Chapter 9

The Self in Cognitive Behavior Therapy

Hamish J. McLeod & Joseph Ciarriochi
University of Wollongong, Australia

Introduction

Approaches within traditional cognitive behavior therapy share certain assumptions concerning human psychology. They assume that the content of one's cognitions concerning the self or otherwise has a direct effect on emotion and observable behavior, that it is sometimes possible for cognitions to be incorrect or distorted in relation to "objective" truth, and that incorrect or distorted cognition may produce emotional and behavioral problems (see e.g., Abramson, Seligman, & Teasdale, 1978; Beck, Rush, Shaw, & Emery, 1979; Bentall, Corcoran, Howard, Blackwood, & Kinderman, 2001). Given these assumptions, the challenging of incorrect or unhelpful cognitions is taken as a central aim of traditional CBT

treatment strategies. This emphasis on cognitive change assumes that emotion regulation is achieved through changing thoughts, attitudes, or beliefs (Hofmann & Asmundson, 2008). This has been referred to as a mechanistic model of human cognition that can be contrasted with the functional contextualist conception that underlies contextual behavioral science (Hayes, Strosahl, & Wilson, 1999). Therapies based on the latter, such as acceptance and commitment therapy (ACT), eschew a focus on cognitive content in favor of achieving change through the alteration of the context in which thoughts are experienced.

Granted, there are important differences between traditional CBT and contextual behavioral approaches such as ACT. Nevertheless, the recognition within the traditional CBT model of the potential effects of cognitive content, particularly those concerning the self, and the analysis of those effects on behavior under certain circumstances (referred to within ACT as "cognitive fusion") makes a contribution to the understanding of human psychology and in particular psychopathology that offers insight even for proponents of alternative conceptions such as contextual behavioral science.

Although the contemporary CBT literature refers to various self-related concepts such as self-esteem, self-focused attention (e.g., rumination), and negative self-talk, it would be an error to treat the self as a unitary "thing" that is prodded and molded into better shape by CBT. Instead, there are multiple aspects of the self that can be described in relation to their content (e.g., critical thoughts directed at the self) and the processes that operate on that content (e.g., biases in automatic attentional or memory processes for self-referent information). In this chapter we will describe the evolution of the self in the CBT literature from its largely descriptive early formulations to the more multifaceted and multilevel contemporary conceptualizations. This includes consideration of how the distinction between conscious and unconscious information processing has influenced cognitive models of distorted self-representation such as that seen in disorders such as schizophrenia. We will also examine the points of agreement and departure between the clinical conceptualizations of the self as it is described in CBT treatment guidelines and the findings from social-cognitive psychology research on the self. We will suggest that a somewhat unitary view of the self in CBT born out of clinical observation has been replaced with a multidimensional set of cognitive processes and structures that span multiple aspects of memory and levels of conscious awareness.

In addition, as the seco
cate, despite differences in
and research on the self are
contextual behavioral scier
be increasing recognition
cepts important in context
cal flexibility, meta-cogniti
implicit or "unconscious"
chapter also.

Content-Based Views o

CBT approaches give the cc
chain of causality between
response. The content of co
are implicated in numerou
(e.g., self-discrepancy theo
cumscribed clinical disorde
1996). The general cognitiv
et al., 1979; J. S. Beck, 199
event that are implicated i
ment. These are automat
beliefs.

Automatic Thoughts

Automatic thoughts ar
ideas and images that occu
experience. They cannot
degree of cognitive control
ing. However, one of the tas
explicitly aware of negative
The association between
behavioral responses is expl
nale for challenging the ac
Beck, 1995). This is consi
helping the client to change
emotions or observable beh

In addition, as the second half of this chapter in particular may indicate, despite differences in assumptions, in some respects CBT theory and research on the self are showing a degree of convergence with current contextual behavioral scientific concepts. For example, there appears to be increasing recognition from within traditional CBT of several concepts important in contextual behavioral science, including psychological flexibility, meta-cognition, the observer perspective, and the effects of implicit or "unconscious" processes. Each of these is discussed in this chapter also.

Content-Based Views of the Self in CBT

CBT approaches give the content of cognition a privileged position in the chain of causality between a stimulus and an emotional or behavioral response. The content of cognitions, appraisals, and beliefs about the self are implicated in numerous broad cognitive models of psychopathology (e.g., self-discrepancy theory; Higgins, 1987) and specific models of circumscribed clinical disorders (e.g., Clark's model of panic disorder; Clark, 1996). The general cognitive therapy model (A. T. Beck, 1976; A. T. Beck et al., 1979; J. S. Beck, 1995) distinguishes three main types of mental event that are implicated in psychopathology and are targeted in treatment. These are automatic thoughts, intermediate beliefs, and core beliefs.

Automatic Thoughts

Automatic thoughts are the ubiquitous, reflexive, often fragmentary ideas and images that occur for all humans as a part of everyday mental experience. They cannot be "switched off" or subjected to the same degree of cognitive control as more reflective deliberate forms of thinking. However, one of the tasks in CBT is to help the patient become more explicitly aware of negative automatic thoughts (NATs; A. T. Beck, 1976). The association between these thoughts and negative emotional and behavioral responses is explicitly taught to the patient as part of the rationale for challenging the accuracy or utility of the thought content (J. S. Beck, 1995). This is consistent with the mechanistic assumption that helping the client to change their thought content will lead to a change in emotions or observable behaviors.

