Rev.Phil.Psych. (2017) 8:391–410
DOI 10.1007/s13164-016-0308-1

Know Thyself? Questioning the Theoretical Foundations of Cognitive Behavioral Therapy

Garson Leder[1]

Published online: 24 March 2016

Abstract Cognitive Behavioral Therapy (CBT) has become the dominant form of psychotherapy in North America. The CBT model is theoretically based on the idea that all external and internal stimuli are filtered through meaning-making, consciously accessible cognitive schemas. The goal of CBT is to identify dysfunctional or maladaptive thoughts and beliefs, and replace them with more adaptive cognitive interpretations. While CBT is clearly effective as a treatment, there is good reason to be skeptical that its efficacy is due to the causal mechanisms posited by the CBT model. This paper will argue that the specific cognitive schemas posited by the CBT model likely do not play a direct role in the development or treatment of psychological illness. Cognitive schemas, as identified in CBT interventions, are likely to be the result of patient confabulation and epistemically under-supported practitioner-based identification. CBT interventions appear to impose coherence on patients' psychological states, rather than identifying and modifying preexistent causally efficacious core beliefs.

1 Introduction

In the latter half of the 20th century cognitive behavioral therapy (CBT) replaced psychodynamic and behavioral therapies as the dominant form of psychotherapy in North America (Westbrook et al. 2011; Norcross and Karpiak 2012). This was largely due to CBT's perceived superior testability and efficacy in comparison to other forms of treatment. Roughly, CBT is a combination of behavioral therapy and cognitive therapy that aims at identifying and replacing maladaptive or dysfunctional thoughts and beliefs and replacing them with more adaptive cognitive interpretations. The CBT model of psychological functioning posits an interconnected triad of thoughts, behavior, and emotions, with thoughts playing the primary role in the development and

✉ Garson Leder
leder.garson@gmail.com

[1] Department of Philosophy, University of British Columbia, Vancouver, BC, Canada

treatment of dysfunctional psychological states (Beck et al. 1976; Beck 2011). Maladaptive or dysfunctional emotional or affective responses are modified by altering thoughts (either directly, or indirectly through behavioral interventions, or both). This model's theory is based on the idea that all external and internal stimuli are filtered through meaning-making, consciously accessible cognitive schemas, or core beliefs, that can represent the world in either adaptive or maladaptive ways.[1] The goal of CBT is to help patients to identify, challenge, and replace the specific dysfunctional or maladaptive beliefs that are postulated to be the primary factor in their psychological disorder.

While CBT is clearly effective as a treatment, there is good reason to be skeptical that the efficacy is due to the causal mechanisms posited by the CBT model. This paper will argue that the specific cognitive schemas posited by the CBT model likely do not play a direct role in the development or treatment of psychological illness. Cognitive schemas, as identified in CBT interventions, are likely the result of patient confabulation and epistemically under-supported practitioner-based identification. CBT interventions appear to impose coherence on patients' psychological states, rather than actually identifying and modifying existent causally efficacious core beliefs.

This discussion will first outline Beck's CBT model of the affective disorders, then highlight the CBT model's reliance upon introspective and retrospective belief reports in identifying and challenging maladaptive cognitions. The discussion will then focus on problems with the CBT model's reliance upon direct introspective access to patients' cognitive processes, and conclude with suggestions for the construction of a more plausible cognitive theory.

2 The CBT Model

At its most basic, the CBT model, first posited by Albert Ellis (1962) and Aaron Beck (1967), is concerned with the relation between cognitions, emotions, and behavior. Cognitions (thoughts, beliefs, and assumptions) are posited as playing the primary role in the formation and treatment of dysfunctional or maladaptive psychological states (Clark and Beck 1999). How individuals interpret the world is supposed to influence, and be influenced by, their behavior and emotions. Maladaptive emotions (such as depressive states) are conceived of as subjective states caused by overly rigid and/or inaccurate cognitive appraisals or evaluations of internal or external stimuli (Clark and Beck 1999). How a stimulus is interpreted by the informational processing system determines the valence, persistence, malleability, and intensity of emotional responses. The CBT model also maintains that behavior influences thoughts, and therefore also alters emotions. Changing maladaptive behavioral patterns is taken as a tool to indirectly change unhealthy cognitive patterns by way of challenging unhealthy cognitions (e.g., safely exposing a patient to an irrationally fear-inducing stimulus is used to challenge and alter negative thoughts and thus extinguish the negative emotional

[1] There is a good deal of ambiguity in the use of the word "schema". Many authors, such as Beck et al. (1976), Beck (2011) and Clark (2004) us the terms schemas and core beliefs interchangeably. Others, such as Young et al. (2003) define schemas as any semantic cognitive filter. I will be following the latter usage.

response). CBT interventions aim to provide patients with less dysfunctional or distorted, and more adaptive and realistic, meaning-making interpretations of the world.

The most influential, studied, and applied cognitive therapy is Aaron Beck's cognitive behavioral therapy (Beck et al. 1976, 1979).[2] The Beckian cognitive model (henceforth, CBT) posits three levels of cognitions that are supposed to filter all experience: automatic thoughts, intermediate beliefs/assumptions, and core beliefs/schemas (Beck 2011; Leahy 1996; Clark and Beck 1999; Westbrook et al. 2011). The most basic level of cognitive processing, core beliefs/schemas, are supposed to "enable individuals to make sense of their environment by breaking it down and organizing it into psychologically relevant facets...[and] direct all cognitive activity whether it be ruminations and automatic thoughts or cognitive processing of external events" (Clark and Beck 1999, p. 52). If things are running well, one's schemas represent the world in ways that do not lead to psychological distress or maladaptive thoughts and beliefs. Things start to go poorly when one's cognitive processes represent the world in overly rigid, negative, or polarized ways.

Automatic thoughts are supposed to sit at the most salient end of the cognitive hierarchy. Automatic thoughts are defined as easily consciously accessible, context-specific beliefs about, attitudes towards, or semantic interpretations of external and internal stimuli. These are surface-level thoughts that superficially explain individuals' thoughts and behaviors. For example, a patient may report having the negative automatic thoughts "I will be picked last" or "I will embarrass myself if I try" when deciding not to join in a group sports activity. Importantly, these thoughts are usually not explicitly held or consciously entertained, but are supposed to be easily identifiable by introspection or elicited by practitioner-based questioning.

Intermediate beliefs are the middle level of the CBT cognitive model and the immediate platform from which automatic thoughts are formed (Beck 2011). They are rules or patterns of association used to interpret and evaluate experiences. For example, the automatic thought "I am being boring" that may ground a patient's feeling of unease during a social situation may be grounded in the intermediate belief that "if I talk too much, then people will think I'm boring". These beliefs often take the form of conditional statements, such as "If I please my partner, then he/she will treat me well" or "If I am criticized, then it means that I have failed" (Clark and Beck 1999). Maladaptive intermediate beliefs are characterized by being overly rigid, based in thought errors such as catastrophizing or all-or-nothing thinking, and containing distorted world views. The CBT model posits that these beliefs are more difficult to identify than automatic thoughts and must normally be inferred by patient and practitioner from the patterns and content of automatic thoughts.

