The cognitive model

Situation, thought, emotion, body and behaviour: how they connect, what counts as a thought, and why the way CBT works is still debated

DRAFT this text has not yet been approved by the author and may change.

17 chapters ~40 min read

How solid is each fact

Proven by research | Clinical consensus | In the author's own words | Interpretation / disputed | Documented fact | Not verified

In short

- **It isn't the event that sets off a reaction, it's how we read it.** This is the central claim of Aaron Beck's cognitive model. The same message from a manager can make one person anxious, another curious and a third irritated.
- **Thoughts, emotions, body and behaviour are four different things.** A thought is a judgement or interpretation ("he's unhappy with me"). An emotion can usually be named in one word ("anxiety"). The body is sensations ("my chest tightened"). Behaviour is what the person did. "I feel like a failure" is a thought, not a feeling.
- **Under a quick thought lie two layers of belief.** In the middle are rules and if–then assumptions people rarely say out loud. At the bottom are short, absolute beliefs about oneself, such as "I'm incompetent" or "I'm unlovable".
- **Beliefs can lie dormant.** In people who had recovered from depression, what mattered most was not how they thought on a good day but how sharply their thinking worsened under mild sadness. Among those whose thinking worsened markedly, 69% relapsed within 18 months, against 30% of those whose thinking barely changed.
- **The model describes well how a reaction unfolds,** and partly explains why some people are more vulnerable than others. The idea that each disorder has its own typical thought content has held up only halfway.
- **It has not been shown that CBT works because it changes thoughts.** Thinking changes with any successful treatment for depression. In one key trial, the behavioural part of the therapy did as well as full cognitive therapy. A review of 35 studies of mechanisms found not one that met all the criteria for proof.

- **The opposite hasn't been shown either.** "There's no need to change thoughts" does not follow from these data: behaviour changes thoughts too, and after cognitive therapy ended, depression came back less often than after medication was stopped.

Timeline

How the cognitive model and the tests of it developed, from the first clinical observations to studies of mechanisms. Details are in the chapters below.

1963 — Beck's paper "Thinking and depression": the distinctive content of thoughts and the distortions in depressed patients
COGNITIVE LINE

1964–1967 — Beck describes schemas as the basis of depressive thinking
COGNITIVE LINE

1979 — "Cognitive Therapy of Depression" by Beck, Rush, Shaw and Emery; the downward arrow technique is first mentioned here
COGNITIVE LINE

1980 — David Burns develops the downward arrow in "Feeling Good"
COGNITIVE LINE

1990 — DeRubeis and colleagues: changes in thinking midway through cognitive therapy predict later improvement
EVIDENCE

1996 — Jacobson and colleagues: full cognitive therapy does no better than its behavioural component
EVIDENCE

1999 — Tang and DeRubeis describe "sudden gains" in therapy
EVIDENCE

2001 — Beck and Perkins' meta-analysis: "each disorder has its own thoughts" holds up only partly
EVIDENCE

2004 — Steven Hayes writes about the "third wave": change your relationship to thoughts rather than their content
THIRD WAVE

2005 — Hollon and colleagues: depression returns less often after cognitive therapy ends than after medication is stopped
EVIDENCE

2006 — Segal and colleagues: how thinking reacts to mild sadness predicts relapse
EVIDENCE

2014 — Beck and Haigh propose a "generic cognitive model" for all disorders
COGNITIVE LINE

2016 — Review by Lemmens and colleagues: none of 35 studies of mechanisms meets all the criteria
EVIDENCE

2018 — Osmo and colleagues: three categories of core beliefs turn out to be four when tested
EVIDENCE

How to read these notes: evidence labels

Every claim here carries a label that tells you how solid it is.

These are open notes on the cognitive model, the second in a series on CBT (cognitive behavioural therapy); they can be read on their own. The label comes first in brackets, followed by the source. **Proven by research** means randomised trials (participants are split into groups by chance and the results compared) or meta-analyses (studies that pool many trials). **Clinical**

consensus means clinical guidelines and established practice. **The author's own account** means what the founder of an approach or their institute holds; it is theory, not a tested fact. **Interpretation** is a reading of the data that is still debated. **Documented fact** covers dates and publications. **Not verified** means we couldn't open the primary source. Where sources disagree, both versions are given.

