Synthese (2024) 203:124
https://doi.org/10.1007/s11229-024-04523-0

ORIGINAL RESEARCH

Belief revision in psychotherapy

J. P. Grodniewicz[1]

Received: 1 September 2023 / Accepted: 8 February 2024

Abstract
According to the cognitive model of psychopathology, maladaptive beliefs about oneself, others, and the world are the main factors contributing to the development and persistence of various forms of mental suffering. Therefore, the key therapeutic process of Cognitive Behavioral Therapy (CBT)—a therapeutic approach rooted in the cognitive model—is cognitive restructuring, i.e., a process of revision of such maladaptive beliefs. In this paper, I examine the philosophical assumptions underlying CBT and offer theoretical reasons to think that the effectiveness of belief revision in psychotherapy is very limited. This is the case, I argue, because the cognitive model wrongly assumes that our body of beliefs is unified, while it is in fact fragmented.

Keywords Belief · Belief revision · Fragmentation of beliefs · The cognitive model of psychopathology · Cognitive restructuring · Cognitive behavioral therapy

1 Introduction

What we think and believe often contributes to our mental suffering. For example, if a person believes that they are unlikeable, they may withdraw from social life and experience a depressed mood as a result of their way of thinking and behaving (Beck, 1967; Ingram et al., 2011). If a person believes that they must be near-perfect in all their endeavors, they may be more likely to suffer from anxiety and panic disorder (Handley et al., 2014). Patients diagnosed with borderline personality disorder often hold beliefs such as: "People will take advantage of me if I give them a chance," "I

✉ J. P. Grodniewicz
j.grodniewicz@gmail.com

[1] Jagiellonian University, Copernicus Center for Interdisciplinary Studies, Kopernika 27, room 26, 31-501 Kraków, Poland

Published online: 11 April 2024

can't cope as other people can," or "People will pay attention only if I act in extreme ways" (Bhar et al., 2008, p. 171). The list goes on.

But what exactly is the role of beliefs in the emergence and persistence of various psychopathologies? According to the so-called "cognitive model:" "self-relevant thoughts, evaluations, and beliefs are key contributors to the development and persistence of psychopathological states" (Clark, 2013, p. 23). It is these beliefs, thoughts, and evaluations that "mediate maladaptive behavioral and emotional responding" (O'Donohue & Ferguson, 2015, p. 19).

If we assume that certain beliefs are causally responsible for the emergence and persistence of psychopathology, it is natural to think that their revision should be a crucial component of treatment or therapy. This *prima facie* compelling inference is central to Cognitive Behavioral Therapy (CBT) (e.g., Ellis, 1962; Beck et al., 1979; Beck, 2011; Hofmann et al., 2013)—a therapeutic approach rooted in the cognitive model of psychopathology.[1] Thus, one of the core therapeutic processes of CBT is *cognitive restructuring*, i.e., the process of modification or revision of the dysfunctional beliefs responsible for our mental suffering.

Despite the fact that CBT (i) is currently one of the most (if not *the* most) popular psychotherapeutic approaches (Cook et al., 2010), (ii) has the status of the evidence-based practice for the broadest set of diagnoses (Society of Clinical Psychology, 2023), and (iii) is often advertised as the "gold standard of psychotherapy" (David et al., 2018), there is no simple answer to the question: "What is responsible for the overall efficacy of CBT?" Moreover, a growing number of authors argue that whatever is responsible for its efficacy, it is not, or at least not primarily, cognitive restructuring (e.g., Jacobson et al., 1996; Hayes, 2004; Dimidjian et al., 2006; Kazdin, 2007; Longmore & Worrell, 2007; Arch et al., 2012; but see Hofmann, 2008; Lorenzo-Luaces et al., 2015).

The goal of this paper is to examine philosophical assumptions underlying the cognitive model of psychopathology and offer theoretical reasons why we should expect the effectiveness of cognitive restructuring to be very limited. I will argue that a central (even if mostly implicit) tenet of the cognitive model is *unificationism*, according to which beliefs of a given subject form a single, consistent set, and are accessible and available at any given time and in any given context.[2] Only if unificationism were true, could we expect cognitive restructuring to be effective.[3] However, we have good reasons to assume that unificationism offers an overly simplistic and

[1] Unless stated otherwise, when I speak about Cognitive Behavioral Therapy I mean the so-called "second-wave" or "classical" CBT emerging directly from Aaron Beck's work on Cognitive Therapy (CT) (Beck, 1967; Beck et al., 1979) and Albert Ellis's work on the Rational Emotive Behavior Therapy (REBT) (Ellis, 1962). In Sect. 6, I will briefly contrast it with the "third-wave" CBT—a set of newer therapeutic approaches, which focus on mindfulness, psychological flexibility, and metacognitive skills (Hayes, 2004; Forman & Herbert, 2009; Hayes & Hofmann, 2017).

[2] This characterization of unificationism, which I discuss in detail in Sect. 4, is based on characterizations offered by Kindermann and Onofri (2021), Egan (2021), and Bendaña and Mandelbaum (2021). It is not meant to compete with these characterizations for the title of the most "accurate" or "exhaustive." It simply focuses on the aspects of unificationism crucial for the present discussion.

[3] The truth of unificationism is necessary for the effectiveness of cognitive restructuring but it is not sufficient. Even if unificationism were true, there could be other reasons why cognitive restructuring is not effective. See the discussion below.

idealized picture of the organization of our beliefs. Our belief systems are fragmented rather than unified.

I want to make clear from the outset that it is *not* my goal in the present paper to argue that CBT is an ineffective psychotherapeutic approach. What I am discussing here is not the question of *whether* CBT works but *why* or *how* it works. In particular, does it work due to the process of belief revision?

The paper proceeds as follows. In Sects. 2 and 3, I spell out the details of the cognitive model and offer a fuller characterization of cognitive restructuring. In Sect. 4, I take a closer look at the role unificationism plays in the cognitive model. In Sect. 5, I argue for the limited effectiveness of cognitive restructuring by appeal to fragmentationism about belief. Finally, in Sect. 6, I briefly discuss some further consequences of my arguments for psychotherapeutic practice.

2 The cognitive model of psychopathology

According to the cognitive model of psychopathology, the main factor responsible for the emergence and persistence of psychological suffering and many mental disorders is maladaptive[4] thoughts, beliefs, and information-processing patterns, often labeled jointly as "cognitions" (see, e.g., Beck et al., 1979; Padesky, 1994; Young, 1990; Brewin, 2006; Clark, 2013; A. T. Beck & Haigh, 2014; Newman, 2015; Arntz, 2018). It is further assumed that these cognitions must be modified to achieve therapeutic change, typically identified with a reduction of symptoms associated with a given psychopathology.

The proponents of the cognitive model traditionally distinguish three main types of cognitions: automatic thoughts, intermediate beliefs, and core beliefs. Let us look at them one by one.

Automatic thoughts are often characterized as "spontaneous thoughts that arise without effort, and that tend to be taken at face value unless a person reflects upon them" (Newman, 2015, p. 119). Consider Tom speaking with his boss. The boss complains about the poor job Tom has done during the pitch meeting with their new client. She has an angry face and slightly raises her voice. In this very moment, Tom thinks to himself: "She hates me." Tom is not voluntarily engaged in the reasoning; plausibly, not even aware of the thought process which led to the emergence of this thought. He *does not* start by asking himself, e.g., "Why is she so angry?", to then go through some further steps consciously, e.g., "I didn't do worse than Sally, and she didn't scream at her," "She was angry just the same last time, even though we cut the

[4] It is quite a challenge to capture the precise meaning of "maladaptive" in the cognitive model. Even a brief overview of the CBT literature suffices to realize that the term is being used as synonymous with a plethora of related terms, e.g., "negative", "irrational", "abnormal", "erroneous", "unrealistic", "dysfunctional", "biased", "inflexible", "illogical", "distorted", "invalid", "faulty", but rarely carefully defined (for a recent philosophical appraisal of this problem see Ratnayake, 2021). Given that my argument against cognitive restructuring presents problems for the idea of therapeutic revision of any beliefs, I will leave the systematic discussion of the meaning of "maladaptive" for another occasion.

deal," only to arrive at the conclusion "She hates me." As it were, the thought "She hates me" just pops up in his head.[5]

It is an assumption of the cognitive model that automatic thoughts can affect the way we feel and behave even if we are paying them very little or no attention. Therefore, one of the key skills a client practices during their cognitive therapy is *identification of automatic thoughts*, which consists in redirecting one's attention to these thoughts and becoming mindful of their contents.