Core beliefs are the most basic and fundamental beliefs about oneself and the world and are the basis of all other consciously accessible cognitions. The beliefs are highly generalized, absolute, and difficult for patients to consciously access. Practitioners are trained to identify them by way of recognizing consistent patterns in patient belief and thought reports. Negative core beliefs may take the form of statements such as "I am a

[2] Since Beck et al's. (1976) formulation, the umbrella term 'CBT' has grown to include a number of closely related therapies. A number of recent 'third wave' cognitive therapies (e.g., mindfulness-based cognitive therapy [MBCT], dialectical behavioral therapy [DBT], and meta-cognitive therapy [MCT]), are often categorized as 'cognitive behavioral therapies', despite differing in theory and practice. This paper will focus on the dominant Beckian model of CBT.

failure" and "I am unlovable", while positive core beliefs are expressed by thoughts such as "I am likable" and "I am worthwhile" (Clark and Beck 1999). The CBT model postulates that all dysfunctional or maladaptive automatic and intermediate thoughts and beliefs are the result of dysfunctional or maladaptive core beliefs (Clark and Beck 1999; Beck 2011).[3]

Patients' intermediate and core beliefs are identified in CBT interventions via Socratic questioning and downward arrow interviewing (Neenan and Dryden 2005; Beck 2011). Socratic questioning (also called 'guided discovery') consists of persistent questioning of the patient's reasons and justifications for having specific automatic thoughts. The aim here is to aid the patient in searching for the (possibly distorted or unhealthy) thoughts or beliefs that explain her thoughts, feelings, and behavior (Neenan and Dryden 2005). For instance, patients may be asked to identify patterns in their behavior, explain why they think they have certain thoughts, and explain what specific thoughts mean to them. Similarly, in downward arrow questioning, the practitioner attempts to identify core beliefs by asking the patient to identify what their previously identified automatic or intermediate beliefs *mean* to them (Beck 2011). This process is repeated until the patient arrives at the lowest level of abstraction (core beliefs). Consider the following example of the downward arrow technique taken from a CBT training handbook (Harwood et al. 2009):

- Situation: At home on a Saturday afternoon.
- Emotions: Depressed (80 %), anxious (60 %).
- Automatic thought: "I should have a date on Saturday night."
- Therapist: What does it mean if you do not have a date on Saturday night?
- Patient: It means that I'll be home by myself on Saturday night.
- Therapist: What does being home alone on a Saturday night mean?
- Patient: It means that I am not out having fun like everybody else.
- Therapist: And what does that mean to you?
- Patient: That I am a loser, nobody loves me, and I'll always be alone.

In this example, the core belief "I'll always be alone" was elicited by the patient attempting to make sense of, or find the deeper meaning in, her higher level thoughts. Again, CBT theory posits that thoughts and beliefs are based on more basic thoughts and beliefs, with core beliefs filtering all semantic interpretations of the world. Patients and practitioners attempt to make sense of maladaptive automatic thoughts by locating more general thoughts and beliefs that would explain why the automatic thoughts are held in the first place. From here, the thoughts can be challenged and modified. This process usually involves the patient keeping a thought record to identify her automatic thoughts, then challenging and weighing the accuracy of the thoughts both in session and through homework. Behavioral interventions, such as increasing pleasurable activities and benign exposure and habituation to perceived harmful or fear-inducing activities, may also be used. In both the approaches the aim is to change how the patient

[3] The Beckian model has been modified to include non-consciously accessible cognitive processes (Beck 1996; Clark and Beck 1999). While cognitive therapy still focuses on identifying and challenging beliefs and and thoughts, the CBT model now postulates that clusters of interrelated schemas called 'modes' play a significant role in cognitive functioning.

thinks about and interprets the world.[4] The patient is challenged (both in person and through homework assignments) to question and find reasons to undermine the distorted or dysfunctional core and intermediate beliefs and replace them with more accurate and adaptive beliefs.

3 Criticisms of the CBT Model

A number of randomized controlled trials, meta-analyses, and meta-analyses of meta-analyses have shown CBT to be an effective therapeutic treatment for a wide range of psychological problems (Leichsenring et al. 2006; Clark and Beck 1999; Butler et al. 2006). Despite its efficacy, CBT is not without critics. Objections to CBT theory come in four main camps: (1) criticisms of the primacy given to cognitions over other psychological processes (such as emotion or non-consciously accessible drives) (e.g., Teasdale and Barnard 1993; Teasdale 1997), (2) criticisms of cognitive theory being overly general or metaphorical (e.g., Coyne and Gotlib 1983; Brewin 1996), (3) criticisms of CBT's (and every other theory's) lack of causally efficacious specific effects (e.g., Wampold and Imel 2015), and (4) criticisms of CBT's assumptions about the representational nature of cognition (e.g., McEachrane 2009; Gipps 2013). What is common to these objections is the idea that if CBT works, it is not because of the reasons given by the theory grounding the therapy.

Supporters of cognitive theories of psychopathology have countered group (1) type criticisms by appealing to the substantial literature on the ubiquity of, and central role for, maladaptive cognitions in cases of psychological distress (Clark and Beck 1999; Beck 2004). The first versions of cognitive therapy may have been vulnerable to group (2) type critiques. Early CBT theorists such as Beck and Ellis were mainly concerned with establishing a dominant role for thought in depression (in contrast to psychodynamic and behavioral learning models), rather than focusing on the specifics of how this might work. Later versions of the cognitive model have addressed this problem by becoming far more specific as to what meaning-making structures are, how they are structured in the informational processing system, and the roles they play in psychological disturbances (Clark and Beck 1999; Beck 2005). The standard CBT response to group (3) objections is to either challenge the accuracy of meta-analyses (Crits-Christoph 1997; Butler et al. 2006) that purport to show an equivalence of effectiveness across different therapies or to claim that if other forms of therapy work it is only because they are changing cognitions- in effect, other therapies are actually doing a form of CBT (Alford and Beck 1998). According to the latter view, all therapies work by challenging specific cognitions- either directly (as in CBT interventions) or indirectly (as in other non-cognitive, but efficacious, treatments). There is some plausibility to this response. But, importantly, this reply rests on the assumption that there are

[4] While CBT adopts many therapeutic methods from behavioral therapy, CBT theorists and behaviorists give differing explanations for the therapeutic change engendered by the use of behavioral therapeutic techniques. Traditional behaviorist theories focus on the alteration of conditioned, non-consciously accessible behavioral rules (rather than consciously accessible inner states) to explain psychological change (Skinner 1977). According to CBT theorists, behavioral methods are successful only insofar as they help modify patients' maladaptive thoughts and beliefs (Beck 1979).

specific beliefs and automatic thoughts that play a primary role in the production of psychopathological states.

The group (4) objections to the CBT model criticize its (and most of cognitive science's) assumptions about the representational nature of belief and thought. Representational theories of cognition take beliefs and thoughts to be causally efficacious mental representations of facts, states of affairs, or propositions. Critics of representationalism have argued that CBT confuses individuals' thought reports (which are represented as having imagistic or linguistic content), with their thoughts (which needn't have any distinct representational content at all) (McEachrane 2009). Rather than respond directly to type (4) objections, CBT theorists take representationalism to be a foundational assumption of cognitive science and clinical cognitive theory (Clark and Beck 1999). This paper will share CBT's assumptions about the representational theory of mind.

This paper raises a fifth set of concerns about the cognitive model. The CBT model is based on the assumption that our cognitions have a coherent, logical, and consciously accessible hierarchical structure; the content of all cognitions is based on more general cognitive content. If patients can introspectively identify their automatic thoughts, then the cognitive model predicts that they should be able to identify the more general schemas that ground these thoughts. However, there are serious problems with both the tenability of CBT's hierarchical model of cognition and its assumptions about the accuracy of the cognitions being identified and challenged in therapy.