Although the range of automatic thought content is potentially vast, this diversity is often simplified by referring to common co-occurring patterns of thinking. The cognitive (or depressive) triad of negative thoughts about the self, the world, and the future is one of the most widely used examples of this (A. T. Beck et al., 1979). The impact of negative thoughts on the self is that patients may exhibit a habitual pattern of self-criticism that is treated as an objectively true representation of the status of the self and contributes to the maintenance of low mood (Dozois, 2007). However, as there is considerable individual variation in the specific content of this negative self-talk, CBT theorists posit a role for more enduring cognitive structures from which these negative automatic thoughts emanate. This leads to the specification of various forms of belief, broadly separated into intermediate and core beliefs.

Intermediate Beliefs

Intermediate beliefs are classified into subtypes of *rules* (e.g., "I should always be self-reliant"), *attitudes* (e.g., "It is terrible to have to ask for help"), and conditional *assumptions* (e.g., "If I ask for help, it means I am incompetent"). Intermediate beliefs sit on the midpoint of malleability between automatic thoughts (the most amenable to modification) and core beliefs (rigidly held unconditional beliefs that are resistant to modification).

These beliefs will be directly relevant to the dysfunction of the self when they codify an unworkable standard (e.g., "I must never upset another person") or arbitrarily impose catastrophic, exaggerated, or otherwise distorted expectations on the self (e.g., "It would be *unbearable* if another person does not like me"). CBT theorists posit that beliefs will exert an effect on affect and behavior in the abstract (e.g., the use of safety behaviors in anxiety disorders; Helbig-Lang & Petermann, 2010) as well as influencing the response to actual stressors (e.g., triggering a strong depressive response following the violation of a self-imposed rule; Showers, Limke, & Zeigler-Hill, 2004).

It is proposed that the form of intermediate beliefs affects their modifiability (J. S. Beck, 1995). Specifically, rules and attitudes are claimed to be less amenable to evaluation than conditional assumptions even though they are permutations of each other. The argument is that becoming aware of a conditional assumption (e.g., "If I ask for help, it means I am incompetent") promotes more cognitive dissonance (i.e., discomfort

caused by the presence of confli
from which the assumption is d
Beck, 1995), which in turn is am
or "irrational" thoughts with
tives. Indeed, the latter is one o
treatments are argued to work (

Core Beliefs

Although core beliefs are
mous with schemas, it has been
them as the *content* of schemas
of global and overgeneralized
tively impervious to change. Tl
asserted to be self-focused an
helplessness, unlovability, or b
generally ascribed to the self as
inadequate," "I am incompetent
beliefs can also reflect proposit
ple are devious"; (Fowler et al.,
beliefs for understanding the o
is most easily understood when
cept of schema, addressed next

Structural Views of the Se

Schemas

The schema concept in CB
chology in an attempt to explai

1 But it should be noted that the re
rigidly held dysfunctional rule even mor
dence. This problem is dealt with in m
strategies such as "rolling with resistanc
2 Unfortunately, this lack of concep
substantial variability in the way that t
Todd, & Reichelt, 2009).

caused by the presence of conflicting content) than awareness of the rule from which the assumption is derived (e.g., "I shouldn't ask for help") (J. S. Beck, 1995), which in turn is amenable to reduction by rejecting distorted or "irrational" thoughts with more evidence-based functional alternatives. Indeed, the latter is one of the mechanisms by which psychological treatments are argued to work (Tryon & Misurell, 2008)[1].

Core Beliefs

Although core beliefs are sometimes referred to as being synonymous with schemas, it has been proposed that it is more accurate to view them as the *content* of schemas (A. T. Beck, 1964)[2]. This content consists of global and overgeneralized propositions that are inflexible and relatively impervious to change. The specific content of core beliefs is often asserted to be self-focused and organized around the main themes of helplessness, unlovability, or both (J. S. Beck, 1995). These themes are generally ascribed to the self as reflected by "I" statements such as "I am inadequate," "I am incompetent," "I am defective." But, the content of core beliefs can also reflect propositions about other people (e.g., "Other people are devious"; (Fowler et al., 2006). The functional implications of core beliefs for understanding the onset and maintenance of psychopathology is most easily understood when they are examined in relation to the concept of schema, addressed next.

Structural Views of the Self in CBT

Schemas

The schema concept in CBT has been co-opted from cognitive psychology in an attempt to explain the origin of negative thoughts about the

1 But it should be noted that the response to cognitive dissonance may be to defend the rigidly held dysfunctional rule even more vigorously despite the presence of contradictory evidence. This problem is dealt with in motivational interviewing treatment by explicitly using strategies such as "rolling with resistance."

2 Unfortunately, this lack of conceptual clarity about the nature of schemas is reflected in substantial variability in the way that the concept is defined by practicing clinicians (James, Todd, & Reichelt, 2009).

self that are characteristic of diagnoses such as major depression and various personality disorders (A. T. Beck, 1964; A. T. Beck et al., 1979). J. M. G. Williams, Watts, MacLeod, and Mathews (1997) identify several features of the schema. It is a stored body of knowledge with a consistent internal structure that provides a template for organizing and making associations between new information. These structures are abstracted from experience and reflect prototypical representations of regularities between stimuli. The learned relatedness between stimuli results in the spreading of activation between schema elements such that exposure to one stimulus (e.g., "Joe complained about the service") will prime likely interpretations of the situation (e.g., that Joe feels he has waited too long for his restaurant meal). In most situations these "top down" processes act as useful heuristic models of the world that conserve cognitive resources and allow rapid responding to change. However, these predictions can be prone to bias and error when the activated schema is not appropriate to the actual situation.

Schemas are often central to the sense of self. They are models of how the self relates to the world, and especially to other people. There is extensive evidence that self-referent information is encoded and recalled more readily than non-self-referent information (Wisco, 2009) and this can have a detrimental effect on psychological functioning when self-schema content is dominated by themes of being weak, unlovable, or defective (Young, Klosko, & Weishaar, 2003).

