized as follows. The explicit measurement of self-esteem is aimed at gauging the individual's private self-valuation as a person, a moral being. This valuation is not fixed, but reflects the hermeneutic imperative of understanding oneself as an identity through time, a character in a narrative of experience and engagement, one who is both a subject and an agent of change. Self-esteem is necessarily an integrative and synthetic understanding. Psychometric reliance on the direct communication of self-esteem can lead to mismeasurement in two

ways. The first is simple insincerity, where the respondent knowingly dissimulates with the clear intention of concealing a private experience of self-esteem. The second is overpracticed insincerity, which is no longer experienced as such by the actor. Rather, responses to test questions are made in a blindly habitual manner that is inconsistent with the private experience of self-esteem. Such a tendency can become so overlearned that the private experience is not even brought to mind in the act of responding. As such, there is no subjective sense of concealment. Importantly, neither of these two forms of mismeasurement or invalidity assumes more than a single, explicitly conscious self-esteem. In both cases, this self-esteem is not the basis of the response; however, neither is it unconscious in any dynamic sense. Even the most practiced performer who could bring her/his own private self-esteem to mind if so motivated; the force of mindless habit allows her/him to avoid having to do so. Genuinely self-deceptive responding presents a more complex, third case for analysis. It can be understood as the sincere expression of a single self-esteem founded on a defensively and selectively constructed personal narrative. There is no *psychometric* invalidity here; the problem is one of existential integrity, of which Kierkegaard, Sartre, and Heidegger had much to say. Thus, in each of the three distinct cases analyzed, we are dealing with a single self-esteem. Where does this leave current thinking about "implicit" self-esteem?

Implicit Measurement

Over the past decade, experimental social cognition has witnessed the burgeoning of research on implicit or unconscious processes (Fazio & Olson, 2003). This development reflects the subdiscipline's disenchantment with introspection, phenomenology, and natural discourse; its heightened preoccupation with irrationality and ego-defense; its fascination with what might be called alienated thought (mindlessness and automaticity); increased doubts about the predictive utility of verbal reports; the functionalist deflation of subjectivity; and a dim but disquieting sense of the "eliminative materialism" (Churchland, 1981) implied to some by advances in neuropsychology. One theme in this movement has been an intensive program of research on "implicit self and identity" (Devos & Banaji, 2003), with considerable focus on implicit self-esteem in particular. The latter was defined by Greenwald and Banaji (1995) as "the introspectively unidentified (or inaccurately identified) effect of the self-attitude on evaluation of self-associated and self-dissociated objects" (p. 11). This definition is significant in that it represents the reduction of self-esteem to an "effect" in a process that occurs outside awareness. Self-esteem is no longer the conscious articulation of one's identity through narrative interpretation and moral reflection, but the causal relation of a hypothetical "self-attitude" whose subjective character and origins are left murky. Reduced in this way, self-esteem has been dessicated and "demoralized" (Greer, 2003). A similar reduction to theoretical relation – here semantic "association" – is given by Greenwald et al. (2002): "Self-esteem is the association of the concept of self with a valence attribute" (p. 5). In reading this, one has the impression that self-esteem is being modeled as artificial intelligence inside Searle's (1990) famous Chinese room! This is the very sort of intentionality-dodging functionalism that Malcolm (1979/80) criticized for its failure to capture the meaningfulness of human experience. Clearly, this is not the same quality of mind that we have been describing in this paper. The difference appears to be categorical. This concern aside, however, there can be little doubt that implicit or indirect measurement of self-esteem sidesteps the problem of misrepresentation that vitiates direct self-report. It does so by indexing behaviour that is not subject to conscious control, occurs outside of awareness, or, at least, is not transparently indicative of self-esteem. Such implicit measures, however, have their own interpretive problems (Bosson, Swann, & Pennebaker, 2000; De Houwer, 2001, 2002; Gregg, 2003; Karpinski, 2004; Mierke & Klauer, 2003). We will discuss only one broad concern in the context of this paper: the conceptual ambiguity of the constructs these tests are assumed to indirectly measure. Is implicit self-esteem as measured by these tests to be taken as a theoretical relation only, a measurement process, or more? To present our arguments, we will focus on the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998), an increasingly popular associative measure that has shown adequate reliability and validity in application to self-esteem (Greenwald & Farnham, 2000).