4 Thoughts, Beliefs, and Confabulation

4.1 Identifying Cognitions

There is an oddity to the CBT model. According to CBT, our meaning-making information processing systems are posited to be actively creating our reality, *unless we are introspecting*. CBT is based the idea that all "stimuli that impinge on the organism" are filtered through cognitive schemas that structure and give meaning to experience (Beck 1967). Our informational processing system is supposed to "actively participate in the construction of reality", and this construction "is not simply an act of representing, copying, or "coding" fixed objects but rather is a process that involves some degree of creativity" (Clark and Beck 1999). Yet we are supposed to be accurate introspectors of our past thoughts and beliefs. CBT theory adopts a constructivist view of cognition, but a more-or-less realist view of introspection. Our cognitive processes are identified as "meaning-making structures" that can either represent the world in maladaptive or adaptive ways, but at the same time we are supposed to be able to accurately identify the cognitions underlying our behavior rather than simply "making sense" of our emotions and behaviors. This is likely not the case.

CBT assumes that patients have, or can be trained to have, direct and accurate access to the content of their own cognitions (Beck and Dozois 2011). CBT practitioners are supposed to aid patients in identifying their own thoughts and beliefs by engaging in directed Socratic questioning aimed at eliciting deeper cognitive schemas. Practitioners guide patients' introspection of their core schemas by identifying common themes in the patients' automatic and intermediate thoughts and directing patients to search for

the underlying structure in their thoughts and beliefs. For example J. Beck states, "asking what a thought means to the patient often elicits an intermediate belief; asking what it means about the patient usually uncovers the core belief" (2011). This identification of cognitions often requires work on the part of both the patient and therapist. A critical part of cognitive therapy is to first *train* patients to recognize, attend to, and record this inner speech or automatic thoughts (Beck 1976, 1979; Beck 2011). Patients often claim to be unaware of having an "internal communication system" or are unused to attending to the content of these thoughts or images (ibid). The training process involves explaining the cognitive model to patients and articulating the logical connection between core beliefs and automatic thoughts and the relation between thoughts and emotions. Beck (1976) states:

> The training in the observation and recording of cognitions makes the patient aware of the occurrence of images and self-verbalizations ("stream of thought"). The therapist trains the patient to identify distorted and dysfunctional cognitions. The patient may need to learn to discriminate between his own thoughts and the actual events. He will also need to understand the relationship between his cognitions, his affects, his behaviors, and environmental events. (p. 146)

Similarly, Beck and Alford (2009) state:

> At the beginning of therapy the patient is generally aware only of the following sequence: event or stimulus— > affect. He must be trained to fill in the link between the stimulus and the affect: stimulus— > cognition— > affect. (p. 310)

There are reasons to be skeptical about the success of this training. Automatic thoughts are identified by simply asking patients to introspect what they were thinking at any given moment (e.g., what they were thinking while feeling sad staying home on a Saturday night). Most automatic thoughts are not explicitly entertained (insofar as they are not salient parts of a patient's inner monologue) and require introspection and practitioner-based prodding to identify. For example, the J. Beck CBT manual states: "Automatic thoughts are usually quite brief, and the patient is often more aware of the *emotion* she feels as a result of the thought than of the thought itself. Sitting in session, for example, a patient may be somewhat aware of feeling anxious, sad, irritated, or embarrassed but unaware of her automatic thoughts until her therapist questions her" (2011, original italics). In cases where the patient is unable to identify any thoughts or confuses thoughts with feelings, practitioners will use questions such as "what would you guess was running through your mind at that time?", "what did this situation mean to you (or about you)?", or even "might you have been thinking __ or __?" (Beck 2011). According to the cognitive model, "the emotion the patient feels is logically connected to the content of the automatic thought" and it is the job of the practitioner to help the patient identify this logical connection. However, a serious problem with this process is that what a thought means to a patient is highly dependent upon the theory of cognition and psychological functioning being deployed by both patient and practitioner. This explicit search for meaning is, according to the CBT model, itself based on meaning-making schemas that need not accurately represent anything (Clark and Beck 1999).

It makes sense that automatic thoughts are based on more basic core and intermediate beliefs only if one adopts a theory of psychological functioning that posits a nested hierarchy of consciously accessible and causally efficacious thoughts. Importantly, many other explanations can also make sense of the same stimuli without reference to a hierarchy of cognitions. For example, most modern psychodynamic theories posit conflicts between subconscious and conscious feelings and drives (the sex drive, self-esteem, etc.) as the basis for psychological distress (Gabbard 2000; Eagle and Wolitzky 1992). For patients who adopt a theory of psychological functioning steeped in the Freudian-inspired psychodynamic model, thoughts such as "I should have a date on Saturday night" may mean that the patient has repressed subconscious-based anger towards the perceived loss of parental affection (or any number of other possible interpretations). In the middle of the 20th century Freudian-inspired drive based theories of psychological processing and object-relation theories (which based mental illness in the feeling of real or perceived loss of object(s) in early childhood) grounded how most of psychology—and much of the educated populace—made sense of their mental lives. The conflict between the Id, Ego, and Super Ego made sense to many people for a long time before falling out of fashion in favor of cognitive and behavioral theories. Just as we should be aware of the influence of theory upon a patient's Freudian interpretation of her lack of Saturday evening plans, we need to also be cautious in accepting at face value a CBT model inspired interpretation of the meaning of a patient's thoughts.

It is also important to be cautious in accepting appeals to common sense. CBT's theory of psychological functioning and assumed nested hierarchy of consciously accessible and causally efficacious thoughts is explicitly intended to be built upon a common sense notion of how the mind works (Beck 1976; Ellis 1994). This is claimed to be another mark of its superiority to the supposedly more unintuitive theoretical assumptions of psychoanalytic and behavioral therapies. The problem, of course, is that common sense and armchair models of cognition do not necessarily map on to how the mind actually works; common sense and folk psychology are bound by culture and context. It may be common sense to some Freudian-inspired individuals to assume that most desires, including infantile and childhood desires for comfort and attention, are sexual in nature (Freud, 1905/2000). It is also common-sensical to adherents of the medical model of mental illness (including many psychiatrists) that negative or maladaptive automatic thoughts are the product of neurochemical imbalances in the brain, rather than of dysfunctional beliefs (Lebowitz 2014; Pescosolido et al. 2010). Thoughts and beliefs, according to this model, are symptoms of neurochemical problems rather than the primary problem itself. The commonness, or intuitiveness, of the common sense assumption that consciously accessible, logically structured beliefs play a primary role in the development and treatment of the affective disorders is dependent upon the acceptance (be it implicit or explicit) of the cognitive model of cognition.

The theory-ladenness of the CBT processes is significant. In order for CBT to work as theorized, patients and practitioners must be able to accurately identify maladaptive automatic thoughts in order to then challenge and modify them (or the core and intermediate beliefs that ground them). This requires that patients identify the *actual* thoughts that explain and cause their feelings and behaviors, not just identify thoughts that *describe* or *make sense of* these states. But the latter is what the cognitive model

would predict. CBT is based on the assumption that our informational processing systems actively create and structure our subjective realities. Whether these representations are accurate or inaccurate should not be important; all that matters for healthy psychological functioning is that our schemas represent the world in adaptive rather than maladaptive ways.

The theory-ladenness of CBT patients' explanation for the meaning of their thoughts and emotions, and their subsequent belief identifications, does not entail that the theory itself is wrong. There are independent reasons to think this. Notably, there are serious flaws with the CBT model's assumptions about the logical connection between automatic thoughts and emotions, and less consciously accessible intermediate and core beliefs, as well as problems with CBT's assumptions about patients' introspective access to their own beliefs. This section will address these problems in order.