Dysfunctional Self-Schemas

Adverse early experiences are assumed to shape self-representation schemas that can lie latent and unrecognized for years until activated by current life stressors. Once activated, this dysfunctional schema (or schemas) biases the processing of ongoing experience and provides the substrate for negative thoughts about the self, the world, and the future. In understanding depression, the proposition is that dysfunctional schemas are a form of psychological diathesis that exert minimal impact on processing during times of normal mood but may be irrationally maintained during symptom exacerbation.

> [The patient] seriously believes and is quite consistent in his beliefs that he is deprived, defective, useless, unlovable, etc. In fact, this internal consistency is often maintained in the face of repeated and dramatic external evidence contradictory to these

beliefs. The beliefs are g
to that Kuhn (1962) de
patient's observations a
by this conceptual frame
a personal paradigm ma
vidual is prepared to rec
adigm cannot accomm
paradigm. (At. T. Beck e

Two critical elements ab
tion. First, negative labels a
and maintained in the face o
emphasis on the *content* of t
are "...a form of semantic me
with the self" (Brewin, 2006
ory, global self-proposition
accessible independently fro
was acquired. Hence, a perso
about themselves (e.g., "I a
encoding experience(s) that
tion. Like other semantic k
facts that are applied to unde
pretation of the meaning of
that they are organized in a
nally consistent to the indiv
when the individual is expos

Segal (1988) delineates
schema and depression. Firs
people hold a greater numb
tive to non-depressed ind
depressed episode and are
accessibility accounts reflec
constructs and these are pr
sode. This accessibility acc
effects that demonstrate pre
is consistent with the m
Furthermore, accessibility a
ratio of negative to positive
only that the negative ones
1993). The third, *negative*
individuals differ from non

> beliefs. The beliefs are generally organized into a system similar to that Kuhn (1962) described as a scientific "paradigm." The patient's observations and interpretations of reality are molded by this conceptual framework. As in the case of scientific beliefs, a personal paradigm may be shaken and modified when the individual is prepared to recognize an anomaly that the existing paradigm cannot accommodate or evidence that disconfirms the paradigm. (At. T. Beck et al., 1979, p. 61)

Two critical elements about the self can be drawn from this description. First, negative labels are applied to the self (e.g., "I am unlovable") and maintained in the face of contradictory experiences. This reflects an emphasis on the *content* of the self-schema and reflects the idea that they are "...a form of semantic memory that describes the qualities associated with the self" (Brewin, 2006, p. 769). Like other forms of semantic memory, global self-propositions are abstracted from experience and are accessible independently from the time and place when the knowledge was acquired. Hence, a person may be consciously aware of a proposition about themselves (e.g., "I am helpless") independently of the original encoding experience(s) that contributed to the acquisition of that proposition. Like other semantic knowledge, these propositions are treated as facts that are applied to understanding the nature of the self and the interpretation of the meaning of events. The second feature of self-schemas is that they are organized in a structure of multiple elements that are internally consistent to the individual but that can be "shaken and modified" when the individual is exposed to disconfirmatory evidence.

Segal (1988) delineates three possible relationships between self-schema and depression. First, *availability* models suggest that depressed people hold a greater number of negative constructs about the self relative to non-depressed individuals and these predominate during a depressed episode and are changed by successful treatment. Second, *accessibility* accounts reflect a greater ease of access to negative self-constructs and these are preferentially accessed during a depressive episode. This accessibility account is consistent with mood congruency effects that demonstrate preferential access to memory information that is consistent with the mood state during retrieval (Blaney, 1986). Furthermore, accessibility accounts do not require any asymmetry in the ratio of negative to positive constructs available for attribution to the self, only that the negative ones are accessed more readily (Segal & Muran, 1993). The third, *negative self-schema* model proposes that depressed individuals differ from non-depressed because of a different structural

relationship or interconnectedness between constructs relating to the self. For example, the depressed person may have a high degree of interrelatedness between negative elements of knowledge about the self. Thus, activating one element of an elaborated network of negative information about the self triggers a spread of activation to related constructs (Wisco, 2009).

Only the negative self-schema model predicts that negative conceptions of the self will remain available following the remission of a depressive episode. Without an enduring negative self-schema, mood congruency effects alone will be sufficient to explain the presence of negative self-referent thoughts and beliefs during a depressive episode. Segal's (1988) review reflects many of the difficulties encountered when the structural or descriptive cognitive models used in therapy are subjected to closer empirical scrutiny. As more empirical investigations of the cognitive therapy model of self-schemas have emerged, it has become clear that emotional processes are as important as cognitive processes in shaping the course of mental illness (David & Szentagotai, 2006). It has also become apparent that simple dichotomies of positive versus negative self-schemas and singular notions of the self-concept are inadequate to explain the existing data.

Singular vs. Multiple Selves

Although there is reliable evidence that information stored in memory contributes to the subjective sense of a continuous "self" across time (Conway & Pleydell-Pearce, 2000), this is not synonymous with the existence of a unitary self-schema. Instead, there is more support for the existence of multiple self-schemas that are differentially activated depending on situational cues (Dalgleish & Power, 2004; Markus, 1990; Power, 2007; Showers, Abramson, & Hogan, 1998; J. M. G. Williams et al., 1997). This leads to the proposal that the phenomenological experience of a singular "I" is an artifice and instead that the self is composed of multiple subsystems (Dimaggio, Hermans, & Lysaker, 2010; Klein, 2010; Power, 2007). For CBT, the existence of multiple possible selves presents both propositions about the psychological substrate of emotional disorders and a target of treatment. For example, if a depressive self-schema is dominating information processing and impeding the operation of more adaptive self-schemas, then treatment could entail attempts to develop or enhance the functioning of more adaptive self-schemas.