> Important This is education, not therapy or diagnosis. Techniques are described as they are used in therapy, not as instructions for working on your own. These notes mention thoughts that life isn't worth living. If you need support now, see If you need help now.

Between event and feeling: the model in one diagram

The core idea of the cognitive model fits on one line: between what happened and what we felt, there is an interpretation.

Picture a Friday, twenty to four. Your manager writes on the work chat: "Anna, come and see me at 5." Nothing else.

Anna is a composite character; the name is made up. But almost anyone who has worked in an office has had a message like that. What would be your first thought?

One situation, three evenings

Here are three possible reactions to the same message:

First thought	Emotion	What the person does
"I've done something wrong"	anxiety, shame	rereads the week's reports, can't get any work done
"Finally, we'll talk about my project"	interest, mild excitement	jots down a few points
"Friday evening again, seriously?"	irritation	texts a friend that she'll be late

One situation, three different evenings. What makes the difference is not the message but how it was read. That is the cognitive model: emotions, bodily reactions and behaviour depend less on the situation itself than on how it is interpreted **the author's own account; Beck & Haigh, 2014**.

Four links in a reaction

CBT breaks a reaction down into parts:

1. **Situation.** An outside event (a message, a conversation), or sometimes an inside one: a bodily sensation, a memory, even another thought.
2. **Automatic thought.** A quick judgement that comes by itself: "I've done something wrong."
3. **Reaction,** which has three sides: **emotion** (anxiety), **body** (tight chest, cold hands, trouble concentrating) and **behaviour** (rechecking reports, messaging a colleague "do you know what's going on?").
4. **Consequences.** Behaviour shapes what happens next and can reinforce the original reading. We come back to this in the chapter on the body and behaviour.

CBT clinicians help people tell thoughts, feelings and behaviours apart **clinical consensus; Beck & Haigh, 2014, p. 16**. The next chapter shows why that matters so much.

This is not about blame

One caveat straight away. The model does not say "your anxiety is your own fault." Anna may have had calls like this end in being fired, in which case her thought didn't come from nowhere. In his later work Beck explicitly describes genetics, early experience and stress as contributing to vulnerability **the author's own account; Beck & Haigh, 2014**. An interpretation is not a reason for blame; it's a place where something can change.

Myth vs fact **Myth:** according to CBT, thoughts create feelings, so people are to blame for their own depression. **Fact:** the model includes biological predisposition, early experience and real stress **the author's own account; Beck & Haigh, 2014**. The interpretation is a convenient place to start the work, not the only cause.

What counts as a thought

The first skill in CBT is telling a thought from an emotion. It's harder than it sounds, because everyday language mixes them up all the time.

Automatic thoughts

An **automatic thought** is a quick appraisal of a situation that arises on its own, without intention, and feels like the truth rather than an opinion. In CBT these are usually the first target of the work **clinical consensus; Wenzel, 2012**. Wenzel describes them as thoughts tied to a specific situation and linked to a worsening of mood **clinical consensus; Wenzel, 2012**.

Their features:

- **Fast.** The thought flashes by in a split second, often while the person is saying or doing something else.
- **Specific.** They are about this situation: "I can't manage this report", not "I'm no good at anything."
- **Believable.** In the moment, they don't feel like a hypothesis. They feel like how things are.
- **Tied to emotion.** The stronger the emotion, the more likely there is a "hot" thought behind it.

Beck and Haigh describe the automatic thought as the form a deeper belief takes in a particular situation. Their example: a depressed lawyer who believes "I am incompetent" thinks "I can't do it" when asked to return a client's call **the author's own account; Beck & Haigh, 2014, p. 6**.

"Automatic" here means "it comes by itself". In the same paper Beck and Haigh also discuss an "automatic processing system", fast, economical and error-prone, as opposed to a slower reflective one **the author's own account; Beck & Haigh, 2014, p. 4**. The two ideas are related but not the same, and it's worth not mixing them up (interpretation).

Thought or emotion: a simple test

In everyday speech, "I feel" often comes in front of a thought. "I feel that he doesn't respect me." "I feel like a failure." Grammatically that's a feeling; in substance it's a judgement: you can check it against facts, agree or disagree with it.