4.2 Problems with the Cognitive Model

CBT posits that automatic thoughts have logical connection to core beliefs; if you think something, you think it for an identifiable and (at least internally) coherent reason. But this is often not the case. Contrary to the CBT model, there is strong evidence that automatic thoughts are often not logically connected to, or derived from, stable and consciously accessible core and intermediate beliefs. The problem here is that individuals are often poor introspectors and retrospectors of the causes of their own cognitive processes; while we often know what we are feeling, we do not often know why we are feeling it. Individuals' self-reports of the causes of their cognitive states and behavior, rather than being based on direct introspective awareness, are often confabulations based on post-hoc rationalizations or a priori causal theories (Nisbett and Wilson 1977; Wilson 2002; Haidt 2006). Individuals will often make sense of their emotions and behaviors regardless of whether the explanation accurately reflects the actual causal story.

Individuals have restricted introspective access to cognitive processes causally responsible for much of their behavior and thoughts (Wilson 2002; Kahneman 2011). Confabulation, or spontaneous unintentionally fabricated or distorted memories, occurs when individuals are put in a position to explain these thoughts and behaviors. One of the earliest and most striking examples of confabulation comes from Nisbett and Wilson's (1977) study of introspective access to cognitive processes. In one experiment, subjects were instructed to pick a preferred article of clothing and explain their choice. All things being equal, individuals tend to have a positive bias towards the rightmost object in a series. Consistent with this, 80 % of the subjects in the Nisbett experiment preferred an article of clothing on the rightmost side of a display of similar (or even indistinguishable) garments. When questioned about their choice, subjects explained their purchase by appealing to a preference for the color, style, material, or texture (even when the garments were indistinguishable), rather than identifying the actual causal reason for their choice (the garment's locations). Wilson and Nisbett concluded that when questioned on their cognitive processes, people base their responses on "a priori, implicit causal theories, or judgments about the extent to which a particular stimulus is a plausible cause of a given response" (Nisbett and Wilson 1977). Similar results were found in Johansson et al. (2005) facial attraction experiment. Subjects were shown photographs of two persons and asked to choose the most

attractive one. After a choice had been made, the experimenters surreptitiously switched the photographs (displaying the one not chosen) and asked the subjects to explain why they chose the photograph in question (even though they in fact did not choose it). In the majority of cases the switch was unnoticed and subjects gave confabulated and specific reasons for their (non)-choice (e.g., a preference for blondes, green eyes, older persons, etc.). Rather than directly introspecting the causal processes of their thoughts, the subjects confabulated and created explanations that made sense given the stimuli and their responses.

One should also be skeptical of CBT's assumption that consciously accessible core beliefs are the logical source of one's stream of conscious inner experience or automatic thoughts. Implicit cognitive biases or non-conscious heuristics (e.g., biases about the attributes of certain ethnicities and gender), rather than core beliefs, often influence our behavior, thoughts, and judgments (Harden and Banaji 2013). For example, white North Americans show an implicit preference for white names (71 %) and faces (88 %) over those of blacks (as evidenced by delayed reaction times in associating black faces and traditionally black-sounding names with words such as "good" and "joy"), as opposed to self-reported explicit biases of 27 % and 36 %, respectively (Nosek et al. 2002). The divergence between explicit and implicit biases can also be found in the *pro-white* implicit biases of American hispanics, asians, and native americans (Nosek et al. 2007). Similar patterns of dissociation between implicit and explicit biases can be found in biases against the elderly, homosexuals, and Muslims (Ibid). While implicit biases, by themselves, do not determine behavior, they have been found to strongly correlate with political choices (Knowles et al. 2010), hiring decisions (Bertrand and Mullainathan 2003), and medical treatment (Harden and Banaji 2013). For example, in a study by Steinpreis et al. (1999), both male and female psychology professors rated fictitious C.V.s of job candidates with male names to be superior, and the male candidates more hireable, in comparison to similar C.V.s containing female names. Given the agreement between professors of both sexes, it is most likely the case that the differences in appraisals is due to implicit (rather than explicit) biases about the academic competence of men compared to women. The significance of these divergences between explicit and implicit biases is that, at least in some cases, automatic thoughts (e.g., "this person is most qualified for the job"), are not based solely on some logically connected consciously accessible belief (e.g., "if one is most qualified, then one should be hired" or "egalitarianism is right"). Implicit biases, or beliefs, about gender or ethnicity that contradict explicitly held (or at least articulated) beliefs can, and do, influence thoughts in the absence of a consciously accessible logical connection.

Environmental factors also play a significant role in influencing cognitions outside of conscious awareness. Priming effects (specific behavioral changes after being exposed to a stimulus) affect both behavior and cognition. For example, exposure to pleasant environments (such as pleasant smells) significantly increases helping behavior, while unpleasant environments (such as unpleasant smells or messy rooms) decrease such behavior (Isen and Levin 1972). Being primed by negative or positively valenced words appears to make individuals more or less likely to act impolitely (Bargh et al. 1996). Neat or messy work environments appear to prime individuals' moral judgments (Schnall et al. 2008), their purchasing habits (Liu et al. 2012), and even their opinion of a therapist's competence (Nasar and Devlin 2011). While behavioral priming studies mainly focus on responses to environmental factors, it is very unlikely that the

true causal processes behind the behavioral responses are noticed by individuals. One of the most interesting aspects behind the priming studies is that the subjects are normally unaware of the priming's effect on their behavior. For example, it is highly unlikely that people would explain their helping behavior as being caused by morally arbitrary factors such as standing in front of a bakery rather than by appealing to some explanation based on their character and personality. Instead of having direct introspective access to the cognitive processes that ground their behaviors (such as the positioning of a garment or cleanliness of a room), individuals often appear to be in the position of cognitive interpreters of their own past feelings and behaviors. By asking for the meaning and cause of thoughts, CBT practitioners are asking patients to identify the causal processes responsible for their cognitions; this is something people are famously bad at.

By employing a post-hoc introspective lens on their cognitive processes, patients are likely often identifying thoughts that *describe* how they feel rather than uncovering their actual thought processes. This is an important difference. The CBT cognitive model maintains that CBT interventions work by accurately identifying and challenging dysfunctional or maladaptive thoughts, not merely by helping patients find ways to conceptualize, then alter, their psychological illness (e.g., "cognitive techniques are aimed at delineating and testing the patient's *specific* misconceptions and maladaptive thoughts" (Beck 1979, italics added) and "[w]hen you [the therapist] ask for patients' automatic thoughts, you are seeking the *actual* words or images that have gone through their mind" (Beck 2011, original italics)). If the CBT model was to consider the beliefs identified and challenged in therapy as just *one* of many equally effective ways for patients conceptualize their psychological problems, then CBT would be on similar theoretical footing with other successful therapies (such as psychodynamic or Freudian psychosexual therapies) with distinct and often incommensurate theoretical rationales. Nevertheless, CBT interventions only require that patients identify thoughts that it would make sense to have given their feelings/behaviors, or that offer coherent explanations for their feelings/behaviors. The CBT model has no method for testing whether a patient actually *had* the non-consciously entertained underlying thought "I should be out on a Saturday night" when tasked to explain what she was thinking while feeling sad and lonely, rather than it being the case that the thought was a post-hoc confabulation given by the patient to *explain* to the practitioner and herself why she felt sad and lonely.