But which thoughts come to our mind is not entirely accidental. It is a further assumption of the cognitive model, that our thoughts are closely related to a more stable and permanent type of cognitions, i.e., beliefs (Dobson & Shaw, 1986).[6] Traditionally the proponents of the cognitive model distinguish two types of beliefs: intermediate and core beliefs. *Intermediate beliefs* are conditional "if… then" rules and assumptions internalized by a subject, e.g., "I am nothing if a person I love doesn't love me," "If you cannot do something well, there is little point in doing it at all.," "If someone disagrees with me, it probably means they don't like me," etc. *Core beliefs*, in turn, are supposed to be the deepest and most general convictions we hold about the world, other people, and ourselves (Beck, 1967; Arntz, 2018). For example, many depressed patients turn out to hold such core beliefs as "I am unlovable," "I'm incompetent," "Nothing good ever happens to me," "People my age have their life figured out," etc.[7]

According to the cognitive model, most maladaptive automatic thoughts are connected to core beliefs via intermediate assumptions. While automatic thoughts are consequences of core beliefs, in the order of discovery, one usually recognizes their core beliefs by inferring them (often with the help of a therapist) from their automatic thoughts. For example, during therapy, Tom might be encouraged to focus on his automatic thought: "My boss hates me." Together with his therapist, he may discover that, even though it was not an element of his conscious reasoning during the squabble after the pitch meeting, he is convinced that if people are angry at him, they probably hate him (intermediate belief). Such a strong assumption might come as a surprise to most of us. But it is not surprising to Tom. After further reflection, he

[5] For the purpose of the current discussion, I will understand automatic thoughts to be involuntary and spontaneous episodes of entertaining a given content. Obviously, thought contents come in different formats, e.g., linguistic, pictorial, auditory, etc. Moreover, we can have multiple thoughts at once. The number of things we are able to think about at one time is plausibly restricted only by the capacity of our working memory (Teasdale et al., 1995; Baddeley, 2007).

[6] Here, I follow the cognitive model in assuming that intermediate and core beliefs are indeed beliefs, and not some other attitudes, e.g., acceptances, aliefs, imaginings, etc. I do not think that this matter is uncontroversial. In fact, it is at least as controversial as the matter of the doxastic status of delusions, hotly debated in philosophy over the last twenty years (Bortolotti, 2022). Nevertheless, this matter requires a separate discussion—one that I cannot offer in the present paper. Another underexplored yet important topic which I have to leave aside here is the way in which the cognitive model conceptualizes the relationship between beliefs and emotions, especially if we assume that emotions have an evaluative, belief-like component (for a discussion of this topic see, e.g., Lacewing, 2004; Whiting, 2006; McEachrane, 2009; Gipps, 2013).

[7] Some tools used to assess the intermediate and core beliefs held by clients are Dysfunctional Attitude Scale (DAS) (Weissman, 1979) and Personality Belief Questionnaire (PBQ) (Fournier et al., 2012). Many of the examples I use come from these questionnaires.

realizes that—for as long as he remembers—he believed that most people he knows dislike or even hate him (core belief).

Recognition of this and similar patterns of maladaptive cognitions guides Tom's therapist to the working hypothesis that Tom struggles with some form of social anxiety (Schulz et al., 2008). She decides that this is one of the things that they should address during therapy.

3 Cognitive restructuring

At this point, Tom's therapist faces an important decision: which therapeutic methods and techniques should she apply to help her client? If she is trained in CBT, it is very likely that one of the main techniques in her therapeutic toolbox will be cognitive restructuring (e.g., Beck et al., 1979; Dobson & Dozois, 2010; A. T. Beck & Haigh, 2014; Newman, 2015; Arntz, 2018). Here is a sample of definitions of cognitive restructuring offered by CBT theorists. Cognitive restructuring is:

> … [the process of] modifying beliefs through the review or production of evidence that contradicts negative or maladaptive conclusions drawn by a client. (Padesky, 1994, p. 268)
> … designed to directly modify specific thought patterns or beliefs… (O'Donohue & Ferguson, 2015, p. 19).
> … [focuses] on the exploration, evaluation, and substitution of the maladaptive thoughts, appraisals, and beliefs that maintain psychological disturbance. (Clark, 2013, p. 24)

Cognitive restructuring is not a single technique whose application looks the same in all contexts. It can be achieved by a broad set of techniques, such as evidence gathering, Socratic questioning, hypothesis testing, analysis of consequences, positivity reorientation, and more (e.g., Beck, 2011; Clark, 2013; Newman, 2015). Nevertheless, what is common to all cognitive restructuring methods, and where they differ from, e.g., exposure or behavioral activation, is that they are supposed to achieve a cognitive change by *directly* targeting maladaptive cognitions (Barber & DeRubeis, 1989; Brewin, 2006; Hofmann et al., 2013). Hofmann and Asmundson (2008) spell this out as follows:

> … patients in CBT are encouraged to generate hypotheses based on their beliefs (theories) about the world, themselves, and their future… *By falsifying these hypotheses, patients are then forced to revise their belief system*, reducing the emotional distress. (Hofmann & Asmundson, 2008, p. 12, emphasis mine)

Despite an often-loose way of speaking about it in CBT textbooks, restructuring is targeted primarily towards beliefs, and it is expected to influence automatic thoughts only indirectly. After all, thoughts are episodes with a relatively short lifespan. When the thought "My boss hates me" pops up in Tom's head, he may focus on it and start investigating its origin, e.g., the underlying belief. He may even start entertaining

another thought, such as "She does not hate me, she's only upset and stressed out" and try to suppress the original one.[8] But it makes little sense to say that in such a case the original thought is *revised* or *modified*. Therefore, when discussing cognitive restructuring, I will focus on how it affects beliefs, for the most part leaving the topic of automatic thoughts aside.

Despite the fact that cognitive restructuring is central to modern CBT, its effectiveness remains the subject of heated discussion. Even under the assumption that CBT *is* an effective therapeutic approach, the question remains whether it is thanks to cognitive restructuring, i.e., whether cognitive restructuring is the mechanism of positive therapeutic change (e.g., Hollon et al., 1993; Jacobson et al., 1996; Hayes, 2004; Dimidjian et al., 2006; Longmore & Worrell, 2007). While some authors deny it rather forcefully (Kazdin, 2007), others find such a denial unwarranted (Hofmann, 2008). Yet others are too cautious to cast a decisive verdict. For example, Hollon and Beck (1994, p. 458) admit that "It is not clear whether these interventions work, when they work, by virtue of changing beliefs or thinking, as specified by theory." Lorenzo-Luaces and his colleagues (2015) suggest that the reason why the results regarding the effectiveness of cognitive restructuring are ambiguous is that it is very difficult to experimentally establish whether cognitive change mediates the symptom change and not the other way round. They admit that "evidence consistent with this hypothesis comes largely from studies in which reverse causality has not been ruled out" (Lorenzo-Luaces et al., 2015, p. 8).

An adequate assessment of the efficacy of cognitive restructuring must consider a plethora of factors I have not even mentioned, from the type of diagnosis and severity of clients' condition to the question whether we focus on the short- or long-term results. However, I am neither willing nor able to offer yet another meta-analysis of the methods of cognitive therapy. The seed of doubt present in the existent research on this topic is all I need to motivate my discussion in the rest of this paper. I will argue that the very idea of cognitive restructuring is based on strong philosophical assumptions about the organization of beliefs—assumptions, moreover, which we have good reasons to reject. Therefore, if further research will eventually lead us to a negative answer to the empirical question about the efficacy of cognitive restructuring, we should not be surprised. Moreover, those among us whose everyday practice potentially involves cognitive restructuring (e.g., therapists, counselors, and other mental health professionals) may want to reconsider their reliance on this process right away.