At this point, it is necessa
cepts of self-knowledge as a d
in the brain (Klein, Cosmi
components such as "self-aspe
that pool of self-referent infor
represent information about t
without it being integrated w
"bad me" schema that comp
attitudes, and beliefs about f
This issue is addressed by pos
mation that are frequently co-
syncratic self-schema within t
This has been referred to as
modularization (Power, 2007)

Compartmentalizatior

Compartmentalization r
Power, 2004) of a set of conce
have a high degree of interco
with other aspects of the self (
compartmentalized self-sche
be related to domains of self-f
texts (e.g. "my interactions w
the self (e.g., "me as an optimi
Showers et al., 1998; Showers
aspect of a compartmentaliz
of related elements with the s
example, activating "me witl
self-evaluations such as beir
underperforming. In contras
such as "me with my spouse'
loved and being dependable.
contrasted with an "integra
and negative features co-activ

Schema compartmentali
descriptive card-sorting task
to generate as many self-aspe
are then given cards with 2(

At this point, it is necessary to make a distinction between the concepts of self-knowledge as a dissociable form of information represented in the brain (Klein, Cosmides, Costabile, & Mei, 2002) and subcomponents such as "self-aspects" (Showers et al., 1998) that exist within that pool of self-referent information. The basic principle is that we can represent information about the self (e.g., "I once got a speeding ticket") without it being integrated within a more elaborated self-schema (e.g., a "bad me" schema that comprises multiple behavioral exemplars, self-attitudes, and beliefs about failing to live up to acceptable standards). This issue is addressed by postulating that networks of self-referent information that are frequently co-activated acquire the properties of an idiosyncratic self-schema within the broader pool of self-referent information. This has been referred to as compartmentalization (Showers, 1992) or modularization (Power, 2007).

Compartmentalization

Compartmentalization refers to the "splitting off" (Dalgleish & Power, 2004) of a set of concerns about aspects of the self such that they have a high degree of interconnectedness to each other and weaker unity with other aspects of the self (Power, 2007). The affective valence of these compartmentalized self-schemas can be positive or negative and they can be related to domains of self-functioning role (e.g., "me as a parent"), contexts (e.g. "my interactions with my boss") and idiosyncratic aspects of the self (e.g., "me as an optimistic person") (Power, de Jong, & Lloyd, 2002; Showers et al., 1998; Showers et al., 2004). By definition, activation of one aspect of a compartmentalized self-schema will facilitate the activation of related elements with the same affective tone. In a negative self-schema example, activating "me with my boss" might make available negative self-evaluations such as being unassertive, liable to be criticized, and underperforming. In contrast a positive compartmentalized self-schema such as "me with my spouse" might activate self-features such as being loved and being dependable. These compartmentalized schemata can be contrasted with an "integrative" self-schema structure where positive and negative features co-activate (Showers et al., 2004).

Schema compartmentalization is typically assessed using a self-descriptive card-sorting task (Linville, 1985, 1987). Participants are asked to generate as many self-aspects as needed to describe themselves. They are then given cards with 20 different positive (e.g., outgoing, capable)

and 20 negative (e.g., weary, disorganized) adjectives and are asked to allocate those that they see as relevant to each idiosyncratic self-aspect. The descriptors can be allocated to more than one self-aspect and once the sorting is completed each self-aspect is rated on a 7-point Likert scale along dimensions of subjective importance, positivity, and negativity. The degree of compartmentalization can be statistically assessed, with perfect compartmentalization occurring when a self-aspect is ascribed entirely positive *or* negative descriptors. Other values derived from the self-concept task reflect subjective judgments about the importance of aspects of the self, self-complexity, and the proportion of negative items used across all self-aspects. Self-complexity scores reflect a combination of the total number of self-aspects generated and the degree of overlap of adjectives across those self-aspects. Hence, individuals with higher self-complexity will generate more self-aspects and show a lower degree of overlap between groups of descriptors ascribed to those self-aspects.

These ways of measuring self-representation and appraisal of aspects of the self have shed some light on the role of self-concept in clinical disorders such as depression and bipolar affective disorder. Showers et al. (1998) conducted a two-year longitudinal study of 132 university students identified at baseline as showing high or low vulnerability to developing depression. Their results indicated that self-structure and content interact to protect *against* low mood. Both the content and structure of self-concept changed in response to stressful life events and the reported amount of negative features of the self was greater when stress was higher. But, although participants who displayed a low vulnerability to depression ascribed more negative content to the self-aspects during the period of higher stress, they rated these negative self-aspects as less important and showed a greater degree of compartmentalization. Showers et al. suggest that "... [r]elegating negative beliefs to distinct or narrowly defined aspects of the self may help one perceive those beliefs as less important" (p. 491).

In summary, it appears that the flexibility of the self-concept in the face of stress (e.g., being able to compartmentalize positive and negative aspects of the self) combined with a change in the appraisal of self-aspects (e.g., viewing negative aspects of the self as less important when stressed) is associated with lower depression and dysphoria. By extension, being unable to flexibly adjust the organization of one's self-concept (i.e., high rigidity) combined with a tendency to view negative self-aspects as highly important increases the negative impact of stress on mood[3].

3 This is consistent with the emphasis within ACT and similar therapies on the important role of psychological inflexibility including rigid self-concept.

