INTRODUCTION: COGNITIVE BEHAVIOR THERAPY AND PHILOSOPHY

Cognitive behavior therapy (CBT) is a broad church embracing many theories, models, and techniques, often now held in conjunction with other approaches to the mind in health and distress (e.g., compassion, mindfulness, dialectics, psychodynamics, etc.). Within this church there are denominations that complement, but also those that compete, with one another, both in their empirical claims and also in some of their underlying theoretical principles (Hofmann and Asmundson 2008).

There are, for example, those who treat the “cognitive” aspect of CBT theory as referring specifically to maladaptive thoughts and beliefs (e.g., Beck 1979). These psychological states are considered distinct from what members of this congregation would stipulate as the non-cognitive states and processes of perception, emotion, and behavior. The causal-explanatory force of appeals to cognition should now be clear: disorders of behavior and feeling are to be explained by reference to disturbances in underlying, ontologically distinct, cognitive states and processes. Other theorists, however, tie what is “cognitive” not specifically to thought and belief, but rather to the processing of meaning (e.g., Teasdale 1997). They do not assume that meaning is an exclusive function of thought or belief, and instead allow it to already belong properly to emotion, behavior, and perception. In such cases the explanatory force of a cognitive model may derive from its identification of the distinct structures of meaning immanent within emotion and behavior themselves, rather than from showing how antecedent thoughts allegedly give rise to such disturbed emotion and behavior.

Because there are such different versions not only of the “C” (Dobson and Dozois 2010), but also of the “B” (Hayes and Brownstein 1986) in versions of what are still often described as “CBT” it makes little sense to offer a philosophical critique “of CBT” per se, and such
a wide-ranging critique is not attempted in this chapter. Nevertheless, underlying several popular CBT models and treatments (henceforth “CBTs”) are, I shall suggest, several questionable assumptions about the significance of cognitive factors in psychopathology. In this chapter I add to previous critique (McEachrane 2009; Lacewing 2004; Whiting 2007) in specifying four related philosophical misunderstandings regarding the nature of the mind in health and distress that, I claim, can sometimes affect CBT’s understanding both of psychopathology and of therapeutic action.

Before I begin I think it might be helpful if I set out, using some examples, what makes for a distinctly philosophical critique of a particular CBT theory. Consider first the following questions that can, and have, been asked of the CBTs:

1. For what range of difficulties, experienced by what kinds of populations, are CBTs effective?
2. Are CBTs more or less effective than other therapies (e.g., psychodynamic, narrative, person-centered or drug therapies) and practical interventions (e.g., employment and dating advice, social sport, and gardening)?
3. How durable are the effects of CBTs compared with other treatments?
4. When a CBT is effective, can this be explained using typical CBT models of psychopathology and therapeutic action, or must we account for it using ideas taken from other cognitive, systemic, or psychodynamic theories of therapeutic action?
5. When a CBT is found to be more or less effective than another therapy, is this to be accounted for:
   a. ideologically (i.e., because of treatment effects due to the convictions of investigators, therapists or patients), or because of
   b. differential treatment efficacy, or because
   c. measures are being used which are tailored to what one but not another therapy would view as meaningful or valuable change (e.g., is it characterological or symptomatic changes that are being measured)?

These are all important questions for therapists, patients, and public health policymakers. Yet there is, I suggest, little that is distinctly philosophical about them. Part of what it means to say this is that they are not answerable through reflection alone, but must instead be answered by empirical investigation. The exception is the last (5c): reflection on what counts as efficacy is a form of theoretical reflection on conceptual matters of meaning, rather than discernment of empirical matters of fact. Even so, such theoretical reflection is not, in the sense in which I shall use the term here, itself particularly philosophical. For whilst all theoretical thought concerns itself with matters of meaning, what I have in mind by distinctly philosophical thought instead concerns itself with whether what is said can genuinely be understood in the way its author seems to hope for it. Philosophical reflection in this sense concerns itself not simply with meaning but more particularly with meaningfulness; not with how, but rather with whether, something can really be understood in the way it invites us to understand it.1

1 In making this distinction I am not intending to be understood as suggesting that a hard and sharp distinction can everywhere and always be drawn between matters empirical and theoretical, or between matters theoretical and philosophical.
In what follows I articulate four ways in which CBTs can sometimes invite the psychological theorist to construe ourselves to ourselves, both in our healthy functioning and in our emotional distress, in ways that, it seems to me, go against what it means to be a human subject. More particularly, I shall consider certain conceptions, tacitly present within the psychopathological theories offered by various CBTs, of what it is to think, feel, and meaningfully respond. My claim will be that whilst these conceptions appear to motivate the practice, conceptually dovetail with the theory, and validate the distinctive scientific self-conception, of various CBTs, they nevertheless simply do not tally with what reflective understanding reveals as the character of human mental and emotional life. Such reflection reveals that our mindedness has a character and richness not always aptly captured by the concepts and theories employed in certain CBT theories. Therapists guided more by such theories than by their own humanity may accordingly run the risk of embodying or offering to their patients a distorted or impoverished understanding of emotional distress and therapeutic action. Given that such CBTs can nevertheless prove helpful to some patients, it follows that they must sometimes do so for reasons other than those they themselves suggest.

Aspects Versus Causes

The CBTs, along with other cognitive approaches in psychology, often present their understanding of psychopathology in a “box and arrow” format. The arrows are intended to represent causal relations; the boxes represent what are intended to be understood as isolable cognitive, behavioral or emotional factors. Such schematic models are appealing because they are clear-cut and appear to offer a cogent rationale for investigation and intervention. Consider the highly condensed summaries of typical cognitive models shown in Box 72.1.