It is important not to overgeneralize. It is certainly not the case that we have no idea about the content of our thoughts and beliefs. Many thoughts are explicitly held, repetitive, and easily identifiable (as is often the case with obsessive-compulsive disorders). The important point here is that most core beliefs and automatic thoughts are not consciously entertained. CBT interventions are based primarily on post-hoc identifications of normally non-salient beliefs and the focus of most interventions is for both the patient and practitioner to become aware of the patient's previously implicitly held thoughts. Cognitive therapy requires that patients try to identify what thoughts would make sense of their actions and feelings at some particular time. It is this theory-laden post-hoc act of thought and belief identification that we should be skeptical of.

In support of the CBT model, there do appear to be strong relationships between negative and overly rigid cognitions or thinking styles with depression (Solomon and Haaga 2004), overly rigid, ruminative, and irrational cognitions with anxiety (Clark and

Beck 2011), and overly rigid, repetitive, and intrusive cognitions with obsessive disorders (Clark 2004). However, these findings only show a correlation between styles of thinking and reported thought content, on the one hand, and the diagnosis of affective and personality disorders, on the other. These studies do not directly support the hypothesis that individuals have stable, consciously accessible core and intermediate beliefs, nor do they directly support the hypothesis that there is an introspectively accessible logical connection between one's automatic thoughts and one's intermediate and core beliefs.

It seems likely, then, that the CBT model has a significant problem. If CBT is to work as theorized, patients must be able to accurately identify their automatic thoughts and the core and intermediate beliefs that ground them. However, rather than identifying actually held thoughts, it is likely that Socratic questioning and downward arrow meaning-questioning produce confabulated post-hoc explanations for the patients' emotions and behaviors. And, crucially, this is exactly what we should expect given CBT's own assumption of world-constructing information processing. CBT interventions appear to be imposing coherence on patients' illness by giving them a way to conceptualize their emotions and behaviors rather than identifying and challenging specific thoughts and core beliefs.

5 CBT Controls for Introspective Accuracy

CBT theorists have been largely unconcerned about the issue of patient confabulation. Therapists are cautioned to be careful about the possibility of patients misidentifying their own cognitions and warned to avoid influencing patients' belief reports, but these suggestions are brief and optimistic. For example, in regards to patient's belief reports, A. Beck's (1976) manual suggests that:

> [T]he therapist should be on guard against accepting facile explanations and should check the reliability of the patient's reports of his introspections. The therapist can acquire confidence that he understands the totality of a particular experience by entering into the patient's "phenomenal world". (p. 30)

The idea here is that by carefully listening to patients' descriptions of their thoughts and beliefs, the therapist should be able to "step into the patient's world" and help identify which beliefs and cognitive patterns are playing the primary roles in patients' psychological distress. At the same time therapists are also prompted to be on guard against leading patients' narratives of their cognitions:

> Since the therapist's questions and other verbal techniques are derived from his own theory, he must be especially vigilant regarding "putting ideas in the patient's head." The therapist should be aware of his leading questions, the patient's suggestibility, and his desire to please the therapist by providing the answers he believes the therapist is seeking. (p. 142)

J. Beck's (2011) CBT manual also warns practitioners to avoid "leading" the patient, while at the same time requiring that the therapist train subjects to accurately identify

their thoughts and beliefs and teach patients about the causal relation between thoughts and emotions. J. Beck states that:

> Whenever you [the therapist] present your interpretations, you will do so tentatively and label them as hypothesis, asking patients whether they "ring true." Correct hypothesis generally resonate with the patient. (p. 198)

> You should regard your hypothesis as tentative until confirmed by the patient… Some patients are intellectually and emotionally ready to see the larger picture early on in therapy; you should wait to present it to others (especially those with whom you do not have a sound therapeutic relationship, or who do not really believe the cognitive model). As mentioned previously, whenever you present your conceptualization, ask the patient for confirmation, disconfirmation, or modification on each part. (p. 205)

While therapists are briefly cautioned to be careful about thought insertion and confabulation, CBT theorists seem confident that the process of practitioner-guided discovery allows patients to gain direct introspective access to the logical relation of their thoughts and emotions. There appears to be a number of possible sources for this confidence.

First, CBT's reliance on veridical belief reports may be thought to be supported by the use of empirically supported questionnaires such as the Cognitive Bias Questionnaire (CBQ, Krantz and Hammen 1979), the Beck Depression Inventory (BDI, Beck and Steer 1987), and the Young Schema Questionnaire (YSQ, Young and Brown 1994) that attempt to measure the accuracy and emotional valence of patients' thoughts. These questionnaires all have statistically significant, though sometimes modest, test-retest reliability (which assesses patients' scores on the same test taken at different times) (Beevers et al. 2007). For example, the Cognitive Bias Questionnaire requires that patients read vignettes involving interpersonal situations then imagine they are in the situation in question and answer a series of multiple choice questions about what they would think and how they would feel. The multiple choice options include obvious over generalizations (e.g., "nobody wants to work with me"), signs of depressive thinking (e.g., "I don't deserve the raise because I'm worthless"), or healthy responses (e.g., "I probably didn't get the job because someone else was more qualified") (Beck 1979). The questionnaire is then scored to identify possible thought errors (e.g., catastrophizing, over generalizations, or all or nothing thinking), distorted world views (e.g., the belief that the world is entirely unsafe), or distortedly valenced thoughts (e.g., overly negative interpretations of events). The CBQ, and questionnaires like it, do seem successful in identifying biases towards distorted or erroneous thought patterns and depressive thinking. However, these questionnaires do nothing to test the veridicality of patients' own belief reports about the specific contents of their automatic thoughts and beliefs. The CBQ, and questionnaires like it, measure whether patients *identify* with certain maladaptive thoughts, not whether they are trustworthy interpreters of their own cognitions. The problem with the CBT model is not that it fails in identifying whether individuals are prone to certain maladaptive psychological states, but rather that it fails in accurately identifying the specific cognitions that are the putative focus of CBT interventions.

Similarly, another common test, the Young Schema Questionnaire, asks patients to evaluate statements such as "I believe that other people can take care of me better than I can take care of myself" on a 1–6 Likert-scale (a score of 1 being "completely untrue of me" and 6 being "describes me perfectly") (Young and Brown 1994). The answers are then scored to identify patients' underlying core schemas. The YSQ has been shown to have a statistically significant test-retest reliability of .5 to .8 (Young et al. 2003). This may be taken as evidence that the questionnaire is measuring persistent, stable thoughts. However, the problem again is that the YSQ, and tests like it, only measure whether the patients' beliefs are consistent with the possession of certain core beliefs, not whether the patient actually has the beliefs in question. And while the test-retest correlation rate may be statistically significant, a 30–50 % difference in answers between tests is also evidence that the questionnaire is identifying general themes (e.g., concern about loss and self-esteem) rather than specific core beliefs (e.g., "my life is out of balance"). Given that patients can behave and feel in ways that are consistent with a number of theoretically distinct psychological explanations, the use of questionnaires is only successful at addressing intra- (rather than inter-) theory issues.