The IAT for self-esteem is essentially a measure of semantic association that is premised on the principle of response compatibility. The logic is clever. In one critical block of the test, a series of *pleasant, unpleasant, me,* and *not-me* words is presented. The individual must categorize each word as either "pleasant or me" or "unpleasant or not-me." The speed of response is then compared against another critical block where comparable words are categorized as either "pleasant or not-me" or "unpleasant or me." To the extent that the individual's self-concept is associated with positively but not negatively valenced nodes in memory, the second set of compound categories represents incompatible pairings, and decisions based on them therefore should take longer. Hence, the difference in speed of

response between the two critical blocks is taken to reflect the positivity of the self-concept, or what has been termed "implicit self-esteem." This makes some sense, although the inference from simple associations to moral predicates is a logical leap. To illustrate, imagine that you spent the next year with the word "BAD" written in large red letters on your hand. You would almost certainly strengthen the association between *me* and *unpleasant* words. This would occur, however, without jeopardizing your self-esteem in the least (one would hope). A simple association can arise for any of a host of reasons. Inferring from it a specific historical process – in this case a distinctive form of moral experience – is a hazardous exercise.

A second issue arises that is more relevant to our analysis: Is the self-esteem measured by the IAT to be taken merely as *indirectly measured* self-esteem as we have defined it here, or should it be considered a self-standing, subterranean form of self-esteem that is dissociated from conscious moral reflection? Farnham, Greenwald, and Banaji (1999) suggest the latter in describing it as a "construct of self-regard" that is "unavailable to introspection" (p. 244). Elsewhere, implicit measurement of self-esteem is claimed to "define constructs that are distinct from, although correlated with, nominally the same constructs measured by self-report" (Greenwald & Farnham, 2000, p. 1034). This would seem to imply two fully fledged self-esteems within the same mind – one conscious and the other not – rather than merely better or worse measurement of one integrative moral stance that is inherent to human agency and personal identity (Taylor, 1985). Commitment to a single self-esteem, it must be said, does nothing to discount the possibility of ambivalence toward the self. One's reflexive moral stance surely changes over time, just as what one "thinks of" the author of an autobiography changes as the narrative unfolds, character is revealed, the meaning of past actions and events become apparent or are reconsidered, and unifying patterns emerge. At any point along the way, one may have mixed feelings about the author. But each feeling in the mix is no more a separate moral stance than the colours of a painting are separate paintings. Similarly, moral reflection on one's own personhood is a synthetic act that renders complex configurations of evaluative thought and feeling meaningful within a culturally bound symbolic framework; it is not the isolated imprint of a vague feeling, dissociated memory, repressed thought, or spontaneous inference on one's behaviour.

We have no quarrel with those who claim that implicit tests of self-esteem are simply indirect measures, of greater or lesser validity, of the construct we have attempted to elucidate in this paper. However, we stand opposed to the further claim that implicit tests are tapping into a second "unconscious" form of self-esteem that is distinct from and possibly inconsistent with its conscious counterpart. This amounts to the introduction of an ill-defined construct that fails to comport with the hermeneutic framework through which self-esteem is articulated as a central moral feature of personal identity. If the proponents of implicit measurement believe they are measuring something distinct from self-esteem as a subjectively identifiable intentional state requiring symbolic self-interpretation, then they are misguided in calling it self-esteem. To conflate two categorically distinct constructs under a single term taken by most to imply moral reflection is to tangle two evolving nomological nets to the detriment of both. The original impetus for developing implicit (*qua* indirect) measures of self-esteem was the misrepresentation that corrupts direct self-reports. If so, then perhaps greater attention should be given to how we might convince our research participants to avoid posturing, overpracticed protectiveness, and conventionality in their responses. The diversion of turning to the unconscious in the hope of measuring a putatively distinct form of moral self-awareness will only serve to confuse matters.

Correspondence should be adressed to Romin W. Tafarodi, Department of Psychology, University of Toronto, 100 St. George Street, Toronto, Ontario M5S 3G3 (Tel: 416-946-3024; Fax: 416-978-4811; E-mail: tafarodi@psych.utoronto.ca).

Résumé

Pendant près de soixante ans, les chercheurs et les praticiens ont tenté de s'entendre sur la définition et la mesure de l'estime de soi. Il n'a pas été possible d'en arriver à un consensus et la précision demeure insaisissable de telle sorte que le débat entourant ce que nous mesurons ou devrions mesurer avec nos instruments se poursuit de nos jours. Dans cet article, nous présentons un compte rendu qui clarifie la nature de l'estime de soi comme un aspect clé de l'identité personnelle et nous examinons sa légitimité en tant que construit hypothétique. La distinction entre l'estime de soi implicite et explicite fait l'objet de discussion dans ce contexte, soulevant des questions critiques quant à l'état théorique de cette distinction.