Such models are, I find, often useful in clinical practice since they clearly depict what may be the main meanings and maintaining factors in the various difficulties with which patients present; they also immediately suggest strategies for intervention. They may also attempt to do justice to the complex, dynamic (non-linear) causality operative in the mind. My objection at this stage is not to their clinical or heuristic utility, and neither am I alleging any excessive simplicity or linearity in their causal schematics; it is instead a rather modest demurral to a quite different aspect of their implicit causal theorization of the mind. The objection could be summarized by saying that such models tend to encourage us to mistake:

(a) a fact about our representations of the mind for (b) a fact about what is represented. The (a) fact about our representations that I have in mind is that we can separately depict various aspects of (say) a depressive state, such as our behavioral dispositions, thoughts, motivation, mood, affects, etc. My objection here is that this cannot by itself be taken to indicate a (b) fact about what is represented: that there are separately existing assumptions and perceptions and moods and motivational states—states “in us”, as we can hardly now but put it—linked by causal connections (the arrows in the diagrams (Box 72.1)).

Take the case of panic. The CBT model separates out a stimulus, a perception of it as threatening, a state of apprehension, and bodily sensations, and places these separated phenomena into a causal sequence. Yet whilst all these ingredients are clearly present in a panic attack, and whilst we can separate them for formal (descriptive) purposes, it is not obvious that they really are, in the normal run of things, best conceived of as separate states within
Box 72.1 CBT formulation summaries

One aspect of contemporary CBT models of panic can be elaborated as follows: A patient experiences a bodily sensation. They catastrophically misinterpret this as a sign of danger (e.g., that they are fainting or having a heart attack). This leads to a further perception of threat, giving rise to further fear, which in turn gives rise to further sensations that are in turn catastrophically misinterpreted. The patient is left in a rapidly cycling, self-fuelling anxiety state (cf. Clark 1989) (Fig. 71.1).

![CBT panic model](image)

The depressed patient as encountered by the CBT therapist is an individual with underlying pessimistic assumptions that organize his or her experience of the world. When activated by critical incidents these assumptions are thought to lead to the development of negative automatic thoughts (e.g., gloomy and non-deliberative thoughts about the current situation, self, or future). These then give rise to negative emotions (e.g., guilt or fear), a lowering in mood, and consequently to a reduction in motivation and activity and engagement. A vicious cycle of depressed thoughts and depressed mood results (cf. Beck 1979; Fennell 1989).

The patient with social anxiety often carries underlying assumptions (e.g., “unless someone shows they like me, they dislike me”). The CBT model describes how these can be activated in social situations, resulting in a belief that they are in “social danger” (e.g., believing that others will treat me badly) (Fig. 72.2). Such beliefs may then cause anxiety, which in turn can cause various somatic and cognitive symptoms (sweating, mind going blank, shaking). These in turn can heighten a self-conscious mode of attention in which the patient considers that they are, for example, likely to make fools out of themselves (cf. Clark and Wells 1995).

(continued)
a person which causally trigger one another. A better description of them may accordingly be as different aspects of the same (anxiety) state. We are “stirred up”—in both our action-readiness and in a somatosensory way—by fearful stimuli, and our being stirred up in this way is itself of a piece with our registering of that stimulus as fearful. The components of our fear response are not in any obvious way isolable and several causally related states in us, but may rather be more naturally understood as merely notionally separable aspects of the state which we are in.

Much the same can be said about the CBT conceptions of depression and social anxiety outlined earlier. It can almost seem as if something of the passive and objectifying voice of depressive anxiety has found its way into the CBT conceptualization of the nature of our mental life per se. Thus we now have assumptions that can be “activated”; meanings become reified into occurring processes and states to be known as “cognitions”; the agent become a locus of “behaviors.” Everywhere we have to do with nouns, and nowhere with verbs: to

---

do with nouns that bolster an appearance that we are referring to several isolable entities and processes, rather than with their root verbs that refer to the action and attitude of a single living entity. The living agent who makes assumptions, understands his experience a particular way, and undertakes actions becomes instead a locus of separate inner and outer states which causally trigger one another. The concept of a unitary living person as featured in our everyday emotional and moral discourse accordingly risks being degraded into that of a bundle of inner and outer states and processes.

CBT models can be useful as heuristics for thinking about psychopathology, and when they do outline genuine causal interactions (two genuinely causal interactions in the earlier discussed models include those between fear and bodily sensations in panic, and inactivity and depressed mood in depression) they significantly help to guide therapy. Furthermore, their use of mechanistic and objectivizing rhetoric which reduces diverse (causal, dispositional, constitutive, etc.) relations to causal relations is not in itself a mortal sin; the CBT theorist is after all not pretending to provide us with a nuanced metaphysics of mind. It is also true that a spuriously mechanistic model of the mind need not lead to a mechanistically pursued therapy; whether that risk obtains in fact would need to be empirically investigated. My criticism at this juncture is instead as follows. First, that the cognitive models can have the somewhat misleading appearance of providing more by way of a scientific psychological explanation than is really available; they are therefore potentially scientifically misleading and may accordingly impede CBT research. Second, they often appear to possess a greater degree of scientific precision than found in other models of therapy—even though this appearance may be partly or solely due only to their having misleadingly dressed up descriptions of different aspects of the patient’s disturbance as a causal-explanatory model of it. For, to repeat, as often as not, what such models are really describing are different aspects of the same state we are in, rather than describing scientifically teased out, causally linked, interruptible, states and processes within us.