A working rule used in psychoeducation: an emotion can usually be named in one word (anxiety, sadness, shame, anger, joy), while a thought is a statement with content (interpretation; standard CBT practice, after Beck & Haigh, 2014, p. 16).

Phrase	What it is	Why
"I'm anxious"	emotion	named in one word
"I feel I'm going to mess it all up"	thought	it's a prediction that can be tested
"My heart is pounding"	bodily sensation	it's about the body, not a judgement
"I feel like a failure"	thought (probably with sadness or shame behind it)	"failure" is a judgement about oneself
"I left the party early"	behaviour	it's an action
"They're probably bored with me"	thought	an interpretation of other people's behaviour

Example **How it sounds in a session.** A client says: "I feel that nobody at work values me." The therapist clarifies: "That's an important thought, let's write it down. And what feeling is behind it: hurt, sadness, anger? How strong, from 0 to 100?" A composite example.

A thought isn't always words

A thought doesn't always come as a sentence. Sometimes it's a flash of a picture: the manager frowning over the report, say. CBT also uses imagery to find thoughts: the person is asked to go back to the situation in their mind. Beck and Haigh note that imagery can help people recognise thoughts **the author's own account; Beck & Haigh, 2014, p. 16**. In workshop recordings Beck demonstrated eliciting automatic thoughts in session through imagery and role-play **documented fact; Beck Institute, 2021**.

That's why the standard therapist's question isn't "what did you think?" but something wider: "What was going through your mind just then?" That way both words and images get caught.

Which thoughts matter

Not every thought needs attention. In a workshop recording Beck distinguishes thoughts loaded with emotion from thoughts that get in the way of goals **documented fact; Beck Institute, 2021**. The first kind are "hot": "I've done something wrong" with anxiety at 80 out of 100. The second can feel almost calm yet stall things: "I'll start when I'm in the mood." Both kinds get worked on; there's no point unpicking a "cold" thought like "I need to buy milk."

Common mistakes

Beginning clinicians and clients alike tend to make the same mistakes clinical consensus; not checked against a single primary source:

- mixing up thought and emotion ("I feel like a failure" gets written in the "emotion" column);
- taking a vague situation ("work in general") instead of one specific episode;
- arguing with the thought straight away, before the person feels understood;
- presenting the model as "you create your own feelings", which leaves the person ashamed.

Three floors under one thought

Why did Anna have this thought rather than "finally, we'll discuss the project"? According to the model, the answer lies deeper, in beliefs she may never have put into words.

The model describes cognition as several levels. Think of a house: upstairs are the rooms where situations happen, below is a floor of rules, and below that is the basement.

Top floor: automatic thoughts

These were the subject of the last chapter. They are what a person can notice and write down: "I've done something wrong."

Middle floor: intermediate beliefs

Intermediate beliefs are the rules, attitudes and assumptions that set out how a person should live. The term was introduced by Judith Beck, a psychologist, Aaron Beck's daughter and author of a widely used CBT textbook documented fact; Wenzel, 2012. Most of the time these beliefs go unspoken clinical consensus; Wenzel, 2012. They come in three kinds:

- **Rule:** a demand on oneself or others. "I must do everything perfectly."
- **Attitude:** an evaluative stance. "Making mistakes is unacceptable."
- **Assumption:** a conditional if–then link. "If I make a mistake, everyone will see what I'm really like."

Assumptions often have a positive and a negative side. Wenzel gives a student's example: "If I get all As, then I'm successful" and "If I get anything less than all As, then I'm a failure" clinical consensus; Wenzel, 2012. Rules like these are often rigid and set an impossible bar.

Intermediate beliefs go together with **compensatory strategies**: habitual behaviour that helps a person avoid running into a painful core belief, such as perfectionism, avoidance, people-pleasing or over-control the author's own account; J. S. Beck, 2020 — only the table of contents was checked; not verified.

The basement: core beliefs

Core beliefs are "fundamental, inflexible, absolute, and generalized" beliefs about oneself, others, the world or the future clinical consensus; Wenzel, 2012, citing J. S. Beck, 2011. They're short: "I'm incompetent", "I'm unlovable", "People are dangerous".