References

Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology, 79*, 631-643.

Brentano, F. (1995). *Psychology from an empirical standpoint* (A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Trans.). London: Routledge. (Original work published 1874).

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78*, 67-90.

Cliff, N. (1982). What is and isn't measurement. In G. Keren (Ed.), *Statistical and methodological issues in psychology and social science research* (pp. 3-38). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coopersmith, S. (1967). *The antecedents of self-esteem.* San Francisco, CA: Freeman.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology, 37*, 443-451.

De Houwer, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33*, 115-133.

Devos, T., & Banaji, M. R. (2003). Implicit self and identity. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 153-175). New York: Guilford Press.

Farnham, S. D., Greenwald, A. G., & Banaji, M. R. (1999). Implicit self-estem. In D. Abrams & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 230-248). Oxford, UK: Blackwell.

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

Fingarette, H. (1969). *Self-deception.* London: Routledge & Kegan Paul.

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age.* Stanford, CA: Stanford University Press.

Goffman, E. (1959). *The presentation of self in everyday life.* New York: Doubleday.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.

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*, 1022-1038.

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*, 1464-1480.

Greer, S. (2003). Self-esteem and the demoralized self: A genealogy of self research and measurement. In D. B. Hill & M. J. Kral (Eds.), *About psychology: Essays at the crossroads of history, theory, and philosophy* (pp. 89-108). Albany, NY: State University of New York Press.

Gregg, A. P. (2003). Optimally conceptualizing implicit self-esteem. *Psychological Inquiry, 14*, 35-38.

Hewitt, J. P. (1998). *The myth of self-esteem: Finding happiness and solving problems in America.* New York: St. Martin's Press.

Janis, I. S., & Field, P. B. (1959). A behavioral assessment of persuasibility: Consistency of individual differences. In C. I. Hovland & I. L. Janis (Eds.), *Personality and persuasibility* (pp. 55-68). New Haven, CT: Yale University Press.

Jaynes, J. (1976). *The origin of consciousness in the breakdown of the bicameral mind.* New York: Houghton Mifflin.

Johnston, M. (1991). Self-deception and the nature of mind. In D. Kolak & R. Martin (Eds.), *Self and identity: Contemporary philosophical issues* (pp. 422-435). New York: Macmillan.

Kaminer, W. (1993). *I'm dysfunctional, you're dysfunctional: The recovery movement and other self-help fashions.* New York: Random House.

Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin, 30*, 22-34.

MacCorquodale, K., & Meehl, P. E. (1948). On the distinction between hypothetical constructs and intervening variables. *Psychological Review, 55*, 95-107.

Malcolm, N. (1979/80). 'Functionalism' in philosophy of psychology. *Proceedings of the Aristotelian Society, 80*, 211-229.

McAdams, D. P. (1993). *The stories we live by: Personal myths and the making of the self.* New York: William Morrow and Company, Inc.

Mead, G. H. (1934). *Mind, self, and society from the standpoint of a social behaviorist.* Chicago, IL: University of Chicago Press.

Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology, 85*, 1180-1192.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological*

and educational measurement (pp. 49-69). Mahwah, NJ: Lawrence Erlbaum Associates.

Raimy, V. C. (1948). Self reference in counseling interviews. *Journal of Consulting Psychology, 12,* 153-163.

Ricoeur, P. (1992). *Oneself as another* (K. Blamey, Trans.). Chicago, IL: University of Chicago Press. (Original work published 1990).

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Sacks, O. (1970). *The man who mistook his wife for a hat and other clinical tales.* New York: Simon & Schuster.

Searle, J. R. (1990, January). Is the brain's mind a computer program? *Scientific American, 262,* 26-31.

Searle, J. R. (1992). *The rediscovery of the mind.* Cambridge, MA: MIT Press.

Tafarodi, R. W., & Milne, A. B. (2002). Decomposing global self-esteem. *Journal of Personality, 70,* 443-483.

Tafarodi, R. W., & Swann, W. B., Jr. (2001). Two-dimensional self-esteem: Theory and measurement. *Personality and Individual Differences, 31,* 653-673.

Taylor, C. (1985). *Human agency and language: Philosophical papers 1.* Cambridge, UK: Cambridge University Press.

Taylor, C. (1989). *Sources of the self: The making of the modern identity.* Cambridge, MA: Harvard University Press.