**Attitudes Versus Cogitations**

Whilst all CBT theorists adhere to the centrality of cognition in emotional disturbance, not all theorists appear to agree on the meaning of the term, a term that in any case is most often left undefined. As described in the introduction, for some the term appears to refer to any aspect of psychological performance (e.g., in perception, action, emotion, judgment, etc.) that involves the conscious or preconscious processing or discrimination of meaning. Under consideration may be enduring attitudes—dispositions of thought, feeling and behavior which may become sedimented in character—or passing emotions, perceptions, cogitations, and imaginings. For others, including the founding fathers of CBT, the term refers more exclusively to our thinking, which in itself is understood as realized in the words and images that pass across our minds. Such thinking is conceived of as the second or third tier of a hierarchy of cognitive elements which move from the most subconscious and dispositional (the underlying schemata) level, through to more readily avowable maladaptive beliefs, up to occurrence automatic thoughts passing across the mind. What follows in this section concerns only the latter sense of cognition (as occurrent thoughts passing across the
Here are some examples of the idea that it is either the content of (in traditional CBT), or the patient’s fused relationship with (in “third-wave” CBTs such as acceptance and commitment therapy—“ACT”), inner representational events that maintain and mediate emotional disturbance:

CBT therapists encourage [patients] to become aware of their thoughts and thought processes. Cognitions are generally classified into negative automatic thoughts and dysfunctional or irrational beliefs. Negative automatic thoughts are thoughts or images that occur in specific situations when an individual feels threatened in some way. (Hofman and Asmundson 2008, p. 4)

When a person is able to fill in the gap between an activating event and the emotional consequences, the puzzling reaction becomes understandable. With training, people are able to catch the rapid thoughts or images that occur between an event and an emotional response. (Beck 1979, p. 26)

ACT aims to alter the context in which thoughts occur so as to decrease the impact and importance of difficult private events. . . . Clinically we want to teach clients to see thoughts as thoughts, feelings as feelings, memories as memories, and physical sensations as physical sensations. None of these private events are inherently toxic to human welfare when experienced for what they are. (Hayes et al. 2004, p. 8)

[In] cognitive fusion … the event and one’s thinking about it become so fused as to be inseparable and that creates the impression that verbal construal is not present at all. . . . A worry about the future is seemingly about the actual future, not merely an immediate process of construing the future. The thought “Life is not worth living” is seemingly a conclusion about life and its quality, not a verbal evaluative process going on now. (Hayes et al. 2004, p. 25)

In these examples we see clearly an idea that crops up not infrequently in the CBT literature: that to think something is to have a thought pass across one’s mind. We are, it is said, often oblivious to such inner events even occurring, perhaps (in the case of ACT) because we are overidentified or “fused” with them. The therapist’s job is accordingly to help us notice and either simply de-fuse from, or more aggressively challenge, these events that supposedly mediate between outer events and our feelings (on which more in the “Feeling and believing” section). But as McEachrane (2009) suggests:

To say of someone that they ‘thought that $p$’ does not imply that they ‘thought of $p$’ or ‘thought about $p$’ or formulated $p$ or that $p$ occurred to them or was in their thoughts. . . . [If] a client says that in a particular situation they thought that, say, ‘I’ll never be like them’ or . . . ‘I’m not a likeable person’, then this does not necessarily mean that the client in that situation thought of these things, formulated these things to herself, that these things occurred to her or were in her thoughts. If you ask me what I think about the economic situation I need not wait on inner mental events to occur which I can then report. Instead I simply tell you how the situation seems to me, and in this telling I express what we could call my “attitudes” rather than report on my occurrent “cogitations.” (p. 86)

It is important to be clear about the scope of this criticism. It is the aim of many psychological therapies to help the patient to take a more flexible, reflective, or “ironic” stance toward their own attitudes (Lear 2003). Psychodynamic therapies, for example, may aim to help a patient develop in their “mentalizing”, to become aware that theirs is but one among many
possible ways of thinking about things, and to increase their inner playfulness or psychological flexibility (Holmes 2009). Further, it is true that the depressed person can indeed become ruminatively lost in an inner world of depressed and depressing cogitations; in this case, talk of “thoughts” can indeed often be taken as talk of inner musings or imagery. Helping the patient become aware that this is going on, and encouraging them to instead harness their attentional resources to, and increase their physical engagement with, the external world are well known to be therapeutically valuable. But treating thoughts as if they are typically inner events leaves a therapist at risk of talking past the patient. Consider the following:

Therapist: Now I’d like to spend a few minutes talking about the connection between thoughts and feelings. Can you think of some times this week when you felt upset?
Patient: Yeah. Walking to class this morning.
Therapist: What emotion were you feeling: sad? anxious? angry?
Patient: Sad.
Therapist: What was going through your mind?
Patient: I was looking at these other students, talking or playing Frisbee, hanging out on the lawn.
Therapist: What was going through your mind when you saw them?
Patient: I’ll never be like them.
Therapist: Okay. You just identified what we call an automatic thought. Everyone has them. They’re thoughts that just seem to pop into our heads. We’re not deliberately trying to think about them; that’s why we call them automatic. Most of the time, they’re real quick and we’re much more aware of the emotion—in this case, sadness—than we are of the thoughts. Lots of times the thoughts are distorted in some way. But we react as if they’re true.