While Showers et al.'s (
the self may be an effectiv
structure in patients with
cate that high negative
diagnostic status even dur
Alatiq and colleagues foun
and greater compartmenta
trols (Alatiq, Crane, Willia
a history of unipolar depre
other two groups. Power e
observed a high degree of p
remitted bipolar patients.
comparison group of people
of chronic illness on sense
cantly lower levels of comp
the tendency to split off self
positive is not simply a resp

Overall, these results p
therapy models of the role
Showers et al.'s (1998) resul
self-concept is constructed
emotional distress. This fit
are rigidly applied across si
that remitted patients with
higher negative compartme
lematic self-organization
However, in addition to exa
in clinical disorders, there i
the information contained
psychopathology.

Perceptions and Apprai

A well-known aphorism in
events cause emotional di
some way to explaining why
sequences in different ind
within individuals; the same
ent reactions in the same in

While Showers et al.'s (1998) results suggest that compartmentalizing the self may be an effective short-term coping response, studies of self-structure in patients with unipolar and bipolar affective disorders indicate that high negative compartmentalization is associated with diagnostic status even during times of clinical remission. For example, Alatiq and colleagues found higher proportions of negative self-attributes and greater compartmentalization in bipolar patients compared to controls (Alatiq, Crane, Williams, & Goodwin, 2010). The participants with a history of unipolar depression returned scores that were in between the other two groups. Power et al. (2002) reported similar findings but they observed a high degree of positive *and* negative compartmentalization in remitted bipolar patients. Their study also included a non-psychiatric comparison group of people with chronic diabetes to control for the effects of chronic illness on sense of self. The diabetes patients showed significantly lower levels of compartmentalization, thereby demonstrating that the tendency to split off self-aspects and view them as entirely negative or positive is not simply a response to having a chronic disabling condition.

Overall, these results provide partial support for aspects of cognitive therapy models of the role of the self in the origin of affective disorders. Showers et al.'s (1998) results suggest that inflexibility in the way that the self-concept is constructed in the face of stress is associated with greater emotional distress. This fits with the proposal that negative self-schemas are rigidly applied across situations (Beck et al., 1979). Also, the finding that remitted patients with bipolar and unipolar affective disorders show higher negative compartmentalization than controls suggests that problematic self-organization can persist beyond acute phases of illness. However, in addition to examining how the self is structurally organized in clinical disorders, there is a clear need to determine how appraisals of the information contained within self-structures affects the expression of psychopathology.

Perceptions and Appraisals of the Self

A well-known aphorism in CBT is that *interpretations* or *appraisals* of events cause emotional distress, not the events themselves. This goes some way to explaining why the same stressor will provoke different consequences in different individuals. This principle can also be applied within individuals; the same class of external stressor may provoke different reactions in the same individual over time. The activation of different

self-schemas across time helps explain this effect but as described above, the appraisal of the available self-concepts (e.g., how important they are) also exerts an effect on affective reactions to stressors (Showers et al., 1998). Therefore, understanding how the self affects psychopathology requires more than the specification of its structural features.

One of the functional features of schema-driven information processing is that prior beliefs and memory for regularities in the environment bias the allocation of attentional resources in new situations and affect what is remembered of an event. Allocation of attention to the self at the expense of awareness of the external world has been implicated in the pathogenesis of various disorders ranging from depression (Pyszczynski & Greenberg, 1987) to social phobia (Moscovitch, 2009). For example, Moscovitch (2009) proposes that the driving force behind many social phobia problems is not the fear of social situations per se but perceived flaws in the self such as skill deficits, character flaws, problems with physical appearance, or deficits in the ability to conceal anxious feelings. Hence, in this formulation, the phobic stimulus is not public speaking or going to a party; it is deficient features of the self that the individual attempts to conceal by deploying safety behaviors. An implication is that cognitive-behavioral treatments should involve exposure to the feared aspects of the self, as well as the feared external trigger situations (Moscovitch, 2009).

The general information-processing view assumes that the moment-to-moment experience of the self reflects the activation of particular self-schemata, termed the *working self-concept* (Markus & Nurius, 1986), *the working self* (Conway, 2005; Conway & Pleydell-Pearce, 2000), and *the experiencing self* (Dalgleish & Power, 2004). It is generally asserted that only one experiencing self is available to consciousness at a time (Power, 2007) and this constrains and "grounds" the available self-views (Conway, 2005; Conway & Pleydell-Pearce, 2000). However, it has been necessary to pose the additional concept of the *reflective self* or *observing self* (Power, 2007) to explain the phenomenological experience of being able to observe the operation of the experiencing self (e.g., noticing that one is having negative thoughts about the self) and the loss of sense of self that can occur in certain states (e.g., dissociative experiences during panic) or in disorders of self such as dissociative identity disorder and schizophrenia (Berrios & Markova, 2003; Power, 2007)[4]. For some patients, the enhancement of reflective self capacity may be necessary to enhance

4 Similar distinctions have been made in the ACT literature with the explicit separation of "self as content" from "self as context" and "self as process" (Hayes et al., 1999).

awareness of alternative funct
negative compartmentalizatic
(Dimaggio et al., 2010).

Specifically targeting the
has been addressed in the Cl
Fannon et al., 2009; Segal & M
sis from attempting to change
ucts (e.g., conditional beliefs
metacognitive strategies that p
ing processes (often with the
spective on the products of t
self-reflection may have deve
strategy for reducing awarene
al., 1998), it is likely that promc
some patients (e.g., those witl
such as paranoid or borderli
therapy conducted in an atmc
passion may be needed in orde
for developing greater self-acc

The fact that some peopl
consciously reflecting on aspe
and clinical practice has had t
unconscious processes on psy
the proposition that some forı
persecutory delusions are c
thoughts relating to low self-w

Conscious vs. Unconscious

The standard CBT approach
events such as negative autom
tional assumptions (A. T. Bec
CBT model emphasizes that th
emotional state and behavior e
their operation. Self-monitor
thought record and questionin
technique (J. S. Beck, 1995) are
mental events into conscious
cally evaluated. So, these men

awareness of alternative functional dimensions of the self or to reduce negative compartmentalization and an overly restricted self-structure (Dimaggio et al., 2010).