Before I move on, I would like to address one possible objection to what I presented in this and the previous section. Someone may protest that the picture of CBT offered here is outdated and that, by discussing it, I am attacking a strawman. CBT used to focus on beliefs and their restructuring, but these times are long gone. Instead of beliefs, modern CBT is interested in *modes*, defined as "networks of cognitive,

[8] It is now well established that suppression is not an effective method of dealing with negative thoughts. In fact, it is proved to be counterproductive—it usually makes the thought one tries to suppress more vivid and accessible (Wenzlaff & Wegner, 2000).

affective, motivational, and behavioral components designed to deal with specific demands" (A. T. Beck & Haigh, 2014, p. 2).[9]

This objection is misguided. While it is true that modes are an important component of modern CBT theory, it is false that the talk about beliefs has been abandoned or that techniques of belief revision are no more among the principal tools of CBT. In a recent summary of the advances in cognitive theory and therapy, Beck and Haigh (2014) claim explicitly that "clinical disorders, which share common underlying processes, can be *differentiated by the nature of their dysfunctional beliefs*" (A. T. Beck & Haigh, 2014, p. 13, emphasis mine), while the goal of a therapeutic process is to "make a *durable impact on disorder-specific dysfunctional beliefs* by relying on a combination of *cognitive restructuring*, attentional modification, and behavioral interventions" (A. T. Beck & Haigh, 2014, p. 18, emphasis mine). This is not an accidental slip of the pen. The same sentiment was there since the idea of modes entered the theoretical stage of CBT. In his (1996) paper "Beyond belief: A theory of modes, personality, and psychopathology," Beck asks: "How does a therapist achieve a more durable modification of the content or structure of a mode?" and answers: "It seems that for such a lasting change to occur, it is necessary to achieve a substantial change in the underlying absolute and conditional rules that shape the individual's interpretations" (p. 16). He concludes the same paper by claiming that "[i]n clinical practice, it is important to concentrate on modifying the dysfunctional beliefs as well as inculcating cognitive skills." (Beck, 1996, p. 19). Thus, even though CBT has evolved significantly over the last 60 years, it would be a mistake to assume that it has abandoned its central idea of belief revision in psychotherapy.

4 The philosophical assumptions of the cognitive model

Up to this point, I have been using the term "belief" rather loosely. It is time to sharpen our philosophical tools. I will now argue that the proponents of the cognitive model almost univocally (even if, for the most part, tacitly) assume *representationalism* regarding the nature, *unificationism* regarding the organization, and *Quineanism* regarding the revision of beliefs. I will borrow the definition of representationalism from Quilty-Dunn and Mandelbaum (2018, p. 2354):

[9] Yet another notion important for modern CBT is this of "schema." Unfortunately, it is difficult to derive a univocal definition of schema even from Beck's own body of work (not to mention other authors in the CBT tradition). Notably, at the very beginning, Beck used to focus only on "cognitive schemas," defined as structures "for screening, coding, and evaluating the stimuli that impinge on the organism" (Beck, 1967, p. 283). At that point, he often used the "schema" and "core belief" interchangeably. Later on, however, e.g., in (Beck, 1996), he spoke about cognitive, orienting, affective, motivational, and behavioral schemas as jointly constituting different modes. Most recently (A. T. Beck & Haigh, 2014), he came back to defining schemas as *cognitive* structures: "Schemas: complex cognitive structures that process stimuli, provide meaning, and activate related psychobiological systems" (p. 2) and beliefs as "representations of abstractions of schema content" (p. 2). Finally, what may add to overall confusion is that today, the talk about schemas is most common in the context of Schema Therapy (ST), according to which schemas comprise "memories, emotions, cognitions, and bodily sensations" (Martin & Young, 2010, p. 318). Thus, ST schemas are more similar to Beckian modes than Beckian schemas.

Representationalism: To have a belief is to stand in a particular relation to a mental representation.

Arguably, representationalism—whose famous proponents include such figures as: Fodor (1987), Dretske (1988), Millikan (1984), Burge (2010), and many more—dominates contemporary cognitive sciences. It is not surprising, given that representationalism about belief perfectly complements the Representational Theory of Mind (RTM)—the apple of the eye of the cognitive revolution during the second half of the 20th century. According to RTM, the best way of explaining cognitive processes is by appealing to transformations and computations performed on mental representations, i.e., information-bearing and truth-evaluable mental objects stored in mind/brain. According to representationalism about belief, in turn, to believe is to be in a particular relation towards one of such representations. What relation? Most generally, to believe that *p* is to take *p* to be the case.

I do not think that the commitment of the cognitive model to representationalism requires an elaborate argument. In the CBT literature, beliefs are constantly referred to as something that subjects *acquire* at one point or another and something they *have* even if they are not fully aware of them. If the treatment is to be effective, patients together with their therapists must "identify" (Hofmann & Asmundson, 2008, p. 3) or "discover and adequately formulate" (Arntz, 2018, p. 342) what is the content of their beliefs. Chris Brewin is fully explicit about representationalism when he introduces the commonalities between different versions of the cognitive model. All these versions "generally assume that previous adversity produces vulnerability in the form of negative representations of the self and the world" (Brewin, 2006, p. 766), which are stored in memory and further contribute to the production of negative thoughts and prolonged negative mood.

Therefore, one way of challenging the cognitive model (and cognitive restructuring in particular) would be through targeting representationalism. The whole idea of modifying or revising specific maladaptive representations, which are the contents of our beliefs, requires there being such representations in the first place. This is part of the critique of the cognitive model developed by contemporary neo-behaviorists (cf. Hayes & Brownstein, 1986; Hayes, 2004) but it is equally available to dispositionalists and eliminativists about beliefs (cf. Schwitzgebel, 2019).[10] However, this is not the critique that I want to pursue in the present paper. For the sake of my argument, I will accept representationalism. What I want to challenge is the second assumption

[10] A related critique of the cognitive model comes from Leder (2017). Leder points out that CBT makes yet another strong assumption: through introspection assisted by techniques such as Socratic questioning, patients are able to accurately recognize specific cognitions causing their mood and behavior. However, given what we know about the limitations of introspection, and our tendency to confabulate to fill the gaps in explanations of our behaviors, this assumption seems at least dubious. Moreover, as was pointed out to me by an anonymous reviewer, one could offer a thorough critique of the cognitive model in general, and cognitive restructuring in particular, based on a strategy similar to this of non-cognitivist theories in ethics. Maybe the content of what the cognitive model calls "core beliefs" is not factual but evaluative in nature. If this is right, we should not expect that providing or examining counterevidence would suffice for their modification. I agree that this is a strategy worth exploring, however, for the lack of space, I cannot do it in the present paper.

made by the cognitive model—equally fundamental for the very idea of cognitive restructuring—unificationism regarding the organization of beliefs.

> **Unificationism**: All beliefs of a given subject (1) form a single, consistent set, and (2) are equally accessible and available at any given time and in any given context.

Condition 1 is supposed to reflect the fact that—as by-and-large rational subjects—we typically do not simultaneously believe that *p* and not *p*, e.g., that there is a coffee mug on the table and that there is no coffee mug on the table. If one does simultaneously believe such contradictory propositions, there is, presumably, something wrong with them. What exactly? On an idealized picture widespread in philosophy as well as in the folk discourse, consistency of one's belief states is often taken to be required for rationality (cf. Davidson, 1985; Quine & Ullian, 1978; Broome, 1999; Kolodny, 2008; Borgoni, 2021; Yalcin, 2021). Simultaneous belief in *p* and its negation is, thus, taken to be a mark of irrationality.

Condition 2 is also *prima facie* compelling. First, if I believe something, I probably believe it at all times (until my belief is changed) and wherever I am. At first sight, it does not seem that there are any beliefs that I hold only on Tuesdays or whenever I am at the farmer's market. Moreover, my actions seem to be responsive to all relevant beliefs that I hold. If I believe carrot cake is in the kitchen, I go to the kitchen once I feel like it is time for cake. But if I realize that the cake is five weeks old, it will stop me from devouring it.

Conditions 1 and 2 of unificationism taken together have far-reaching consequences for how our belief systems are supposed to behave. One of the consequences concerns the mechanism of belief revision. Following Bendaña and Mandelbaum (2021, p. 79), I will call it Quineanism:

> **Quineanism**: Belief revision is sensitive to global properties of an agent's total set of beliefs (that is, the beliefs taken as an entire set).

If (*per* unificationism) all beliefs of a given subject form a single, consistent set, and are equally accessible at all times and contexts, revision of any given belief will affect the entire belief set.[11] Moreover, acquisition of a new belief contradictory to the beliefs already held by a subject, will trigger contradiction resolution—either the new belief, or some of the old ones will have to go. This idea should already sound familiar to us, because this is exactly what is at play in cognitive restructuring.[12] Without the assumption that our beliefs form a consistent set (Condition 1 of unificationism) and that belief revision is sensitive to global properties of this set (Condition 2 of unificationism and Quineanism), Hofmann and Asmundson could not conclude, as

[11] This, of course, presupposes also that beliefs are evidence-responsive. While it is an orthodox assumption in epistemology (see, e.g., Velleman, 2000; Shah, 2003; Helton, 2020), it is not uncontested. Moreover, as suggested in footnote 6, the nature of what the cognitive model calls "core beliefs" may differ from the nature of more common, everyday beliefs, e.g., in that core beliefs are not (or less) evidence-responsive.