Here the therapist is interested in the patient’s cogitations, and takes the patient’s reply (“I’ll never be like them”) as literally describing something going through the mind. But nothing in what the patient says confirms the correctness of the idea that they are experiencing “automatic thoughts”; instead they may simply be describing their beliefs. The possible discrepancy is even clearer with some of the reported thoughts described by Hayes et al. (2004) (“A worry about the future is seemingly about the actual future, not merely an immediate process of construing the future. The thought “Life is not worth living” is seemingly a conclusion about life and its quality, not a verbal evaluative process going on now.”) You ask me: “Why are you troubled?” and I say “I’m worried about being made redundant, and the possible effect of this on my family life.” Whilst it may indeed be possible for me to be overly worried and lose perspective, it is not at all clear that this is best described as a matter of fusion with “immediate processes of construing” or with “verbal evaluative processes going on now.” By contrast with what Hayes suggests, we can and indeed should say that a thought that life is not worth living is indeed a judgment about life and its quality, regardless of whether the judgment is expressed in inner speech; and we also can and indeed should say that a worry about the future is indeed about the future, regardless of whether it manifests in a current episode of worrying.

The risks of misconstruing the phenomenology of thought do not include merely that of the therapist and patient talking past one another. An oft discussed concern in therapeutic circles is when we can count externalization—the capacity to develop a noticing, more distant and reflective, attitude toward one’s own thoughts—as a healthy, and when as an unhealthy, procedure. Externalization can, for example, be a healthy strategy for the patient who has become swept along by ruminative cycles of cogitations—as typified by certain
cases of depression. It can also be helpful when the patient is unable to appreciate that they do have one amongst other possible attitudinal stances toward the world—as typified by certain cases of borderline personality disorder. But it can be unhelpful when the therapeutic task is that of helping a patient take ownership of her attitudes, stand behind them as her own, live from and embody them, count herself as their true subject—rather than, say, using irony or intellectualization or projection (attributing them elsewhere) as a defense, or seeing them as unwanted foreign intrusions. And so one risk in misconstruing all thought as cognition which the patient can relate to, rather than live from, may be that of encouraging an unhealthy form of externalization or self-alienation. For some of our thoughts—in particular certain emotionally painful ones which register personal losses—precisely do need to be engaged with and suffered, rather than reflectively noticed, if the patient is ever to be able to acknowledge and thereby move past their painful experiences.

Another risk is that of underestimating the therapeutic task. By misconstruing thought as thinking—by conflating attitudes with cogitations—the CBT therapist may start to attempt to help their patient change what he thinks (his attitudes) through changing his relation to, or the content of, his thinking (his cogitations). Yet it is surely natural to suppose that our cogitations are often enough a function of our attitudes and not, on the whole, vice versa. If, that is, my attitude is that my life is not worth living then, when I am invited to attend only to my own cogitation, I shall most likely be attending to it—say to my occurrent thought of “this is all hopeless”—from just this attitude. In such a case my challenging of, or distancing myself from, cogitations that reflect my attitudes risks becoming little more than a strategy promoting bad faith or self-alienation. What is really required is, rather, a change of attitude, a change which may at times be facilitated by attention to one’s cogitations but which is unlikely to proceed solely from such a change. In CBT terms what may be required is not work on “automatic thoughts” but rather deeper work on “core beliefs” or “schemata”; the risk of misconstruing attitudes as cogitations is that this requirement gets overlooked. The patient may of course benefit in all sorts of ways by being able to attend to their occurrent ruminations which can often go unnoticed. First, however, they need to possess a healthy standpoint from which to assess their own mental contents and processes—and this can itself take significant therapeutic work.

**FEELING AND BELIEVING**

Perhaps the central claim of much CBT is that disturbances of emotion arise from disturbances of cognition. This is often presented as a modern-day form of the Stoic principle that “men are disturbed, not by things, but by the principles and notions which they form concerning things” (Epictetus 2004, section V; Robertson 2010). As discussed earlier, much will turn on the question of whether we are to understand cognition as (a) referencing any meaningful uptake or information processing of “things”; or (b) whether we are to understand it as referring instead to cognitive items such as acts of interpretation, the application or formation of “principles and notions”, or to those of our attitudes which are most aptly described as “underlying beliefs.” I shall return to (a) in the conclusion; in this section I shall consider this more restricted latter sense (b) of cognition, and ask whether emotional disturbance is aptly theorized as consequent upon it.
Within the CBTs it is the “ABC” model of emotional disorder, imported from Albert Ellis’s Rational Emotive Behaviour Therapy (Ellis 1991; Wilson and Branch 2005), that most explicitly embodies the notion of cognitions as thoughts or beliefs. “A” here stands for the activating events or situations experienced; “Bs” are the subject’s beliefs about or interpretations of these events; “Cs” are the emotional or behavioral consequences. The ABC model accordingly construes the cognitive state of belief as a determining intermediary between experience and emotional response. The viability of conceiving of the “C” in CBT along the lines of the “B” in ABC has been questioned by CBT theorists on both theoretical and empirical grounds (Kohlenberg and Tsai 1994; Teasdale 1997; Rachman 1997). One rather obvious empirical objection to the Stoic claim that we are affected not by things but by the notions we form regarding them comes from medical and conditioning approaches to anxiety: I may be made mildly anxious directly by very strong coffee (even if I mistakenly believe it to be decaffeinated); I may be directly emotionally affected by hormonal changes or head injuries; I may feel compelled to perform compulsions to reduce obsessional anxiety even whilst knowing there is no rational basis for this; and I can be afraid of stimuli, such as spiders or birds, which I believe to be perfectly safe (Whiting 2007, p. 239)—hence the continued relevance of simple behavioral therapies for certain phobic or compulsive behaviors. Another objection comes from the observation that, whilst certain thoughts and beliefs may be depressogenic or anxiogenic, this may only occur for people who are already caught up in a particular frame of mind, perhaps driven by certain unconscious desires and emotions (Whiting 2006, p. 241). This, in fact, is an objection anticipated by a key founder of CBT, Aaron Beck, who claimed only to identify the maintaining factors, and not the underlying causes, of depression (Beck et al. 1979; Clark and Steer 1996). The question to be answered in this chapter, however, asks whether there are distinctly philosophical grounds for questioning the idea that disturbed emotions are typically driven by beliefs. My claim is that this idea can sometimes be seen to arise from a misunderstanding, inscribed in certain CBT theories, of what it means to be an emotionally reactive human subject.