Specifically targeting the adaptive functioning of the reflective self has been addressed in the CBT literature only relatively recently (e.g., Fannon et al., 2009; Segal & Muran, 1993). This reflects a shift in emphasis from attempting to change the content of self-schemas and their products (e.g., conditional beliefs and negative automatic thoughts) onto metacognitive strategies that promote a greater awareness of one's thinking processes (often with the aim of encouraging a nonjudgmental perspective on the products of those processes). However, given that low self-reflection may have developed as a simplistic but partly effective strategy for reducing awareness of feared aspects of the self (Showers et al., 1998), it is likely that promoting self-awareness may be threatening for some patients (e.g., those with greatly elevated interpersonal sensitivity such as paranoid or borderline personality disorder patients). Hence, therapy conducted in an atmosphere of collaboration, safety, and compassion may be needed in order to promote the self-exploration necessary for developing greater self-acceptance (Gilbert, 2009).

The fact that some people experience considerable difficulty with consciously reflecting on aspects of the self has meant that CBT theory and clinical practice has had to account for the impact of conscious and unconscious processes on psychopathology. This issue is exemplified by the proposition that some forms of psychopathology such as mania and persecutory delusions are defenses against experiencing *conscious* thoughts relating to low self-worth (Bentall et al., 2001).

Conscious vs. Unconscious Self-Processes

The standard CBT approach includes promoting awareness of mental events such as negative automatic thoughts, unhelpful rules and conditional assumptions (A. T. Beck et al., 1979; J. S. Beck, 1995). The basic CBT model emphasizes that these stimuli will have a deleterious effect on emotional state and behavior even if the patient does not fully recognize their operation. Self-monitoring homework tasks such as keeping a thought record and questioning strategies such as the "downward arrow" technique (J. S. Beck, 1995) are deployed in order to explicitly bring these mental events into conscious awareness so that they can be systematically evaluated. So, these mental events are "preconscious" in that they

can operate outside of conscious awareness but are potentially available to introspection.

But some self-processes that are implicated in cognitive models of psychopathology are thought to operate entirely outside of awareness because their function is to prevent the confrontation of negative views of the self. This type of explanation is typically invoked where a distorted self-view develops to protect fragile self-esteem (Bentall et al., 2001). An example is the proposal that persecutory delusions form a defense against implicit (unconscious) low self-esteem that allows maintenance of positive explicit self-esteem by ascribing the source of negative experiences to the malign actions of others (Kinderman & Bentall, 1996).

The Self Esteem—Implicit Association Test (SE-IAT)

Empirical findings thus far collected from a self-esteem variant of the Implicit Association Test (SE-IAT; Greenwald, McGhee, & Schwartz, 1998) demonstrate that investigating unconscious processes can provide a more comprehensive understanding of the multiple processes that contribute to the experience of the self. The SE-IAT is a reaction time task that uses speed of responding as an index of the degree of association between concepts (Greenwald & Farnham, 2000). The stimuli are manipulated along self-other and positive-negative dimensions. Response times are faster when the target word and attribute matches the implicit association held by the participant (e.g., self-clever) than when they fail to match it (e.g., self-stupid). Unlike explicit assessments of self-concept, the SE-IAT has the advantage of being a more direct index of the spreading activation property of semantic networks where exemplars of a particular category that are more closely associated produce faster responding (e.g., "bird-canary" vs. "bird-ostrich") (Collins & Quillian, 1969). When assessing self-esteem, response latencies are interpreted as an index of how much the self-concept is associated with positive versus negative material. MacKinnon and others (MacKinnon, Newman-Taylor, & Stopa, 2011) used this measure to test the hypothesis that people with persecutory delusions would display the paradoxical combination of high explicit but low implicit self-esteem (Bentall et al., 2001). Patients with persecutory delusions actually showed similar levels of implicit self-esteem to healthy control subjects but lower explicit self-esteem. This directly opposes the "delusions as defense" model.

Although MacKinnon et al.'s
refinement of self-concept-based
methods used reflect progress tow
mata and related unconscious pr

Summary Comparison of Sc of the Self

The literature reviewed in this
the self in clinical CBT and relate
Unitary notions of a single "self"
of describing how self-referent in
and reorganized in the mind. Fu
self have been significantly ex
unconscious cognitive process
extensions of the original cogni
on the importance of metacogn
pathology (e.g., Janeck, Calama
1 presents a possible formulati
concepts from clinical models
social-cognitive research.