CBT theorists may also take patient and practitioner identification of cognitions at face value based on the proven efficacy of CBT interventions. The CBT model posits that specific psychological dysfunctions are caused by specific maladaptive thoughts and beliefs and prescribes a uniform treatment plan for each unique dysfunction. This uniformity makes CBT easier to study than less rigid forms of psychological intervention such as psychodynamic approaches which focus heavily on the patient-practitioner relationship and the uniqueness of each patient. CBT's superior testability has led to it becoming the most tested, and most empirically supported, form of psychological intervention. One serious problem, however, is that other forms of psychotherapy, with distinct theoretical foundations, also seem to work. While there is significant debate over whether other therapies work as well or better, there is little doubt that a number of therapies with seemingly disparate theoretical groundings (most notably psychodynamic approaches), are also effective psychological treatments. For example, Grissom's (1996) meta-analysis of 32 meta-analyses and Luborsky et al's. (2002) meta-analysis of 17 meta-analyses found statistically insignificant differences between effect-sizes between all theory-based treatments. These findings are consistent with meta-analyses by Wampold et al. (1997) and Assay and Lambert (1999). The accuracy of these meta-analyses is also supported by a number of direct comparisons between CBT and psychodynamic approaches that claim no statistically significant differences between the two approaches (Cuijpers et al. 2010; Wampold et al. 2002). CBT may work, but given that other psychological approaches work as well, CBT's efficacy is not strong evidence for its distinct theoretical model.

In response to the apparent lack of statistically significant differences in success rates between CBT and other theory-based psychological treatments, a number of theorists have attempted to identify common factors that underlie the seeming disparate treatments (Frank and Frank 1991, Wampold and Imel 2015; Messer and Wampold 2002). CBT theorists, and most notably A. Beck (Alford and Beck 1998; Clark and Beck 1999; Beck 2004), have argued that the process of cognitive change identified by the cognitive model is the primary causally efficacious common factor found in effective psychological treatments. According to A. Beck, "a common denominator of the various systems is the ascription of cognitive mechanisms to the process of therapeutic

change…[I]mprovement in the clinical condition is associated with changes in cognitive structuring of experience irrespective of the type of therapy" (2000). The idea here is that any therapy that works does so insofar as it changes how we think about the world. However, even if it is the case that cognitions play a primary role in the efficacy of therapeutic treatments, this does not mean that the CBT model is accurate. The CBT model maintains that identifying and challenging the *specific* thoughts and core beliefs that are the primary causes of the patient's symptoms is the agent of change in psychological interventions (Beck 2011; Beck 1979). While it is possible that other treatments such as psychodynamic interventions work by way of indirectly challenging specific core beliefs, this paper has argued that this is likely not the case. Rather, it is seems that challenging a patients thoughts helps give her new, adaptive ways to conceptualize her mental life regardless of what specific thoughts or beliefs she previously held.

6 Introspection and Emotions

There may be a concern that the arguments in this paper are inconsistent. This paper questions CBT's claims about the identification of patients' cognitions (e.g., 'the world is unsafe') while generally accepting patients' claims about the identification of their emotional experiences (e.g., 'I am feeling panicked') in order to diagnose mental disorders and judge the efficacy of psychotherapy. Both processes rely on introspection, and it may be thought that both processes are equally unreliable. However, the possible appearance of inconsistency is avoided if we are clear about the differences in the degree of specificity required by CBT belief and thought reports, on the one hand, and emotional experience reports, on the other.

The diagnosis and treatment of the affective disorders relies in large part on patients' claims about their emotional experiences. This would be problematic if patients in fact had very unreliable introspective access to their emotional experiences. But this appears to not be the case. Individuals certainly have restricted introspective access to the *causes* of their emotions (just as they have restricted access to the causal processes underlying cognitions) (e.g., Nisbett and Wilson 1977). Retrospective reports of emotional experiences, like retrospective reports of cognitions, are also susceptible to systematic biases and distortions (Robinson and Clore 2002; Thomas and Diener 1990). It is also unclear how fine-grained individuals' introspective access to their mental states can be, and whether or not the act of introspection necessarily alters or distorts the experience (e.g., Schwitzgebel 2008; Hurlburt 2011). However, it is generally agreed that individuals have *coarse-grained* awareness of their emotional experiences; we can usually, though not always, tell when we are currently experiencing strong emotions such as depression, joy, panic, or fear even if we cannot identify the exact valence or guarantee that the act of introspection has not altered the experience. This is important because the diagnosis of affective disorders is concerned with patients' coarse-grained experiences (e.g., feelings of low mood, anxiety, or panic), rather than with their awareness of the causes, and fine-grained nature of, their emotional experiences (which may in fact be mistaken or confabulatory).

It is not inconsistent to claim that patients can be, and often are, mistaken about the exact nature and causes of their emotional distress (as well as often being mistaken

about the causal processes responsible for much of their behavior), while also cautiously accepting their coarse-grained, real-time reports of their current emotional experiences. It is these coarse-grained dysfunctional emotional experiences that most therapies for affective disorders are attempting to treat and ameliorate. And, as this paper has argued, CBT mistakenly assumes that these coarse-grained emotional experiences are primarily caused by fine-grained, and accurately identifiable, thoughts and beliefs.

7 Conclusions

CBT works, but likely not for the reasons given by the CBT model. CBT is based on the cognitive model of psychological functioning which postulated a nested hierarchy of consciously accessible cognitions consisting of automatic thoughts, intermediate beliefs, and core beliefs. However, there is good reason to be skeptical that core and intermediate beliefs are accurately and reliably consciously accessible or that they exist in the form postulated; they may serve as useful descriptions or ways to conceptualize psychological illness, but patients are likely not accurately identifying causally efficacious cognitive structures. Furthermore, it is likely that the Socratic method and downward arrow techniques proscribed by the CBT model lead to confabulation rather than accurate identification of dysfunctional or maladaptive automatic thoughts, and thus identification of deeper logically connected cognitions. While it may be the case that changes in cognitive processing are the basis of successful therapeutic treatments, the specific model posited by CBT theorists is likely false.

There remains the question that if CBT does not work by accurately identifying and challenging beliefs and thoughts, why does it work? There are a number of possible answers. First, CBT may work, not by accurately challenging specific cognitive content, but by challenging maladaptive cognitive processes. Recent cognitive theories have argued for a change of therapeutic focus from the cognitive content of automatic thoughts and schemas, to thoughts about thinking. Mindfulness-based cognitive therapy (Segal et al. 2004, 2012) maintains that affective change is not just about changing content of depressive thinking, but also about changing one's relationship to one's thoughts. Mindfulness-based therapies posit that it is the change in one's perspective about one's negative thoughts, rather than challenging the thoughts themselves, which leads to direct and lasting change in psychotherapy. Related views can be found in Dialectical Behavioral Therapy (DBT) and Acceptance and Commitment Therapy (ACT), which both focus implicitly on "decentering" one's relationship to one's cognitions (Segal et al. 2012). Similarly, Meta-Cognitive Therapy (MCT), developed by Wells (2009), focuses on metacognitions (or "beliefs about thinking") rather than on specific cognitive content. MCT "proposes that disturbances in thinking and emotion emerge from metacognitions that are separate from these other thoughts and beliefs emphasized in CBT" (Wells 2009). Instead of challenging the content of specific core beliefs or automatic thoughts, MCT aims to challenge the beliefs about thinking (e.g., "if I worry about my symptoms, I won't miss anything important") from which these other cognitions are supposedly derived (Wells 2009).