A standard cognitive model of depression has it that aversive early experience can lead to the formation of dysfunctional assumptions which, when activated by a critical incident, lead to cycles of negative automatic thoughts and other behavioral, motivational, affective, cognitive, and somatic symptoms (Fennell 1989, p. 171). These assumptions or beliefs are said to take such forms as “If someone thinks badly of me, I cannot be happy”, or “I must do well at everything I undertake”, “I am inferior as a person”, “My worth depends on what other people think of me” (Fennell 1989, pp. 171–8). Therapy then consists in strategies such as teaching the patient to become aware of depressing thoughts as they occur, or undertaking tasks to test the truth of fixed negative beliefs (Williams 1997, p. 265). The first thing to be said about such propositions is that, whilst the therapist probing for core beliefs may well elicit them, they are often far more naturally taken to express a patient’s feelings. McEachrane (2009 p. 92) quotes Ellis (1994, pp. 32–33) attempting to convince a patient of the opposite of this:

“I know I’m doing better of course, and I’m sure it’s because of what’s gone on here in these sessions. And I’m pleased and grateful to you. But I still feel basically the same way—that there’s something really rotten about me, something I can’t do anything about, and that the others are able to see. And I don’t know what to do about this feeling.”

“But this ‘feeling’, as you call it, is largely your belief—do you see that?”
“How can my feeling be a belief? I really—uh—feel it. That's all I can describe it as, a feeling?”

“Yes, but you feel it because you believe it. If you believed, for example, really believed you were a fine person, in spite of all the mistakes you have made and may still make in life, and in spite of anyone else, such as your parents, thinking that you were not so fine; if you really believed this, would you then feel fundamentally rotten?”

“Oh, Hmm. No, I guess you’re right; I guess I then wouldn’t feel that way.”

However the logical error here seems to be Ellis’s rather than his patient’s. Ellis appears to be assuming that, since I am unlikely to feel rotten about myself if I truly believe I am a good and fine person, then it must be the case that my feeling rotten about myself is a product of my believing myself to be rotten. But what, we might ask, if we start with the idea that beliefs are often products of feelings rather than vice versa—might we not expect to arrive at the same state of affairs? That is to say: If I felt bad about myself then may I not, as a consequence, be disposed to form negative beliefs about myself? And then again isn’t it also true that sometimes we feel that there is something rotten about us without forming any associated belief—just as, say, a phobic person may find herself feeling that spiders or birds are dangerous without believing anything of the sort? It may be hard to imagine truly believing that one is a fine person whilst feeling one is rotten, but this may reflect nothing more than the powerfully constraining impact of feeling on believing.

One way in which a CBT theorist might try to salvage the idea that it is maladaptive beliefs that drive emotional disturbance could be as follows. Beliefs, they may say, possess intentionality and carry mental content (they are about this or that). Feelings, on the other hand, are merely “positive” or “negative” (“hot” or “cold”) emotions, sensations, or states of being. Feelings therefore need to be driven by beliefs before they can properly be said to be about anything or to express any kind of understanding. The reply to this is that it is hard to understand why we should accept such an impoverished conception of our feelings as merely aversive or hedonic sensations. For feelings often precisely are about something: I feel furious that he has wasted my time; she is sad about his departure. And since our feelings are very naturally taken to be already about states of affairs, it is hard to see what help they need from intermediary beliefs. Taking “cognitive” in sense (a) from earlier—as referring to any meaningful uptake of things—emotions and feelings can be understood as already cognitive, i.e., as constituted by cognition rather than obtaining in any kind of simple or complex causal relation with separable non-affective cognitive states (cf. Lacewing 2004). Helping the patient understand that their unwanted feelings are not simply free-floating meaningless sensations, but have a meaningful content which situates them in an intentional context, is in fact, and of course, the work of pretty much all non-behavioral therapy.

To consider further the relationship between emotion and belief: imagine that I always dreamed of being a psychologist but then fail my exams and become morose. This emotional state of mine worsens and I start to believe that I cannot be happy unless I am a psychologist. A CBT therapist who becomes wedded to the idea that intermediary beliefs or adherence to rules underpins emotional distress may propose that it is precisely this belief of mine that is
driving my despondency. It is indeed surely possible that such additional beliefs may serve
to maintain or worsen depression. What isn't clear though is why much of my sadness may
not be thought of as arising simply because I have not realized my life's ambition. What is
significant here is not an intervening belief about what I need to be happy, but more gener-
ally how I see my life, my outlook on, felt sense about, or evaluative perception of, my self, my
situation, and my future (cf. McEachrane 2009, pp. 94–95). In a depressive frame of mind I
may arrive at beliefs such as "I cannot be happy unless I am a psychologist," but such beliefs
are, I suggest, in the run of things more likely to result from my sense of my situation rather
than vice versa.