Three levels of conscious p
of self-constructs in CBT and r
depicted in Figure 1 with the
degree of permeability betwee
Although psychological therap
bally label core beliefs that un
and emotions, these beliefs a
learning experiences, most of
scious awareness. This differ
processing is particularly imp
tion of processes such as "sche
the unconscious or barely pre
tion is not fully reportable usir
utilize research methods that
Self-Esteem IAT and variants
answers to questionnaire.

ss but are potentially available

licated in cognitive models of
entirely outside of awareness
ıfrontation of negative views of
ally invoked where a distorted
steem (Bentall et al., 2001). An
elusions form a defense against
ıt allows maintenance of posi-
ıurce of negative experiences to
& Bentall, 1996).

tion Test (SE-IAT)

rom a self-esteem variant of the
nwald, McGhee, & Schwartz,
ɔnscious processes can provide
he multiple processes that con-
SE-IAT is a reaction time task
ex of the degree of association
, 2000). The stimuli are manip-
ve dimensions. Response times
ɔute matches the implicit asso-
clever) than when they fail to
ıssessments of self-concept, the
e direct index of the spreading
where exemplars of a particular
produce faster responding (e.g.,
& Quillian, 1969). When assess-
ıterpreted as an index of how
positive versus negative mate-
on, Newman-Taylor, & Stopa,
hesis that people with persecu-
cal combination of high explicit
., 2001). Patients with persecu-
evels of implicit self-esteem to
licit self-esteem. This directly

Although MacKinnon et al.'s (2010) results demonstrate a need for refinement of self-concept-based theories of persecutory delusions, the methods used reflect progress toward more convincing measures of schemata and related unconscious processes.

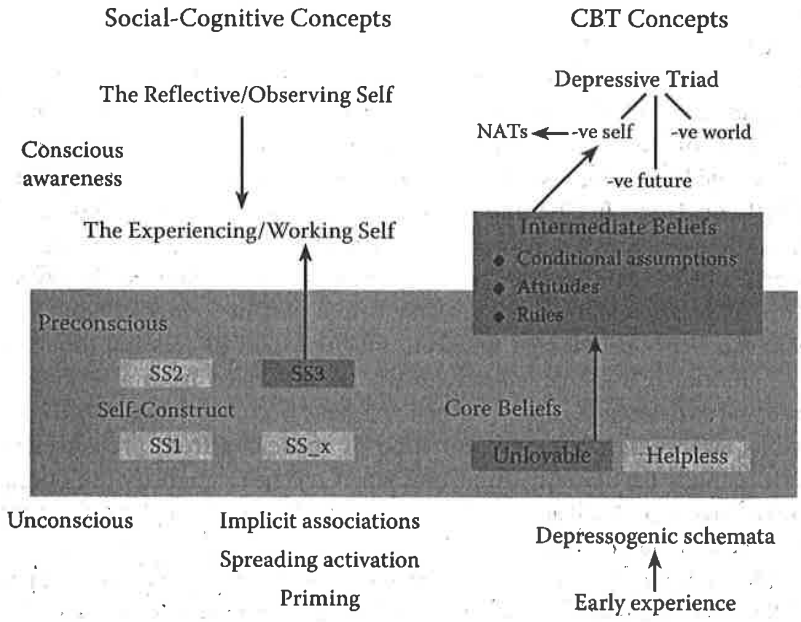
Summary Comparison of Social-Cognitive & CBT Concepts of the Self

The literature reviewed in this chapter suggests that the conception of the self in clinical CBT and related areas of empirical research is evolving. Unitary notions of a single "self" have been replaced by multifaceted ways of describing how self-referent information is organized, stored, accessed, and reorganized in the mind. Furthermore, content-based models of the self have been significantly extended by incorporating conscious and unconscious cognitive processes that affect self-construction. These extensions of the original cognitive models reflect the greater emphasis on the importance of metacognitive processes in understanding psychopathology (e.g., Janeck, Calamari, Riemann, & Heffelfinger, 2003). Figure 1 presents a possible formulation of the ways that multiple self-referent concepts from clinical models of CBT correspond to those derived from social-cognitive research.

Three levels of conscious processing are invoked in the descriptions of self-constructs in CBT and related social-cognitive research. These are depicted in Figure 1 with the preconscious level shaded to highlight a degree of permeability between fully conscious and unconscious levels. Although psychological therapy may help a patient to identify and verbally label core beliefs that underpin much of their distressing thoughts and emotions, these beliefs are abstractions derived from a myriad of learning experiences, most of which will have occurred without full conscious awareness. This differentiation of conscious from preconscious processing is particularly important when trying to quantify the operation of processes such as "schema activation" (Segal, 1988) that fall within the unconscious or barely preconscious realm. Because schema activation is not fully reportable using verbal means, it has become necessary to utilize research methods that rely on behavioral response times (e.g., the Self-Esteem IAT and variants of the Emotional Stroop tasks) rather than answers to questionnaire.

By definition, self-referent schemata are summaries of experience, not a veridical record of all of the details of all of the specific learning events that shaped any given schema. Hence, Figure 1 places the examples of depressogenic schemata and early experiences within the class of unconscious phenomena. The related constructs and processes in the social-cognitive literature include notions of implicit association between stimuli (that influence response times independent of explicitly expressed beliefs and attitudes), spreading activation between conceptually or semantically related information, and priming effects. All of these phenomena exert a measurable effect on behavioral functioning but are independent of the types of controlled information processing usually assessed with questionnaires or other explicit assessment procedures. Furthermore, we argue that this underlying information is not accessible in any complete way but that it is possible to become aware of the derivatives of this information (e.g., core beliefs, self schemas).