[12] At least in the most popular conceptualization of cognitive restructuring. I will briefly discuss an alternative conceptualization in Sect. 6.

they do in the quote which I have presented in Sect. 3, that patients confronted with evidence contradicting their maladaptive beliefs "are forced to revise their belief system" (2008, p. 12). Cognitive restructuring presupposes unificationism and Quineanism and can be challenged by undermining them. This is the goal of the next section.

5 The fragmentation of belief and cognitive restructuring

Unificationism has been questioned in the past (e.g., Cherniak, 1986; Lewis, 1982; Stalnaker, 1984, 1991) but, recently, it has gone under heavy scrutiny due to the rising popularity of fragmentationism (e.g., Egan, 2008, 2021; Rayo, 2013; Davies & Egan, 2013; Greco, 2014; Elga & Rayo, 2021; Mandelbaum, 2014, 2016, 2019; Yalcin, 2018, 2021; Kinderman & Onofri, 2021; Borgoni, 2021; Porot & Mandelbaum, 2021; Sommer et al., 2022). According to fragmentationism, both conditions of unificationism are false.

Firstly, the body of beliefs of a given subject does not form a single, consistent set. Rather, it is divided into separate fragments or compartments. The consistency requirement applies, at most, intra-fragmentally (all beliefs contained in a given fragment are consistent) but not inter-fragmentally (a subject can simultaneously hold inconsistent beliefs as long as they belong to different fragments). There is a simple yet convincing case to be made in support of this claim. As suggested by Sommer et al. (2022), given the size of our belief corpuses, their global consistency is computationally intractable:

> …achieving consistent beliefs appears to be an intractable problem for any reasonably large set of beliefs. There is not enough time for consistency checking of a large set of beliefs (intractability of consistency checking). People are not capable of deducing all the implications of the beliefs they hold to check for consistency (lack of closure). They cannot be expected to notice relevant information from any and all domains for each belief (isotropy). Finally, holding all beliefs in working memory simultaneously to optimize their global consistency is a task that is almost certainly beyond the scope of the human mind (the Quinean nature of belief). (Sommer et al., 2022, p. 6)

What begs for explanation is, therefore, not inconsistency but occasional and local consistency of our beliefs. Notably, Sommer and his colleagues (2022, pp. 13–14, fn. 4) stress that fragmentation does not *cause* inconsistency. Quite the opposite—fragmentation, understood primarily as coactivation of certain beliefs, allows for some degree of (occasional and local) consistency.

Secondly, according to fragmentationism, not all our beliefs are equally accessible and available at all times and contexts. For example, if we saw the film *The Sound of Music* a very long time ago, we may be unable to answer the question "What was the name of the youngest von Trapp child?" but, simultaneously, able to confidently answer "yes!" to the question "Was the youngest von Trapp child's name 'Gretl'?" (Egan, 2021, p. 111). How can this be, given that answering both these questions seems to require possession of the same belief, i.e., that the name of the youngest von

Trapp child in *The Sound of Music* was "Gretl"? Fragmentationists answer that the relevant belief is accessible for the purpose of answering the yes/no question, but not the wh-question.

Here, then, is a brief formulation of fragmentationism:

> **Fragmentationism**: Beliefs of a given subject (1) *do not* form a single, consistent set but are divided into separate fragments, and (2) *are not* equally accessible and available at any given time and in any given context.

According to supporters of fragmentationism, fragmentation is neither a sign of psychological malfunction nor cognitive failure. Rather, for better or worse, this is just the way information seems to be organized in the human mind. Here's an early statement of fragmentationism by David Lewis:

> I used to think that Nassau Street ran roughly east–west; that the railroad nearby ran roughly north–south; and that the two were roughly parallel… My system of beliefs was broken into (overlapping) fragments. Different fragments came into action in different situations, and the whole system of beliefs never manifested itself all at once. (Lewis, 1982, p. 436)

Lewis highlights that once he noticed the inconsistency, he immediately revised relevant beliefs. So, he was not pathologically *stuck* in the inconsistent way of thinking. It just so happened that, at least for some time, the relevant pieces of information were functionally separated from one another.[13]

Fragmentationism has been used to explain a number of psychological phenomena, from spontaneous recovery of extinguished associations to implicit bias and various strategies we employ to deal with cognitive dissonance (for overviews see Porot & Mandelbaum, 2021; Bendaña & Mandelbaum, 2021; Kinderman & Onofri, 2021). My goal in the remainder of this paper is to argue that it can also help us account for the limitations of cognitive restructuring. Crucially, in line with other supporters of fragmentationism, I do not assume fragmentation to be a symptom of a psychological dysfunction or disorder. Therefore, it is not my claim that the psychotherapeutic process of cognitive restructuring does not work because the minds of patients seeking therapy are exceptionally fragmented and thus non-susceptible to cognitive restructuring. Rather, I assume that fragmentation is a universal feature of (the organization of information in) our minds and thus even people who do not suffer from psychological problems are generally not very susceptible to interventions such as cognitive restructuring. So—the argument goes—if it does not work in general, no surprise it would not work in therapy.

[13] This is a good place to offer a caveat for readers who are skeptical towards the "belief-talk" in general. Even though fragmentationism is often spelled out by appeal to the container metaphor with different beliefs "belonging to different fragments" or "stored in separate containers", at its heart fragmentationism is a hypothesis about the patterns of information access. All the important problems and questions about fragmentationism can be spelled out without making any assumptions about the nature of beliefs or even their existence as anything more than a folk-psychological buzzword.

According to the cognitive model, successful cognitive restructuring results in the modification of core beliefs. Core beliefs are supposed to be the deepest convictions a person holds about themselves, the world, and other people. They are often first acquired through negative early childhood experiences, and immediately begin to function as "filters" through which the person sees the world and themselves in it (Riso & McBride, 2007). This, in turn, leads to further reinforcement of core beliefs, because the person is biased towards gathering the information consistent with them (Padesky, 1994; Beck et al., 2004).

On the unificationist picture assumed by the cognitive model, the core beliefs are the ones which lie close to the center of the metaphorical *web of beliefs*. Once the patient, with the help of the therapist, identifies them (e.g., using the *downward arrow* technique, starting with the easily accessible automatic thoughts and asking: "And if that's true, so what?", "What's the worst part about it?", "What does it mean about you?", etc. (Beck, 2011) they can start recalling and producing evidence contradicting these beliefs. Eventually, in light of the overwhelming evidence, the patient has to modify their core beliefs, which (by Quineanism) leads to the revision of the whole belief system and—according to the cognitive model—reduction of symptoms.

A radically different picture emerges when we assume fragmentationism. Given that fragmentationism does not presuppose the existence of a single, consistent web of belief, it has to offer an alternative characterization of core beliefs. Bendaña and Mandelbaum (2021, p. 81) do it by invoking an additional assumption about the fragmented systems of beliefs, i.e., redundancy.[14]

> **Redundancy**: Different tokens of any particular belief may be stored in different fragments.

If we assume redundancy, we can characterize core beliefs as beliefs which are spread around the largest number of fragments.[15] Notice, that this characterization fits very nicely with the cognitive model's assumptions about the acquisition of core beliefs and the way they affect the everyday life of people struggling with mental health problems. Even though not all fragments of our belief systems are accessible all the time, if one has tokens of a given core belief spread around multiple fragments, it is very likely that the thoughts and behaviors guided by this belief will emerge across different contexts and diverse life situations. Moreover, as expected by the cognitive model, the core, i.e., the most redundant beliefs are plausibly the ones which are

[14] Bendaña and Mandelbaum do not focus on maladaptive core beliefs discussed in the cognitive model of psychopathology, but positive core beliefs (that we are good, smart, reliable, etc.) constituting the positive self-image postulated on the ground of dissonance theory (cf. Thibodeau & Aronson, 1992; Aronson, 1997; Mandelbaum, 2019). However, from the point of view of the bare organization of belief, there is no difference between positive and negative core beliefs, so I take this part of their argumentation to be equally applicable to the problems discussed in this paper.

[15] This formulation of redundancy presupposes representationalism, and it is not exactly clear how redundancy (or something functionally identical) could be accounted for on the grounds of, e.g., dispositionalism (Quilty-Dunn & Mandelbaum, 2018, p. 2358 fn. 6). Nevertheless, given that the cognitive model itself presupposes representationalism, I am happy to stick with this definition for the purpose of current discussion.

acquired over a long period of time in multiple contexts and situations, for example, through repetitive, negative early childhood experiences.