What, then, are we to make of the not infrequent CBT insistence that beliefs have a key
pathogenic role to play in depression and other emotional disorders? My own belief is that,
whilst certain CBT therapists may sometimes be driven by an overly intellectual (belief-
or rule-mediated) conception of our emotional perception of our worlds, what has more often
happened is that they have stretched the meaning of “belief” to fit the phenomenology and
save the theory. For example, I once asked my CBT supervisor what she meant by her talk
of her patients’ “beliefs,” and she replied that you knew that what you had to do with was a
belief when a patient cried in articulating it. This clearly has little to do with the normal
notion of belief, given that the infinity of our everyday beliefs (that it is not raining, that
most people have legs, that the next symbol is an ellipsis . . . ) hardly provokes an infinity of
tears. What instead we have to do with here may rather be the result of a common enough
situation in psychiatry and psychotherapy generally: that the practical meanings of clinical
terms become constituted more by their actual clinical uses than by associated theory,
where such clinical uses are themselves a product of a distinctive lexicon being forced to
carry whatever are the communicative and therapeutic burdens of the clinical encounter.
The result is an implicit and unacknowledged argot, the non-standard meanings of which
are implicitly known to the practitioners in question.

This explanation should not, of course, be mistaken for exculpation. CBT theorists have
as much a professional duty to communicate clearly and accurately as all psychologists, and
misusing ordinary terms (“belief,” “thought”) may be no less unhelpful than deploying the
language of science when it adds nothing to what is better stated in everyday language. And
the fact remains that we do ordinarily distinguish between beliefs and feelings, and use the
distinction to mark situations in which, say, we feel scared of a cat even whilst believing
it to be safe. The person who is scared of the cat may have all sorts of inarticulate feelings
about cats which could helpfully be made more explicit in therapy, but it isn’t clear what—
other than saving an exclusively “cognitive” (in sense (b)) model—is to be gained from
describing these as implicit beliefs. Furthermore cognitive neuroscience has now provided
several models outlining different streams of information processing for propositional and
embodied understandings (cf. Power and Dalgleish 1999; Teasdale 1997). Rather than think
of the emotionally significant meanings which arise as a patient experiences their world
as encoded in propositionally structured interpretations or beliefs, such theorists instead
invite us to consider the prime significance of non-linguistic and somatosensory forms of

2 For an example of the latter, Mollon (2007, p. 13) complains of the scientistic rhetoric of the CBT
literature, where “facing your fears is called ‘exposure,’ refraining from an activity is called ‘response
prevention,’ learning to relax is called ‘stress inoculation,’ and revising your thoughts is called ‘cognitive
restructuring.”
meaning. The clinical upshot is that therapy must aim at helping a patient change their way of being in the world through changing such non-belief-based meanings.

In this section I have added to such empirical critique by providing philosophical reasons for thinking that an “ABC”-style model falsifies what it actually means to be an emotionally reactive, meaning sensitive, human subject. If, when a CBT theorist tells us that emotional reactions are shaped by our “cognitions” they mean by “cognitions” something like our beliefs, and if by our “beliefs” they mean something other than how we feel about or see our situations, then the CBT theorist appears to run the risk of radically misconstruing the foundations of our affectivity by putting the cart of our belief before the already-meaningful horse of our emotion.

**On Articulating our Assumptions**

In this section I inquire into the relation between (a) a patient’s verbal articulations in therapy of their underlying self-understandings, and (b) the nature of such underlying self-understandings. Typical CBT models expound something like the following: (a) that the first, rather preliminary, step of cognitive therapy is to help the patient clearly identify their emotionally problematic core beliefs, rules and assumptions. And (b) that the second task is to encourage them to quasi-scientifically test out these assumptions, either through rational engagement (e.g., Kuehlwein 2002) leading to what is sometimes called “cognitive restructuring,” or more practically through “behavioral experiments” (e.g., Bennett-Levy 2004).

In what follows I will challenge such a conception of what is happening in therapy when assumptions are articulated, questioned, or put to the test. My aim is not to question the CBT practice, but rather to question the conception of self-understanding that it encourages for the theorist: that articulation of one’s own self-understanding in therapy is a rather preliminary stage occurring prior to the genuine transformations that can occur only when such self-understandings are challenged and healthier alternatives considered and embraced. The criticism offered here will be that those CBTs operating according to the above mentioned description risk underlaying both (a) the intrinsically transformative nature of the acknowledgement of one’s deepest fears, and (b) the way in which verbal, behavioral, and experiential therapeutic techniques function not merely to test hypotheses, but rather to effect a more fundamental change in the form of the patient’s thought. A more fundamental change, that is, that first of all enables it to become penetrable by the light that experience and rational enquiry can shine into the fearful recesses of the mind.

To start to explore what this means, consider first what happens when an underlying dysfunctional assumption gets put into words, perhaps through the application of the “downward arrow” technique. (This technique has the therapist repeatedly ask of, say, a patient voicing self-critical thoughts, questions such as “and if that were true, what would that say about you?” until a definitive underlying negative fear is articulated.) Here is how Fennell (1989) articulates the therapeutic strategy:

Rather than challenging the thoughts themselves, the therapist asks: “Supposing that was true, what would that mean about you?” This and similar questions … are repeated until
it is possible to formulate a statement general enough to encompass not only the original problem-situation, but also other situations where the same rule is operating. Once a dysfunctional assumption has been identified, questioning and behavioural experimentation are used to find a new, more moderate and realistic rule. (pp. 204–205)