While these meta-cognitive and decentering approaches are offered as rivals to CBT, the differences may be merely superficial. Both MBCT and MCT share many of the

same theoretical commitments about cognitive primacy and therapeutic focus on consciously accessible cognitions. MCT, like CBT, assumes that consciously accessible thoughts or beliefs play a primary role in therapy and focuses on accurately identifying and challenging beliefs about beliefs (rather than CBT's focus on beliefs about the world, self, and future). And MBCT, like CBT and MCT, aims at altering patients' perspectives on their negative cognitions; MBCT focuses on decentering and detaching oneself from one's thoughts, while CBT aims to challenge patients' views about the rationality or validity of their thoughts. Both MBCT and CBT work by patients identifying specific cognitions; they differ only in how the patient is instructed to treat these beliefs. It is unclear, then, whether the new theoretical and therapeutic focus on thoughts about thoughts offers a genuine theoretical challenge to the CBT model.

Another plausible explanation is that CBT may work for the same reasons that other effective therapies work; CBT fosters a challenging and caring therapeutic allegiance between patient and practitioner and offers a plausible explanation and method of treatment for the patient's problems. The "common factors" theory postulates that non-specific (to any given theory) common factors (such as a healing setting, a coherent theory/rationale, a healing ritual, and an emotionally charged confiding relationship) play a dominant role in psychological change (as opposed to the specific factors postulated by distinct theories) (Frank and Frank 1991; Anderson et al. 2010; Messer and Wampold 2002). This response, while plausible, is underdeveloped. It still must be explained why these factors are necessary for successful therapy and what these common factors have in common. Most common factors theorists take as their model Jerome Frank's idea that therapy is best understood as a form of rhetoric (Frank and Frank, 1961/Frank and Frank 1991). According to Frank, psychological healing is a matter of persuasion with the common factors being necessary components. What is left unexplained, and what the cognitive model purports to answer, is why persuasion is the mechanism of change in psychotherapy. Rather than being a rival to the CBT model, the common factors approach implicitly assumes something like a cognitive model of psychopathology; therapy, like rhetoric, is just supposed to be a matter of convincing the patient to accept more adaptive beliefs. Therapy may work by imposing coherence upon a person's mental life, but it still must be explained why and how this might work.

References

Alford, B.A., and A.T. Beck. 1998. *The integrative power of cognitive therapy*. New York: Guilford Press.

Anderson, T., K.M. Lunnen, and M. Ogles. 2010. Putting models and techniques in context. In *The heart and should of change: Delivering what worlds in therapy*, 2nd edn, eds. B.L. Duncan, S.D. Miller, B.E. Wampold, and M.A. Hubble, 143–166. Washington, DC: American Psychological Association.

Asay, T.P., and M.J. Lambert. 1999. The empirical case for the common factors in therapy: quantitative findings. In *The heart and soul of change: What works in therapy*, eds. M.A. Hubble, B.L. Duncan, and S.D. Miller, 23–55. Washington, DC, US: American Psychological Association.

Bargh, J.A., M. Chen, and L. Burrows. 1996. Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology* 71(2): 230–244.

Beck, A.T. 1967. *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.

Beck, A.T., A.J., Rush, B.F., Shaw, and Emery, G. 1976. *Cognitive therapy and the emotional disorders*. New York: International Universities Press.

Beck, A.T. 1996. Beyond belief: A theory of modes, personality, and psychopathology. In *P.M. Salkovskis* eds. Frontiers of cognitive therapy, 1–25. New York: Guilford Press.

Beck, A.T. 1979. *Cognitive therapy of depression*. New York: Guilford Press.

Beck, A.T. 2004. A cognitive continuum. In *Cognition and psychotherapy*, eds. A. Freeman, M. Mahoney, P. Devito, and D. Martin, 107–220. New York, NY, US: Springer.

Beck, A.T. 2005. The current state of cognitive therapy: a 40-year retrospective. *Archives of General Psychiatry* 62(9): 953–959.

Beck, J.S. 2011. *Cognitive behavior therapy, second edition: basics and beyond*. New York: Guilford Press.

Beck, A.T., and B.A. Alford. 2009. *Depression: Causes and Treatment* (2nd ed). Philadelphia: University of Pennsylvania Press.

Beck, A.T., and D.J. Dozois. 2011. Cognitive therapy: current status and future directions. *Annual Review of Medicine* 62: 397–409.

Beck, A.T., and R.A. Steer. 1987. *BDI, Beck depression inventory: manual*. TX: Psychological Corporation San Antonio.

Beevers, C.G., D.R. Strong, B. Meyer, P.A. Pilkonis, and I.W. Miller. 2007. Efficiently assessing negative cognition in depression: An item response theory analysis of the dysfunctional attitude scale. *Psychological Assessment* 19(2): 199–209.

Bertrand, M., & Mullainathan, S. (2003). *Are Emily and Greg more employable than Lakisha and Jamal*? *A field experiment on labor market discrimination.* National Bureau of Economic Research (Working Paper 9873).

Brewin, C.R. 1996. Theoretical foundations of cognitive-behavior therapy for anxiety and depression. *Annual Review of Psychology* 47(1): 33–57.

Butler, A.C., J.E. Chapman, E.M. Forman, and A.T. Beck. 2006. The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review* 26(1): 17–31.

Clark, D.A. 2004. *Cognitive-behavioral therapy for OCD*. New York, New York: Guilford Press.

Clark, D.A., and A.T. Beck. 1999. *Scientific foundations of cognitive theory and therapy of depression*. New York: Wiley & Sons.

Clark, D.A., 2011. *Cognitive therapy of anxiety disorders: science and practice*. New York, USA: New York: Guilford Press.

Coyne, J.C., and I.H. Gotlib. 1983. The role of cognition in depression: A critical appraisal. *Psychological Bulletin* 94(3): 472.

Crits-Christoph, P. 1997. Limitations of the dodo bird verdict and the role of clinical trials in psychotherapy research: Comments on Wampold et al. (1997). *Psychological Bulletin* 122(3): 216–220.

Cuijpers, P., F. Smit, E. Bohlmeijer, S.D. Hollon, and G. Andersson. 2010. Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *The British Journal of Psychiatry: the Journal of Mental Science* 196(3): 173–178.

Eagle, M.N., and D.L. Wolitzky. 1992. Psychoanalytic theories of psychotherapy. In *History of psychotherapy: A century of change*, eds. D.K. Freedheim, H.J. Freudenberger, J.W. Kessler, S.B. Messer, D.R. Peterson, H.H. Strupp, and P.L. Wachtel, 109–158. Washington, DC, US: American Psychological Association.

Ellis, A. 1962. *Reason and emotion in psychotherapy.* Secaucus, NJ: Lyle Stuart.

Ellis, A. (1994). *Reason and emotion in psychotherapy*, Revised and Updated. New York: Carol Publishing.

Frank, J.D., and J.B. Frank. 1991. *Persuasion and healing: A comparative study of psychotherapy*, 3rd edn. Baltimore, MD, US: Johns Hopkins University Press.

Freud, S. 2000. *Three essays on the theory of sexuality* (J. Strachey, Trans.). New York: Basic Books. (Original work published in 1905).

Gabbard, G.O. 2000. *Psychodynamic psychotherapy in clinical practice*, 3rd. edn. Washington, DC: American Psychological Association.

Gipps, R.G.T. 2013. Cognitive behavior therapy: a philosophical appraisal. In *The Oxford handbook of philosophy and psychiatry*, eds. K. Fulford, M. Davies, R. Gipps, G. Graham, J. Sadler, G. Stanghellini, and T. Thorton. Oxford: Oxford University Press.

Grissom, R.J. 1996. The magic number .7 + −.2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology* 64: 973–982.

Haidt, J. 2006. *The happiness hypothesis*. Cambridge, MA: Basic Books.