Judith Beck groups negative beliefs about the self into three broad categories the author's own account; as described by Osmo et al., 2018:

Category	How it might sound
Helplessness	"I'm weak", "I can't cope", "I'm trapped", "I'm worse than others"
Unlovability	"I'm undesirable", "I'll be rejected", "There's something wrong with me if nobody loves me"
Worthlessness	"I'm bad", "I'm worthless"

The phrasings in the table are illustrations, not questionnaire items.

Fact **Three categories turned out to be four.** In 2018 Brazilian researchers built a questionnaire on negative core beliefs and tested it on 1,083 adults. Judith Beck's three categories did not hold up as such: helplessness split into two, "helplessness/inferiority" ("I'm worse than others") and "helplessness/vulnerability" ("I can't protect myself"). Together with unlovability and worthlessness, that made four groups of beliefs about the self, plus a separate group of beliefs about other people proven by research; Osmo et al., 2018. It is one study on a non-clinical sample, so the conclusion is cautious: the categories are a useful working framework, not a law of nature.

Even CBT authors describe these categories differently. Wenzel names themes of lovability, adequacy (competence) and helplessness clinical consensus; Wenzel, 2012. For depression, Beck and Haigh list seeing oneself as "worthless, unlovable, incompetent, or helpless" the author's own account; Beck & Haigh, 2014, p. 13.

Myth vs fact **Myth:** everyone is one of three "types": unlovable, helpless or worthless. **Fact:** these are working categories for case conceptualisation, not personality types or diagnoses. One person can hold several core beliefs, and testing them on data produced four categories, not three proven by research; Osmo et al., 2018.

Deeper still: schemas

A **schema** is a relatively enduring internal structure that organises information: it shapes both what a person believes and what they notice in the first place. A core belief is the content of a schema **clinical consensus; Wenzel, 2012**. A schema works like a filter: someone with a schema of inadequacy will spot one mistake in a report and miss ten compliments. The "evidence" gathered this way strengthens the belief further **clinical consensus; Wenzel, 2012**.

Beck and Haigh describe several properties of schemas, including density and "charge": a schema activated often becomes denser, and a highly charged one is easily triggered **the author's own account; Beck & Haigh, 2014, p. 6**.

One important detail: in their model, as a disorder gets worse, conditional beliefs turn unconditional. "If I can't cope, I'm helpless" becomes simply "I'm helpless" **the author's own account; Beck & Haigh, 2014, p. 7**.

"Schemas" in Beck's cognitive therapy and "early maladaptive schemas" in Jeffrey Young's schema therapy are not the same thing; schema therapy is covered in Episode 12.

One thought on every level: Anna's example

What the whole "building" looks like in one composite example, and how therapy gets down to it.

The downward arrow

To get from an automatic thought down to beliefs, CBT uses the **downward arrow** technique. It was first mentioned by Beck, Rush, Shaw and Emery in 1979 and later elaborated by David Burns in 1980 **documented fact; Wenzel, 2012**. The principle: the therapist doesn't argue with the thought but accepts it for the moment and asks what it means, again and again, until there's nothing further beneath it **clinical consensus; Wenzel, 2012**.

Here's how it might go with Anna (a composite example):

1. "I've done something wrong" — *automatic thought*. Therapist: "Let's suppose that's true. What does it mean to you?"
2. "He'll tell me off, and he'll see I can't cope." "And if he sees that, what then?"
3. "If I make mistakes, it means I'm bad at my job" — *assumption*. "And what does that say about you?"
4. "If people see my mistakes, they'll realise what I'm really like" — *assumption*. "And what are you really like, according to that thought?"
5. "I'm incompetent. I'm not up to it" — *core belief* (category: helplessness).

This is often where a pause comes. Signs that the core level has been reached: the statement is short, absolute and about oneself; asking "what does that mean?" no longer adds anything new; the emotion clearly rises; the person recognises the belief as an old one clinical consensus; Wenzel, 2012; Manhattan CBT.