Fennell’s own assumption is that dysfunctional assumptions can be hard to unearth because, “rather than [being] discrete events occurring in consciousness, they are generalized rules which may never have been formulated in so many words” (1989, p. 204). Why such assumptions should not have been put into words is not considered by her, nor is the emotional experience of the patient who articulates them—which, in my clinical experience at least, typically involves a quite particular mixture of distress and relief. This empirical issue is not, however, our key concern here—which is instead philosophical and concerns what happens to such tacit assumptions when they are voiced. Fennell’s presentation, which is quite standard in the CBT literature, encourages the view that their voicing involves merely their being put into words. But reflection on what is meant by the voicing of our deepest troubling thoughts—reflection that is philosophical in so far as it concerns itself with how the objects of psychological investigation ought to be characterized, rather than itself depending on the results of empirical enquiry—reveals something far richer. The articulation of one’s previously unarticulated fears involves, that is, not merely voicing but also acknowledgement (Finkelstein 1999), a correlative increase in self-understanding, an emancipation which comes from making visible or thinkable those assumptions which otherwise continue to invisibly constrain our meaningful experience and, so long as the resultant fears can be contained until they dissipate, an increased capacity to tolerate reality.

To take up the last of these: the patient who presents under the influence of a dysfunctional assumption which they have not yet articulated to themselves can often be seen to be failing to fully distinguish between fear (or wish) and belief. Whether they always really believe, for example, that others laugh at them as they walk down the street, or whether this is simply what they fearfully imagine or think, is not always entirely clear. An important reason for this is that they can be presenting in a state of mind which somewhat ablates the very distinction between fearing and believing. Their “safety behaviors” (which they perform to prevent what they imagine will be) the experienced realization of their fears—e.g., never meeting the gaze of others for fear of meeting a hostile look, or never failing to grip onto the shopping trolley for fear that they would collapse) similarly keep in place the constitutive lack of clarity of the anxious state. However when the CBT therapist encourages them to test out their fears by surveying the beliefs of others, or by undertaking a behavioral experiment (which may involve seeing if matters really do deteriorate if they don’t perform their safety behaviors), what happens is not merely that the truth of an assumption is evaluated, nor that they simply become more reflectively aware of what they already thought, but that a clear distinction between fear and genuine belief begins to be instated in the mind. A fearful state of mind that demotes reality contact becomes displaced by an empirical hypothesis.

3 The psychodynamic theorist has a ready explanation of such phenomenology: that the articulation of such assumptions involves the patient coming to acknowledge that which has previously been defended against as too emotionally painful.

4 In psychodynamic terms, their thinking is more a matter of “primary” than “secondary process.”
Or perhaps the patient is invited to put a probability figure to their fear: what is the actual likelihood that you will be scoffed at if you show your face in the town center? Ninety percent? Twenty percent? Once again the CBT strategy invites the transformation of what would once have been called a “neurotic” state (in which an avoidance of reality has conspired with and inspired a fusion of fear, belief, and experience—a state which may sometimes be compensated for by structurally similar fusions of wish, belief, and experience) into an empirical hypothesis. Various CBT strategies have the clarifying effect of helping the patient shape up their fearful state so that it can be brought into contact with reality, first expressing it in the necessary bivalent (true/false) or probabilistic form of a genuinely reality-oriented proposition. They have this effect, that is, despite often understanding themselves as merely being in the business of promoting the voicing and then the testing of beliefs. What is significant, however, is not so much the disproving, but the emotional shift of rendering such fears provable or disprovable: reality testing (i.e., becoming able to distinguish between thought and fantasy) is a precondition of, and not a synonym for or consequence of, hypothesis testing.

In this section I have suggested that whether we are dealing with the verbal articulation, or with the testing, of a patient’s deepest dysfunctional assumptions, what is principally significant is a change in the form of the patient’s fears. For what renders such assumptions dysfunctional is not simply their content but also their neurotic form—their insulation from empirical testing and rational thought, their fusion with wish or fear. What I am suggesting is that mere reflection on the phenomenology of the patient’s experience in therapy shows that it is a change in such form that ultimately makes for the possibility of a change in unhelpful content. The CBT therapist’s armory of tools for unearthing and engaging with such assumptions has the effect of shining the light of reason into the darker recesses of the patient’s mind, encouraging their thought to be governed not by fearful fancy but rather by (what psychoanalysts would call) “the reality principle.” In the process their thought becomes articulated—which is to say, not merely voiced but structured—so that it now more clearly embeds a distinction between appearance and reality, is less insulated from reality testing, and is less a hostage of their deepest fears and wishes.

A much debated concern in the scientific CBT literature is whether the benefits of CBT are mediated by cognitive changes in the patient, whether such benefits may occur before cognitive techniques have even been applied, and whether cognitive techniques really add much to behavioral therapy (Longmore and Worrell 2007). One way of putting the philosophical point I am making here is that expressive and behavioral techniques can already be seen to be cognitive—insofar as they involve a change in the form of the patient’s fearful or depressive preoccupations (cf. Carey and Mansell 2009). The various techniques of CBT serve to free the patient from those inchoate fears which have hitherto not found adequate, clearly delineated, expression. Making the fears less inchoate and more thinkable risks making them feel more real, and the avoidance of this scary possibility often seems to have been a significant part of the patient’s difficulties. Therapy, however, provides an emotionally “containing” environment in which the therapist’s clarity regarding, say, the relatively benign and possibly even nourishing nature of social reality—by contrast with what the patient fearfully imagines to be the case—can become internalized by the patient. It is for this reason that a good therapeutic relationship must not merely be collaborative (Leahy 2008) but also be containing, and an effective CBT therapist’s containing manner shows itself in such trans-model common factors as her confidence in her methods, her clarity, reality orientation, compassion, and warmth.
To summarize the four objections: The first proposed that various CBT formulations conflate the formal relations between different aspects of the same state with causal relations between discrete inner states. The second had it that some CBT models construe emotionally laden perspectives too much as occurrent inner processes, and too little as a subject’s attitudes. The third considered that such attitudes can sometimes be misdescribed in CBT models as beliefs—when what we really have to do with here are feelings that are themselves already about the events and situations that arouse them. The fourth argued that CBT models can underplay the significance of changes in the form of (a subject’s ownership of) such attitudes when they focus instead on changing their content. In these concluding paragraphs I first consider whether a unifying diagnosis can be given of these diverse tendencies to distort what it means to be a living human subject before going on to ask what can be done about it.