Figure 1: A comparison of social-cognitive and CBT constructs relating to the self



Notes: SS= self-schema; -ve = negative; NAT = Negative automatic thoughts

The constructs specifi
potentially responsive to in
conscious awareness for m
structs such as the self-co
(Dalgleish & Power, 2004; S
Figure 1 are presented to c
mentalized into subunits tl
ideas correspond to the nc
down to fundamental conce
1995). This reduction of co
versally endorsed. For exa
lates up to 13 maladaptive
others that include themes
entitlement (Young et al., 2
CBT models propose tha
thoughts, feelings, and bel
awareness. One implicatio
involves helping the patien
the self, conditional assur
rules that they impose on t
tiny is depicted in Figure 1
boundary of preconscious a

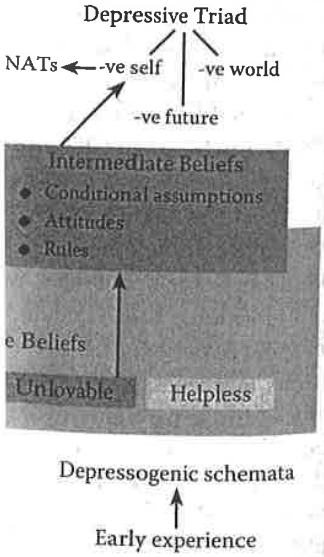
The final point of comp
processes that are most re
CBT, emphasis is placed o
triad on emotional state an
(negative views of the self) a
(NATs) that arise briefly in
an effect on emotions and
are less clear in specifying
"read" or appraised by the
NATs are judged to be "tru
is not a focus of much ther
problem of infinite regress
a homunculus or "true sel
ings but it leaves an import
cognitive conceptions of th

5 Young et al. (2003) also diffe
early in life from conditional mala
modified over time.

are summaries of experience,
s of all of the specific learning
ce, Figure 1 places the examples
xperiences within the class of
onstructs and processes in the
of implicit association between
ependent of explicitly expressed
tion between conceptually or
ming effects. All of these phe-
vioral functioning but are inde-
formation processing usually
xplicit assessment procedures.
ıg information is not accessible
to become aware of the deriva-
self schemas).

and CBT constructs relating

CBT Concepts



The constructs specified at the preconscious level in Figure 1 are potentially responsive to introspection but may operate largely without conscious awareness for most people. Social-cognitive models use constructs such as the self-concept or self-schema to convey these ideas (Dalgleish & Power, 2004; Showers, 1992). The subfields of SS1, SS2 etc. in Figure 1 are presented to convey that the self-concept may be compartmentalized into subunits that are differentially activated. In CBT these ideas correspond to the notion of Core Beliefs, which may be distilled down to fundamental concerns about lovability or helplessness (J. S. Beck, 1995). This reduction of core beliefs to only two main themes is not universally endorsed. For example, Young's Schema therapy model postulates up to 13 maladaptive unconditional schemas about the self and others that include themes of fear of abandonment, defectiveness, and entitlement (Young et al., 2003)[5]. However, at the functional level, most CBT models propose that self-schemas can exert an influence on thoughts, feelings, and behavior even if they are outside of conscious awareness. One implication of this is that CBT treatment frequently involves helping the patient become aware of negative attitudes toward the self, conditional assumptions about self-worth, and dysfunctional rules that they impose on themselves. This potential for conscious scrutiny is depicted in Figure 1 by situating "Intermediate Beliefs" across the boundary of preconscious and conscious processes.

The final point of comparison in Figure 1 addresses the concepts and processes that are most readily subject to full conscious awareness. In CBT, emphasis is placed on the impact of the cognitive (or depressive) triad on emotional state and behavior. The self-referent subtype of these (negative views of the self) are expressed in Negative Automatic Thoughts (NATs) that arise briefly in consciousness and dissipate quickly but exert an effect on emotions and behavior. Clinical models of traditional CBT are less clear in specifying the mechanisms by which these thoughts are "read" or appraised by the individual. There is a tacit assumption that NATs are judged to be "true," but the observer that makes this judgment is not a focus of much therapeutic attention. This partly avoids the messy problem of infinite regress to unsatisfactory structural concepts such as a homunculus or "true self "that observes and judges thoughts and feelings but it leaves an important gap in the clinical models. However, social-cognitive conceptions of the self have achieved some clarity by specifying

5 Young et al. (2003) also differentiate unconditional maladaptive schemas that are learned early in life from conditional maladaptive schemas that are shaped by later experience and are modified over time.

modes of processing as an alternative to a unitary and static observer of the self. Figure 1 shows that there is a distinction drawn between the working or experiencing self and the observing or reflective self. Ultimately, these are still descriptive concepts that serve to more accurately portray phenomenological experience but do not explain how these experiences arise. But they do provide a more nuanced way of separating the experience of a thought (e.g., "I am a failure") from the capacity to observe, judge, or appraise that thought. The development of "third wave" mindfulness-based variants of CBT reflects a move towards incorporating these distinctions more completely into therapy (e.g., Hayes, Strosahl, & Wilson, 1999).

Conclusion

Despite a pedigree stretching back to early Greek philosophy, the concept of self infiltrated psychiatric and psychological writing only in the nineteenth century with the conceptualization of "disorders of the self" such as schizophrenia (Berrios & Markova, 2003). CBT as a theoretical and therapeutic endeavor has embraced the proposition that self-belief, self-criticism, self-esteem, and related concepts are central to the understanding of a range of psychopathological conditions. However, there is considerable variation in the way that the self is invoked both as a basis for suffering and as a focus of change efforts such as "schema modification." This partly reflects the problem of treating the self as a unitary concept. As Berrios and Markova note: *"The self was never meant to be a solid object like a stone, a horse, or a weed, nor even a concept to be considered as semantically tantamount to changes in blood flow or test scores" (p. 10).* The evidence presented above indicates that progress is being made toward developing a more multifaceted and dynamic view of the self. Furthermore, as regards the relationship between traditional CBT and the alternative contextual behavioral approach to therapy represented by ACT, as indicated, in some respects CBT theory and research on the self are showing a degree of convergence with current contextual behavioral scientific understandings. This convergence will hopefully benefit both traditions, especially with respect to the conceptualization of the self and the use of this concept in the psychological treatments of the future.