Harden, C.D., and M.R. Banaji. 2013. The nature of implicit prejudice: implications for personal and public policy. In *The behavioral foundations of public policy*, ed. E. Shafir, 13–31. New Jersey: Princeton University Press.

Harwood, T.M., Beutler, L.E., and M. Charvat. (2009). Cognitive-behavioral therapy and psychotherapy integration. In Dobson K, eds. Handbook of cognitive-behavioral therapies, 94–130. New York: Guilford Press.

Hurlburt, R.T. 2011. *Investigating pristine inner experience: Moments of truth*. Cambridge: Cambridge University Press.

Isen, A.M., and P.F. Levin. 1972. Effect of feeling good on helping: cookies and kindness. *Journal of Personality and Social Psychology* 21(3): 384–388.

Johansson, P., L. Hall, S. Sikström, and A. Olsson. 2005. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 310(5745): 116–119.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Knowles, E.D., B.S. Lowery, and R.L. Schaumberg. 2010. Racial prejudice predicts opposition to Obama and his health care reform plan. *Journal of Experimental Social Psychology* 46(2): 420–423.

Krantz, S., and C.L. Hammen. 1979. Assessment of cognitive bias in depression. *Journal of Abnormal Psychology* 88(6): 611.

Leahy, R.L. 1996. *Cognitive Therapy: Basic Principles and Applications*. New Jersey: Jason Aronson, Incorporated.

Lebowitz, M.S. 2014. Biological conceptualizations of mental disorders among affected individuals: A review of correlates and consequences. *Clinical Psychology: Science and Practice* 21(1): 67–83.

Leichsenring, F., W. Hiller, M. Weissberg, and E. Leibing. 2006. Cognitive-behavioral therapy and psychodynamic psychotherapy: techniques, efficacy, and indications. *American Journal of Psychotherapy* 60(3): 233–259.

Liu, J., D. Smeesters, and K. Vohs. 2012. Reminders of money elicit feelings of threat and reactance in response to social influence. *Journal of Consumer Research* 38(6): 1030–1046.

Luborsky, L., R. Rosenthal, L. Diguer, T.P. Andrusyna, J.S. Berman, J.T. Levitt, and E.D. Krause. 2002. The dodo bird verdict Is alive and well—mostly. *Clinical Psychology: Science and Practice* 9(1): 2–12.

McEachrane, M. 2009. Capturing emotional thoughts: the philosophy of cognitive-behavioral therapy. In *Emotions and Understanding: Wittgensteinian perspectives*, eds. Y. Gustafsson, C. Kronqvist, and M. McEachrane. New York: Palgrave Macmillan.

Messer, S.B., and B.E. Wampold. 2002. Let's face facts: common factors are more potent Than Specific Therapy ingredients. *Clinical Psychology: Science and Practice* 9(1): 21–25.

Nasar, J.L., and A.S. Devlin. 2011. Impressions of psychotherapists' offices. *Journal of Counseling Psychology* 58(3): 310–320.

Neenan, M., and W. Dryden. 2005. *Cognitive therapy*. London: SAGE Publications.

Nisbett, R.E., and T.D. Wilson. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review* 84(3): 231–259.

Norcross, J.C., and C.P. Karpiak. 2012. Clinical psychologists in the 2010s: 50 years of the APA division of clinical psychology. *Clinical Psychology: Science and Practice* 19(1): 1–12.

Nosek, B.A., M. Banaji, and A.G. Greenwald. 2002. Harvesting implicit group attitudes and beleifs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice 6*(1): 101.

Nosek, B.A., F.L. Smyth, J.J. Hansen, T. Devos, N.M. Lindner, K.A. Ranganath, C.T. Smith, K.R. Olson, D. Chugh, A.G. Greenwald, and M.R. Banaji. 2007. Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology* 18(1): 36–88.

Pescosolido, B.A., J.K. Martin, J.S. Long, T.R. Medina, J.C. Phelan, and B.G. Link. 2010. "A disease like any other"? A decade of change in public reactions to schizophrenia, depression, and alcohol dependence. *American Journal of Psychiatry* 167: 1321–1330.

Robinson, M.D., and G.L. Clore. 2002. Beliefs, situations, and their interactions: towards a model of emotion reporting. *Psychological Bulletin* 128: 934–960.

Schnall, S., J. Benton, and S. Harvey. 2008. With a clean conscience: cleanliness reduces the severity of moral judgments. *Psychological Science* 19(12): 1219–1222.

Schwitzgebel, E. 2008. The unreliability of naive introspection. *Philosophical Review* 117(2): 245–273.

Segal, Z.V., J.D. Teasdale, and J. Mark. 2004. Mindfulness-based cognitive therapy: theoretical rationale and empirical status. In *Mindfulness and acceptance: Expanding the cognitive-behavioral tradition*, eds. S.C. Hayes, V.M. Follette, and M.M. Linehan, 45–65. New York: Guilford Press.

Segal, Z.V., J.M.G. Williams, and J.D. Teasdale. 2012. *Mindfulness-based cognitive therapy for depression*. New York: Guilford Press.

Skinner, B.F. 1977. Why I am not a cognitive psychologist. *Behaviorism* 5(2): 1–10.

Solomon, A., and D.A. Haaga. 2004. Cognitive theory and therapy in depression. In *Cognitive therapy across the lifespan*, eds. M.A. Reinecke, and D.A. Clark, 12–40. Cambridge, UK: Cambridge University Press.

Steinpreis, R.E., K.A. Anders, and D. Ritzke. 1999. The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41: 509–528.

Teasdale, J.D. 1997. The relationship between cognition and emotion: the mind-in-place in mood disorders. In *Science and practice of cognitive behaviour therapy*, eds. D.M. Clark, and C.G. Fairburn, 67–93. New York, US: Oxford University Press.

Teasdale, J.D., and P.J. Barnard. 1993. *Affect, cognition, and change: Re-modelling depressive thought (vol. xv)*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Thomas, D.L., and E. Diener. 1990. Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology* 59(2): 291–297.

Wampold, B.E., and Imel, Z.E. 2015. *The Great Psychotherapy Debate: The Evidence For What Makes Psychotherapy Work*. New York: Routledge.

Wampold, B.E., G.W. Mondin, M. Moody, F. Stich, K. Benson, and H. Ahn. 1997. A meta-analysis of outcome studies comparing bona fide psychotherapies. *Psychological Bulletin* 122(3): 203.

Wampold, B.E., T. Minami, T.W. Baskin, and S. Callen Tierney. 2002. A meta-(re) analysis of the effects of cognitive therapy versus "other therapies" for depression. *Journal of Affective Disorders* 68(2): 159–165.

Wells, A. 2009. *Metacognitive therapy for anxiety and depression (vol. xvii)*. New York: Guilford Press.

Westbrook, D., H. Kennerley, and J. Kirk. 2011. *An introduction to Cognitive Behaviour Therapy: Skills and Applications*. London: SAGE Publications.

Wilson, T.D. 2002. *Strangers to Ourselves: Discovering the Adaptive Unconscious*, New Ed edn. Cambridge: Mass.: Belknap Press.

Young, J.E., J.S. Klosko, and M.E. Weishaar. 2003. *Schema Therapy: A Practitioner's Guide*. New York: Guilford Press.

Young, J.E., and G. Brown. (1994). Young Schema Questionnaire, 2nd edition. In Young, J.E., *Cognitive Therapy for Personality Disorders: A Schema-Focused Approach* (rev. edition). Sarasota, FL: Professional Resource Press.