Crucially, fragmentationism provides reasons why we should not expect cognitive restructuring to be very effective. Even if (i) a client, together with the therapist, produces and analyzes evidence contradicting a given core belief, and (ii) the evidence are sufficiently strong and (iii) the client is sufficiently evidence-responsive to revise this core belief, the revision is only local, i.e., it occurs only in a fragment or fragments activated during therapy.[16] *Contra* to what is hoped for in the unificationist picture assumed in the cognitive model, such a local change does not result in the revision of the whole belief system. In contexts other than this of a therapeutic session (such as stressful or triggering everyday life situations) other fragments get activated—fragments where a given core belief has not been revised. In such contexts, thoughts, behaviors, and emotions of the client are still influenced and governed by the maladaptive core beliefs.[17]

Let us illustrate it with an example. It is fictionalized, but it reflects elements of multiple actual cases. An adult therapy client, John, struggles with building lasting romantic relationships. During therapy, it turns out that many of John's behaviors and automatic thoughts seem to be guided by the core belief "I have to be on guard at all times or else others will hurt me." Considering the origin of this belief, John admits that he has had it for "as long as he remembers." This comes as no surprise to John's therapist. She knows that John was raised in an abusive home with alcoholic parents whose behavior was unpredictable and often aggressive. Many of her previous clients with similar childhood stories held similar core beliefs. This core belief naturally becomes one of the main targets of cognitive restructuring. John's therapist encourages him to consider whether the claim is not too strong and catastrophic; does he really have to be on guard "at all times"? What about when he is with his partner, Jill? During therapy sessions (and at home when he journals or does another exercise assigned by the therapist as homework), John is perfectly able to consider and reflect upon the fact that Jill gives him ample evidence of trustworthiness, and that she behaves not only predictively but lovingly. In result, John is ready to revise the maladaptive core belief. Asked by the therapist, he admits with full conviction "I don't have to be on guard when I'm with Jill." Nevertheless, during minor quarrels with Jill, John finds himself overreacting. In such situations he keeps noticing thoughts such as "I shouldn't have opened up to her," "She's just like the others," etc. Additionally, despite what he says during therapy, when he and Jill are intimate with each other, John feels tense and is unable to relax. It is extremely difficult for him to

[16] Bendaña and Mandelbaum (2021, p. 81) call this feature "multiple resistance:" the beliefs that are most redundant are, therefore, most resistant to revision.

[17] As rightly pointed out by an anonymous reviewer, we should keep in mind that the context of psychotherapy is special in that it is characterized by increased trust, safety, and openness, as well as "warmth, accurate empathy, and genuineness" (Beck et al., 1979, p. 45). This is an important point. However, even though the presence of such conditions may increase clients' susceptibility to cognitive restructuring in the fragments activated during therapy, we should not expect that it enables them to access all or most fragments of their belief systems. After all, the specific context of a psychotherapy session is just one among many contexts we find ourselves in every day, each associated with its own set of activated or easily accessible fragments. Moreover, due to redundancy, consciously focusing on a specific issue does not activate all or even most fragments containing beliefs relevant to this issue. I come back to this topic below.

focus on the present moment. Instead, his mind wanders; he finds himself thinking about how his other romantic relationships ended and imagines his relationship with Jill ending in a similar, dramatic way.

Why does John keep thinking and behaving as if he believes "I have to be on guard at all times or else others will hurt me"? One explanation—available to the supporters of unificationism—would be to say that the cognitive restructuring has not worked or has not worked *yet.* For some reason, despite what he asserts during the session, John has not yet revised his core belief and thus the belief keeps affecting different areas of his life. What John says when he reports to his therapist "I don't think that I have to be on guard when I'm with Jill" is false. This is not what he believes. Whether he can admit it or not, he still believes that he must be on guard at all times.[18]

A different explanation—one I find much more compelling—is that John's system of beliefs is fragmented. John has in fact revised the belief that he must be on guard at all times in the fragment (or fragments) of his belief system activated during therapy and accessible for the purposes of calm reflection. Therefore, he speaks truthfully when he says that he believes that he does not have to be on guard when he is with Jill. Nevertheless, the maladaptive core belief that he must be on guard at all times is spread around multiple fragments of his belief system. After all, learning that he has to be on guard at all times was a big part of his upbringing. So, he *also* believes that he must be on guard at all times. In result, despite the attempts of restructuring, in many situations John keeps thinking and behaving according to the redundant, core belief.

A very important feature of the fragmentationist explanation is that it validates clients' experience. As stated by Marsha Linehan, the author of the Dialectical Behavior Therapy (DBT) and a towering figure in the clinical psychology of the past fifty years: "Validation communicates to the patient in a non-ambiguous way that her behavior makes sense and is understandable in the current context." (Linehan, 1993, p. 221). The unificationist assumption that, despite what he says, John does not believe (or does not *really* believe) that he can relax when he is with Jill, could only reinforce his fears that he cannot trust his "own emotional reactions, cognitive interpretations, or behavioral responses" (Linehan, 1993, p. 222). The fragmentationist perspective allows us to acknowledge the full complexity of John's mental state: John believes that he does not have to be on guard with Jill *and* John believes that he has to be on guard at all times.

Finally, many CBT practitioners and theorists highlight the importance of homework, and meta-analyses confirm that greater homework compliance is related to better treatment outcome (e.g., Mausbach et al., 2010). One of the most popular types of CBT homework is the so-called "thought record" which is effectively an extension of cognitive restructuring work outside the therapy session, i.e., "[i]dentifying automatic thoughts and beliefs when patients notice a dysfunctional change in affect, behavior, or physiology, and then evaluating and responding to their cognitions

[18] As noted by an anonymous reviewer, supporters of unificationism should provide an explanation of why cognitive restructuring has not worked (or has not worked *yet*) in John's case. It is not clear what could it be. Was he not concentrated or motivated enough? Were his beliefs particularly unresponsive to evidence? If so, why? At the same time, fragmentationism has a clear and simple explanation ready at hand.

through Socratic questioning, behavioral experiments, and/or reading therapy notes that address their cognitions" (Beck, 2011, p. 28). The positive relationship between homework compliance and treatment outcomes is another piece of empirical data fitting nicely into the fragmentationist picture. If someone wanted to stick to cognitive restructuring while assuming fragmentationism, a natural question would be: how can we revise maladaptive core beliefs in as many fragments as possible? One way to go about it would be to extend restructuring outside the context of a therapy session, to a client's home, workplace, etc. As might be expected, *sparse* restructuring, i.e., restructuring occurring in more contexts is more effective than restructuring occurring only during therapy sessions.

Unfortunately, different contexts are not simply different places. Even though we do not have a precise definition of factors that make something a fragment-activating-context, it should be expected that contexts are constituted by multiple events happening both in the outside world and in clients' mental life. A client can be expected to attempt cognitive restructuring in contexts which we can schematically characterize as <at home, while having a minute for myself>, <during a lunch break at work>, and even <in the train on my way to a stressful meeting>, but it is much less plausible that they will be able to engage in cognitive restructuring in such contexts as <in the midst of an argument with my spouse>, <during a difficult and important exam>, etc. In such situations, a person lacks resources necessary to perform any demanding cognitive tasks (e.g., Eysenck & Calvo, 1992), including the revision of their beliefs. Even though it might be the case that "*moderate* levels of emotional arousal and distress may be a necessary component of effective cognitive change" (Hunt, 1998 p. 381, emphasis in the original) this is not so for high arousal. This is why emotion regulation facilitated by a therapist is an important component of almost all therapeutic sessions. One could suggest that, for example in the case of an argument with one's spouse, the person could simply stop arguing and focus on themselves for a moment, in order to try and change the way they think. However, plausibly, the more would change in the person's emotional state, immediate surroundings, and activity they are engaged in, the further away they move from the context associated with the problematic ways of thinking.

To sum up, even though sparse cognitive restructuring is, unsurprisingly, more effective than restructuring occurring only during therapy sessions,[19] we should still expect its efficacy to be limited. This is not to say that cognitive restructuring is not effective at all. Once again, restructuring the fragments of our belief systems activated during therapy or self-help work may well be quite effective. Nevertheless, the basic idea of classical CBT was that this change would propagate to other contexts and areas of our life. Fragmenationism gives us reasons to doubt it.