My suggestion here is that what unifies these disturbances of vision is their being expressive of what could be called an “alienated” conception of human subjectivity. Imagine that the self were no longer an expressive bodily being located immanently by its feelings within a meaningful intersubjective world—but had retreated inward, away from the world, the body, and even the mind, becoming instead a disengaged inner spectator trying to make sense of the intersubjective world from without. Perhaps such a self is a prototype of a scientist-observer who is in the business of trying to control and predict the world by constructing inner representations or interpretations of it. The effects of such a retreated conception of the self would be several. The mind and body, having been denuded of subjectivity, will now more naturally appear as a domain of merely causally (rather than meaningfully) inter-related, objectified states and processes. Mind becomes a domain of inner states and processes rather than a matter of how we are embedded in our intentional worlds. The self will no longer speak from its attitudes, but will rather be reduced to speaking about them. Our minds and bodies become not so much what we could call the flesh of the self, but instead show up as domains of inner and outer processes that require to be controlled from within. Therapy, accordingly, would get theorized as no longer in the business of self-transformation, but instead becomes a technology—grounded in what will appear to be, amongst the therapies, uniquely scientific causal models of inter-related inner processes—for helping us manage our feelings. Our relationships too will be reduced to a merely external form—which is to say, that they will become theorized not as constitutive of who we are, but rather seen merely to causally connect us to that which is essentially other. So too the being of the patient will no longer be thought of as partly immanent in the emotional flux of the therapeutic relationship; that relationship will instead risk being reduced in the theorist’s vision to something which is merely a collaboration between distinct relata.

---

5 George Kelly, psychologist of “personal constructs,” proposed just such a scientist-observer conception of the human subject (Kelly 1955). His work significantly influenced the early CBT theorist Albert Ellis.
Precisely such a mechanistic and self-alienated conception of the mind can, it seems to me, be what we often find in the patient who presents wanting to know how to better “control their anxiety,” “change their thoughts” or “manage their minds.” Such a patient has, we could say, become alienated from their own inner life which, accordingly, is seen as an independent domain painfully afflicting them and requiring management or excision. Now it bears recollection that in this chapter I have not been concerned to critique the CBTs en bloc, nor on the whole to question the cogency of CBT therapy, but instead to scrutinize some ways in which some CBT theories may be inflected in ways that go against what it means to be an emotionally alive human subject. The risk for the CBTs that I have identified may now be described as one of being encouraged by an infelicitous theoretical model of joining the earlier-described patient in such a de-subjectivized vision of their psyche. But by helping the patient to give articulate structure to his fears, to think, to be nourished by reality contact, and to distinguish fearful fantasy from genuinely representational belief, the CBT practitioner can be understood as doing far more than, say, helping the patient to test out his hypotheses: she is helping to restore her patient’s subjectivity.

The final question is: What can the CBTs do about all of this? The answer, I think, is already possessed by many CBT theorists, and has already been mentioned. It is to considerably broaden the conception of what counts as “cognitive”—to include not only a patient’s thoughts, interpretations, and beliefs, but also any aspect of her meaningful engagement with her world. Cognition now becomes the name for an inherent property of perception, emotion, action readiness, interaction, etc., and we can drop the intellectualist consideration that perception, say, must be supplemented by acts of interpretation before it can be said to disclose a meaningful world for us. An example of this broadening of the domain of the cognitive was provided for me at a recent teaching session on treating childhood trauma by a well-known CBT author at a well-known UK cognitive therapy training center. A model of self-disturbances in trauma was outlined that stressed the priority of emotion, the central role of emotional and interpersonal avoidance (read: defenses) in the maintenance of psychopathology, and the therapist’s use of self in therapy. The presenter then commented that this approach could still be counted as “cognitive” since it concerned itself with the meaning of the patient’s experience. The obvious question to ask at this point, however, is: What therapy would now not be properly called “cognitive”?

At present it is fair to say that various features tend to characterize the therapeutic practice of those who consider themselves “cognitive” in orientation. These include a readiness to use information-sharing and instruction (“psycho-education”), an active and structured encouragement of within- and between-session behavior change, and a focus on collaboration rather than containment in the therapeutic relationship. My suspicion is that, as for other therapies too, being “cognitive” means, in practice, to be identified with the interests, blindspots, and those fairly subtle helpful and unhelpful habits of mind and action tacitly
embedded in particular psychological communities. The clinician’s temptation may often be to attempt to justify their unique orientation with reference to the empirical or theoretical credentials of their model. But if we read “cognitive” in its most theoretically plausible sense—i.e., as referring generally to structures and processes of meaning rather than merely to thoughts and beliefs—the impossibility of distinguishing CBT from, say, psychodynamic therapy by referring to the former’s focus on cognition should now be clear.

ACKNOWLEDGEMENTS

This chapter has benefitted considerably from the thoughtful comments of Louise Braddock, Michael Lacewing, and Rebecca Murphy.

REFERENCES