[19] This might be an additional motivation to further investigate the benefits of integrating cognitive restructuring with experiential techniques (Greenberg et al., 1989), such as psychodrama, which may increase the efficacy of cognitive restructuring by triggering the activation of more fragments during a psychotherapeutic session itself. It is, however, unlikely that it would result in the effective restructuring of maladaptive beliefs in all fragments.

6 Consequences for psychotherapeutic practice

I have suggested that most theorists working within the paradigm of the cognitive model and CBT endorses a form of unificationism. Nevertheless, it would be naïve to assume that such a vast therapeutic tradition, associating thousands of practitioners over the last seventy years, will be entirely monolithic. Moreover, "the development of cognitive therapies was not closely tied to a single recognizable strand of basic research and theory in psychology," and thus "the development of effective therapies has preceded theoretical understanding" (Brewin, 1996, p. 36). Thus, in CBT literature, one can find traces of thinking seemingly compatible with fragmentationism. In her classic textbook, Judith Beck provides the following example:

> Sally, too, has a core belief of incompetence. Fortunately, when she is not depressed a different schema (which contains the core belief, "I'm reasonably competent") is activated much, but not all, of the time. But when she is depressed, the incompetence schema predominates. (Beck, 2011, p. 33)

According to this characterization, cognitive schemas[20] behave at least to some extent similarly to the fragments of a belief system. Sally believes both that she is incompetent and that she is reasonably competent, but the two beliefs are insulated from one another and not active simultaneously. The belief that she is incompetent becomes active and accessible during the acute phase of her depression. Similarly, Artnz (2018, p. 343) claims that "Sometimes people have dual belief systems… believing the core belief in certain conditions but not in others."

While these authors seem to recognize the fragmentation of belief and thus move away from the classical version of the cognitive model, they nevertheless recommend cognitive restructuring as the main therapeutic process. If what I said earlier in the paper is on the right track, this is problematic. If our belief systems are fragmented, we should expect cognitive restructuring to be, at most, moderately effective. At best, it leads to revision of beliefs in the fragments activated during therapy sessions and not the ones activated during other life situations including those of high distress. Thus, the authors should either drop their declarative commitment to belief fragmentation or, preferably, focus their practical recommendations on therapeutic processes other than cognitive restructuring.

Moreover, it is worth pointing out that throughout the years different authors proposed competing characterizations of cognitive restructuring itself. To this point I have been focusing on the classical, Beckian model, according to which cognitive restructuring is the process of correcting the irrational thinking by directly modifying maladaptive beliefs. Alternatively, some authors defended the *activation-deactivation model*, according to which the goal of restructuring is not the modification of maladaptive beliefs but an activation of competitive, positive ways of thinking which, in result, become more accessible than the maladaptive ones and thus are brought to bear in more contexts and life situations (Barber & DeRubeis, 1989; Kwon

[20] Judith Beck's use of "schema" in this context is similar to the way it is used in (A. T. Beck & Haigh, 2014) (cf. footnote 10 above).

& Oei, 1994; Brewin, 2006). *Prima facie*, this model appears to be more compatible with the hypothesis of belief fragmentation. Nevertheless, it remains puzzling how, in the process of therapy, can we assure a stable pattern of activation of fragments containing adaptive beliefs and deactivation of fragments containing maladaptive ones.

Finally, let me stress that fragmentationism undermines the idea of belief revision in psychotherapy, i.e., cognitive restructuring, but not necessarily the whole project of CBT.[21] There are psychotherapeutic approaches belonging to the CBT tradition broadly construed, which are more compatible with fragmentationism because they ascribe zero or only minimal importance to cognitive restructuring. These are so-called "third wave" Cognitive Behavioral Therapies, such as Dialectical Behavioral Therapy (DBT) (Linehan, 1993), Compassion Focused Therapy (CFT) (Gilbert, 2010), or Acceptance and Commitment Therapy (ACT) (Hayes et al., 1999). Instead of attempting the belief revision, these approaches focus on clients' acquiring and developing new skills and attitudes, e.g., mindfulness, acceptance, and compassion.[22] Such skills, in turn, enable clients to distance themselves from the contents of overwhelming narratives about the world, others, and oneself without disputing or modifying them, and act in accordance with their goals and values *despite* the nudging of maladaptive thoughts and beliefs (Zettle, 2005). According to the third-wave approaches, it is unlikely that in therapy we can free clients from the problematic patterns of thinking which contribute to their mental suffering. Even if—after a long time and only to some extent—this happens, it is even more unlikely that the belief revision was the *mediator* of therapeutic change. What changes the lives of therapy clients are first and foremost not cognitive but behavioral processes, e.g., exposure, behavioral activation, and mindfulness; what clients are encouraged to do, rather than what they are encouraged to think or believe. Given that cognitive restructuring encounters limitations indicated by many clinical psychologists, and that these limitations are to be expected considering the fragmentation of belief, the approach exemplified by the third-wave therapies might be the way to go with psychotherapeutic practice within the CBT tradition.[23]

7 Conclusions

My goal in this paper was to offer a philosophical appraisal of the therapeutic process of cognitive restructuring—the core process of classical Cognitive Behavioral Therapy. I did not aspire to offer an empirical examination of its efficacy. Instead, I have taken mixed or pessimistic opinions about its efficacy voiced by some clinical

[21] As I have already mentioned, there may be other more general arguments against CBT as a whole. They, however, lie outside the focus of the present paper, which aims to examine specifically the idea of belief revision in psychotherapy, central to the classical version of CBT.

[22] Cf. (Zawidzki, 2019).

[23] At the same time, in our future research, we should consider whether phenomena similar to fragmentation are not present also in the domains of behavioral dispositions and skill-based performance. If they are, resigning from attempts of belief revision in psychotherapy may not be enough, and we should think much more generally about ways of moving the therapeutic work outside the narrow context of therapy sessions. I am grateful to an anonymous reviewer for bringing up this point.

psychologists as my point of departure and asked the following question: do we have good theoretical reasons to assume that cognitive restructuring is effective? To answer this question, I first offered a reconstruction of the main philosophical assumptions underlying the cognitive model of psychopathology and classical CBT. I argued that their core assumptions, and the source of the very idea of cognitive restructuring, is unificationism regarding the organization of beliefs and Quineanism regarding their revision. Once we reject these assumptions and accept that our belief systems are fragmented and largely inconsistent, with different fragments active in different contexts and belief revision happening locally instead of globally, our expectations regarding the efficacy of belief revision in psychotherapy drop.

Acknowledgements I presented earlier versions of this paper at the Lancaster Philosophy of Psychiatry Workshop, Meeting of the European Society of Philosophy and Psychology in Leipzig, "Talking as Cure" Conference in Cambridge, and The Annual Meeting of the Association for Advancement of Philosophy and Psychiatry in New Orleans. I am grateful to the audiences at these conferences for illuminating discussions. Special thanks to Sarah Arnaud, Konrad Banicki, Anna Drożdżowicz, Andrew Evans, Grzegorz Gaszczyk, Mateusz Hohol, Pablo Hubacher Haerle, Bartosz Kleszcz, Garson Leder, Sahanika Ratnayake, Katherine Rickus, Jared Smith, Jan Topczewski, Kinga Wołoszyn, Tad Zawidzki, and two anonymous reviewers for many helpful comments and suggestions.

Funding The research was supported by the National Science Centre, Poland (grant number: 2022/47/D/HS1/00923).

Declarations

Conflicts of interest The author has no competing interests to declare that are relevant to the content of this article.

References

Arch, J. J., Eifert, G. H., Davies, C., Plumb Vilardaga, J. C., Rose, R. D., & Craske, M. G. (2012). Randomized clinical trial of cognitive behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for mixed anxiety disorders. *Journal of Consulting and Clinical Psychology*, *80*(5), 750–765. https://doi.org/10.1037/a0028310.

Arntz, A. (2018). Modifying core beliefs. In S. C. Hayes, & S. G. Hofmann (Eds.), *Process-based CBT: The science and core clinical competencies of cognitive behavioral therapy* (pp. 339–350). New Harbinger Publications, Inc.

Aronson, E. (1997). Back to the future: Retrospective Review of Leon Festinger's a theory of cognitive dissonance. *The American Journal of Psychology*, *110*(1), 127–137. https://doi.org/10.2307/1423706.

Baddeley, A. (2007). *Working memory, thought, and action* (pp. xviii, 412). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198528012.001.0001.

Barber, J. P., & DeRubeis, R. J. (1989). On second thought: Where the action is in cognitive therapy for depression. *Cognitive Therapy and Research*, *13*(5), 441–457. https://doi.org/10.1007/BF01173905.
Beck, A. T. (1967). *Depression*. New York.
Beck, A. T. (1996). Beyond belief: A theory of modes, personality, and psychopathology. *Frontiers of cognitive therapy* (pp. 1–25). The Guilford.
Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond* (2nd ed.). Guilford Press.
Beck, A. T., & Haigh, E. A. P. (2014). Advances in Cognitive Theory and Therapy: The generic cognitive model. *Annual Review of Clinical Psychology*, *10*(1), 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734.
Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (Eds.). (1979). *Cognitive therapy of depression*. Guilford Press.
Beck, A. T., Freeman, A., & Davis, D. D. (2004). *Cognitive therapy of personality disorders, 2nd ed* (pp. xx, 412). Guilford Press.
Bendaña, J., & Mandelbaum, E. (2021). The fragmentation of belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.
Bhar, S. S., Brown, G. K., & Beck, A. T. (2008). Dysfunctional beliefs and psychopathology in Borderline personality disorder. *Journal of Personality Disorders*, *22*(2), 165–177. https://doi.org/10.1521/pedi.2008.22.2.165.
Borgoni, C. (2021). Rationality in Fragmented Belief systems. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.
Bortolotti, L. (2022). Delusion. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/delusion/.
Brewin, C. R. (1996). Theoretical foundations of cognitive-behavior therapy for anxiety and depression. *Annual Review of Psychology*, *47*, 33–57. https://doi.org/10.1146/annurev.psych.47.1.33.
Brewin, C. R. (2006). Understanding cognitive behaviour therapy: A retrieval competition account. *Behaviour Research and Therapy*, *44*(6), 765–784. https://doi.org/10.1016/j.brat.2006.02.005.
Broome, J. (1999). Normative requirements. *Ratio*, *12*(4), 398–419. https://doi.org/10.1111/1467-9329.00101.
Burge, T. (2010). Origins of Objectivity. *Origins of Objectivity*. Oxford University Press. https://oxford.universitypressscholarship.com/view/https://doi.org/10.1093/acprof:oso/9780199581405.001.0001/acprof-9780199581405.
Cherniak, C. (1986). *Minimal rationality*. MIT Press.
Clark, D. A. (2013). Cognitive Restructuring. In *The Wiley Handbook of Cognitive Behavioral Therapy* (pp. 23–44). American Cancer Society. https://doi.org/10.1002/9781118528563.wbcbt02.
Cook, J. M., Biyanova, T., Elhai, J., Schnurr, P. P., & Coyne, J. C. (2010). What do psychotherapists really do in practice? An internet study of over 2,000 practitioners. *Psychotherapy: Theory Research Practice Training*, *47*(2), 260–267. https://doi.org/10.1037/a0019788.
David, D., Cristea, I., & Hofmann, S. G. (2018). Why cognitive behavioral therapy is the current gold Standard of Psychotherapy. *Frontiers in Psychiatry*, *9*, 4. https://doi.org/10.3389/fpsyt.2018.00004.
Davidson, D. (1985). Incoherence and irrationality. *Dialectica*, *39*(4), 345–354. http://www.jstor.org/stable/42969059.
Davies, M., & Egan, A. (2013). Delusion: Cognitive approaches—bayesian inference and compartmentalization. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini, & T. Thornton (Eds.), *The Oxford Handbook of Philosophy and Psychiatry* (Vol. 1, pp. 689–730). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199579563.013.0042.
Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dunner, D. L., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology*, *74*(4), 658–670. https://doi.org/10.1037/0022-006X.74.4.658.
Dobson, K. S., & Dozois, D. J. A. (2010). Historical and philosophical bases of the cognitive-behavioral therapies. *Handbook of cognitive-behavioral therapies* (3rd ed., pp. 3–38). Guilford Press.
Dobson, K. S., & Shaw, B. F. (1986). Cognitive assessment with major depressive disorders. *Cognitive Therapy and Research*, *10*(1), 13–29. https://doi.org/10.1007/BF01173379.
Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes* (pp. xi, 165). The MIT Press.
Egan, A. (2008). Seeing and believing: Perception, belief formation and the divided mind. *Philosophical Studies*, *140*(1), 47–63. https://doi.org/10.1007/s11098-008-9225-1.

Egan, A. (2021). Fragmented models of belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.
Elga, A., & Rayo, A. (2021). Fragmentation and Information Access. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.
Ellis, A. (1962). *Reason and emotion in psychotherapy* (p. 442). Lyle Stuart.
Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The Processing Efficiency Theory. *Cognition and Emotion*, *6*(6), 409–434. https://doi.org/10.1080/02699939208409696.
Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. MIT Press.
Forman, E. M., & Herbert, J. D. (2009). New directions in cognitive behavior therapy: Acceptance-based therapies. *General principles and empirically supported techniques of cognitive behavior therapy* (pp. 77–101). John Wiley & Sons, Inc.
Fournier, J. C., DeRubeis, R. J., & Beck, A. T. (2012). Dysfunctional cognitions in personality pathology: The structure and validity of the personality belief Questionnaire. *Psychological Medicine*, *42*(4), 795–805. https://doi.org/10.1017/S0033291711001711.
Gilbert, P. (2010). *Compassion focused therapy: Distinctive features* (pp. viii, 237). Routledge/Taylor & Francis Group.
Gipps, R. G. T. (2013). Cognitive behavior therapy: A philosophical appraisal. In K. W. M. In M. Fulford, R. G. T. Davies, G. Gipps, J. Z. Graham, G. Sadler, Stanghellini, & T. Thornton (Eds.), *The Oxford Handbook of Philosophy and Psychiatry* (Vol. 1, pp. 1245–1263). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199579563.013.0072.
Greco, D. (2014). Iteration and fragmentation. *Philosophy and Phenomenological Research*, *88*(1), 656–673. https://doi.org/10.1111/phpr.12086.
Greenberg, L. S., Safran, J., & Rice, L. (1989). Experiential Therapy. In A. Freeman, K. M. Simon, L. E. Beutler, & H. Arkowitz (Eds.), *Comprehensive Handbook of Cognitive Therapy* (pp. 169–187). Springer US. https://doi.org/10.1007/978-1-4757-9779-4_9.
Handley, A. K., Egan, S. J., Kane, R. T., & Rees, C. S. (2014). The relationships between perfectionism, pathological worry and generalised anxiety disorder. *Bmc Psychiatry*, *14*(1), 98. https://doi.org/10.1186/1471-244X-14-98.
Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, *35*(4), 639–665. https://doi.org/10.1016/S0005-7894(04)80013-3.
Hayes, S. C., & Brownstein, A. J. (1986). Mentalism, behavior-behavior relations, and a behavior-analytic view of the purposes of science. *The Behavior Analyst*, *9*(2), 175–190. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2741891/.
Hayes, S. C., & Hofmann, S. G. (2017). The third wave of cognitive behavioral therapy and the rise of process-based care. *World Psychiatry*, *16*(3), 245–246. https://doi.org/10.1002/wps.20442.
Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change* (pp. xvi, 304). Guilford Press.
Helton, G. (2020). If you can't change what you believe, you don't believe it. *Noûs*, *54*(3), 501–526. https://doi.org/10.1111/nous.12265.
Hofmann, S. G. (2008). Common misconceptions about cognitive mediation of treatment change: A commentary to Longmore and Worrell (2007). *Clinical Psychology Review*, *28*(1), 67–70. https://doi.org/10.1016/j.cpr.2007.03.003.
Hofmann, S. G., & Asmundson, G. J. G. (2008). Acceptance and mindfulness-based therapy: New wave or old hat? *Clinical Psychology Review*, *28*(1), 1–16. https://doi.org/10.1016/j.cpr.2007.09.003.
Hofmann, S. G., Asmundson, G. J. G., & Beck, A. T. (2013). The science of cognitive therapy. *Behavior Therapy*, *44*(2), 199–212. https://doi.org/10.1016/j.beth.2009.01.007.
Hollon, S. D., & Beck, A. T. (1994). Cognitive and cognitive-behavioral therapies. *Handbook of psychotherapy and behavior change* (4th ed., pp. 428–466). Wiley.
Hollon, S. D., Shelton, R. C., & Davis, D. D. (1993). Cognitive therapy for depression: Conceptual issues and clinical efficacy. *Journal of Consulting and Clinical Psychology*, *61*(2), 270–275. https://doi.org/10.1037//0022-006x.61.2.270.
Hunt, M. G. (1998). The only way out is through: Emotional processing and recovery after a depressing life event. *Behaviour Research and Therapy*, *36*(4), 361–384. https://doi.org/10.1016/S0005-7967(98)00017-5.
Ingram, R. E., Atchley, R. A., & Segal, Z. V. (2011). *Vulnerability to depression: From cognitive neuroscience to prevention and treatment* (pp. xii, 260). The Guilford Press.

Jacobson, N., Dobson, K., Truax, P., Addis, M., Koerner, K., Gollan, J., Gortner, E., & Prince, S. (1996). A component analysis of cognitive-behavioral treatment for Depression. *Journal of Consulting and Clinical Psychology*, *64*, 295–304. https://doi.org/10.1037/1522-3736.3.1.323a.

Kazdin, A. E. (2007). Mediators and mechanisms of change in Psychotherapy Research. *Annual Review of Clinical Psychology*, *3*(1), 1–27. https://doi.org/10.1146/annurev.clinpsy.3.022806.091432.

Kinderman, D., & Onofri, A. (2021). The fragmented mind: An Introduction. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.

Kolodny, N. (2008). Why be disposed to be coherent? *Ethics*, *118*(3), 437–463. https://doi.org/10.1086/528783.

Kwon, S. M., & Oei, T. P. S. (1994). The roles of two levels of cognitions in the development, maintenance, and treatment of depression. *Clinical Psychology Review*, *14*(5), 331–358. https://doi.org/10.1016/0272-7358(94)90032-9.

Lacewing, M. (2004). Emotion and cognition: Recent developments and therapeutic practice. *Philosophy Psychiatry & Psychology*, *11*(2), 175–186. https://doi.org/10.1353/ppp.2004.0054.

Leder, G. (2017). Know Thyself? Questioning the theoretical foundations of cognitive behavioral therapy. *Review of Philosophy and Psychology*, *8*(2), 391–410. https://doi.org/10.1007/s13164-016-0308-1.

Lewis, D. (1982). Logic for equivocators. *Noûs*, *16*(3), 431–441. https://doi.org/10.2307/2216219.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder* (pp. xvii, 558). Guilford Press.

Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review*, *27*(2), 173–187. https://doi.org/10.1016/j.cpr.2006.08.001.

Lorenzo-Luaces, L., German, R. E., & DeRubeis, R. J. (2015). It's complicated: The relation between cognitive change procedures, cognitive change, and symptom change in cognitive therapy for depression. *Clinical Psychology Review*, *41*, 3–15. https://doi.org/10.1016/j.cpr.2014.12.003.

Mandelbaum, E. (2014). Thinking is believing. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, *57*(1), 55–96. https://doi.org/10.1080/0020174X.2014.858417.

Mandelbaum, E. (2016). Attitude, Inference, Association: On the Propositional structure of Implicit Bias. *Noûs*, *50*(3), 629–658. https://doi.org/10.1111/nous.12089.

Mandelbaum, E. (2019). Troubles with Bayesianism: An introduction to the psychological immune system. *Mind & Language*, *34*(2), 141–157. https://doi.org/10.1111/mila.12205.

Martin, R., & Young, J. (2010). Schema therapy. *Handbook of cognitive-behavioral therapies* (3rd ed., pp. 317–346). Guilford Press.

Mausbach, B. T., Moore, R., Roesch, S., Cardenas, V., & Patterson, T. L. (2010). The Relationship between Homework Compliance and Therapy outcomes: An updated Meta-analysis. *Cognitive Therapy and Research*, *34*(5), 429–438. https://doi.org/10.1007/s10608-010-9297-z.

McEachrane, M. (2009). Capturing Emotional Thoughts: The Philosophy of Cognitive-Behavioral Therapy. In Y. Gustafsson, C. Kronqvist, & M. McEachrane (Eds.), *Emotions and Understanding* (pp. 91–106). Palgrave Macmillan. https://doi.org/10.1057/9780230584464_6.

Millikan, R. G. (1984). *Language, Thought and other Biological categories*. MIT Press.

Newman, C. F. (2015). Cognitive Restructuring/Cognitive Therapy. In C. M. Nezu, & A. M. Nezu (Eds.), *The Oxford Handbook of cognitive and behavioral therapies* (pp. 118–156). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199733255.013.22.

O'Donohue, W. T., & Ferguson, K. E. (2015). Historical and philosophical dimensions of contemporary cognitive-behavioral therapy. In C. M. Nezu, & A. M. Nezu (Eds.), *The Oxford Handbook of cognitive and behavioral therapies* (pp. 8–28). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199733255.013.27.

Padesky, C. A. (1994). Schema change processes in cognitive therapy. *Clinical Psychology & Psychotherapy*, *1*(5), 267–278. https://doi.org/10.1002/cpp.5640010502.

Porot, N., & Mandelbaum, E. (2021). The science of belief: A progress report. *WIREs Cognitive Science*, *12*(2). https://doi.org/10.1002/wcs.1539.

Quilty-Dunn, J., & Mandelbaum, E. (2018). Against dispositionalism: Belief in cognitive science. *Philosophical Studies*, *175*(9), 2353–2372. https://doi.org/10.1007/s11098-017-0962-x.

Quine, W. V. O., & Ullian, J. S. (1978). *The web of belief*. Random House.

Ratnayake, S. (2021). *It's been utility all along: An alternate understanding of cognitive behavioral therapy and the depressive realism hypothesis*. Philosophy, Psychiatry & Psychology.

Rayo, A. (2013). A Plea for semantic localism. *Noûs*, *47*(4), 647–679.

Riso, L. P., & McBride, C. (2007). Introduction: A return to a focus on cognitive schemas. In L. P. Riso, du P. L. Toit, D. J. Stein, & J. E. Young (Eds.), *Cognitive schemas and core beliefs in psychological problems: A scientist-practitioner guide* (pp. 3–9). American Psychological Association. https://doi.org/10.1037/11561-001.

Schulz, S. M., Alpers, G. W., & Hofmann, S. G. (2008). Negative self-focused cognitions mediate the Effect of Trait Social anxiety on state anxiety. *Behaviour Research and Therapy*, *46*(4), 438–449. https://doi.org/10.1016/j.brat.2008.01.008.

Schwitzgebel, E. (2019). Belief. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/belief/.

Shah, N. (2003). How truth governs belief. *The Philosophical Review*, *112*(4), 447–482.

Society of Clinical Psychology (2023). *Treatments*. https://div12.org/treatments/.

Sommer, J., Musolino, J., & Hemmer, P. (2022). A hobgoblin of large minds: Troubles with consistency in belief. *WIREs Cognitive Science*, *14*(4), e1639. https://doi.org/10.1002/wcs.1639.

Stalnaker, R. (1984). *Inquiry*. MIT Press.

Stalnaker, R. (1991). The problem of logical omniscience. *I Synthese*, *89*(3), 425–440. http://www.jstor.org/stable/20116982.

Teasdale, J. D., Dritschel, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus-independent thought depends on central executive resources. *Memory & Cognition*, *23*(5), 551–559. https://doi.org/10.3758/BF03197257.

Thibodeau, R., & Aronson, E. (1992). Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin*, *18*(5), 591–602. https://doi.org/10.1177/0146167292185010.

Velleman, D. (2000). *The possibility of practical reason*. Oxford University Press.

Weissman, A. N. (1979). The Dysfunctional Attitude Scale: A Validation Study. *Publicly Accessible Penn Dissertations*. https://repository.upenn.edu/edissertations/1182.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology*, *51*, 59–91. https://doi.org/10.1146/annurev.psych.51.1.59.

Whiting, D. (2006). Why treating problems in emotion may not require altering eliciting cognitions. *Philosophy Psychiatry & Psychology*, *13*(3), 237–246. https://doi.org/10.1353/ppp.2007.0024.

Yalcin, S. (2018). Belief as question-sensitive. *Philosophy and Phenomenological Research*, *97*(1), 23–47. https://doi.org/10.1111/phpr.12330.

Yalcin, S. (2021). Fragmented but rational. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind*. Oxford University Press.

Young, J. E. (1990). *Cognitive therapy for personality disorders: A schema-focused approach* (pp. xi, 79). Professional Resource Exchange, Inc.

Zawidzki, T. (2019). Metacognitive skill and the therapeutic regulation of emotion. *Philosophical Topics*, *47*(2), 27–52. http://www.jstor.org/stable/26948105.

Zettle, R. D. (2005). The evolution of a Contextual Approach to Therapy: From Comprehensive distancing to ACT. *International Journal of Behavioral Consultation and Therapy*, *1*(2), 77–89.