The whole building

Level	For Anna
Situation	"Come and see me at 5"
Automatic thought	"I've done something wrong" → anxiety 80 out of 100, tight chest, rechecking reports
Compensatory strategies	checks her work three times, never asks for help, takes on extra, stays late
Intermediate beliefs	rule "I must do everything perfectly"; attitude "Making mistakes is unacceptable"; assumptions "If I make a mistake, everyone will see what I'm really like" and "If I work harder than everyone, I won't be found out"
Core belief	"I'm incompetent"

Same situation, different basements

The same words from the manager touch different things in different people (composite examples):

Person	Chain of thoughts	Where the downward arrow leads	Category
Anna	"I've done something wrong"	"I'm incompetent"	helplessness
Marina	"He's angry with me" → "He doesn't value me any more" → "Sooner or later everyone leaves me"	"I'm unlovable"	unlovability
Oleg	"I've let the team down" → "I cause everyone problems" → "I just get in the way"	"I'm bad, I'm worthless"	worthlessness

Why this is done with a therapist

The downward arrow isn't used in a first session. Work on core beliefs usually starts once there's trust (a therapeutic alliance) and the person has learned to work with automatic thoughts **clinical consensus; Wenzel, 2012; Manhattan CBT**. The reasons are practical:

- sustained attention to core beliefs in session can intensify negative feelings, and symptoms can temporarily rise between sessions, so a plan is needed for what to do if things get harder **clinical consensus; Wenzel, 2012**;
- core beliefs are "almost never" changed after a single session **clinical consensus; Wenzel, 2012**;
- leading questions from the therapist ("so you see yourself as a failure?") can easily swap the client's wording for someone else's **clinical consensus; Manhattan CBT**.

> Important It isn't a good idea to go all the way down to the "basement" alone: you can end up face to face with "I'm worthless" with no support. If, while looking at your thoughts, you notice hopelessness or thoughts that life isn't worth living, that is a reason to seek help straight away; see If you need help now.

What's done with a belief once it's found

More detail is in Episodes 5 and 6. In brief, Wenzel describes more than ten approaches **clinical consensus; Wenzel, 2012**, among them:

- **Cognitive continuum.** A 0–100 scale from "completely incompetent" to "completely competent". Real people and facts are placed on it, and the global "I'm incompetent" becomes a question of degree.
- **Positive data log.** Each day, facts are noted that support a new, more accurate belief ("I handle most tasks well and sometimes make mistakes, like everyone").
- **Behavioural experiment.** A prediction is drawn from the rule ("if I ask a colleague to check my text, she'll decide I'm not up to it") and tested in real life.
- **Historical test.** Going through periods of life and collecting what supported the belief and what contradicted it.

An honest caveat: there's very little research on whether core belief work adds anything beyond working with automatic thoughts and behaviour **clinical consensus; Wenzel, 2012**. Wenzel's hypothesis, that such work consolidates gains and protects against relapse, remains a hypothesis (interpretation; Wenzel, 2012).

Body and behaviour: how a reaction keeps a belief alive

A reaction doesn't end with an emotion. What a person does next can quietly preserve the belief it grew from.

Safety behaviours

When you're scared, you want to do something to feel better: check, avoid, seek reassurance ("tell me it's fine"), distract yourself. Beck and Haigh call these safety behaviours: they reduce distress for a while but maintain the disorder over time **the author's own account; Beck & Haigh, 2014, p. 9**. The mechanism is simple: if all goes well, the person credits the safety behaviour rather than the absence of danger. The belief is never tested.

Anna's vicious circle

Anna's compensatory strategies work: she's praised for being reliable. But they have a cost, exhaustion. And each time she "gets away with it", her mind files it not as "I'm competent" but as "good thing I double-checked" (interpretation; the same logic as safety behaviours in Beck & Haigh, 2014). So her behaviour reinforces the assumption "if I work harder than everyone, I won't be found out", which in turn guards "I'm incompetent".

The body

The bodily side of a reaction is not just background. In panic, according to the model, a bodily sensation itself becomes the situation that gets read catastrophically: "my heart is racing, so it's a heart attack" **clinical consensus; not verified in these notes, covered in Episode 8**. The sensation then feeds the fear, and the fear feeds the sensation.

That's why CBT changes behaviour as well as thoughts: dropping safety behaviours and running behavioural experiments make it possible to test a belief through experience **the author's own account; Beck & Haigh, 2014**.

Self, world, future — and different disorders

Beck proposed that thoughts differ in content from one disorder to another. The data have borne this out only in part.

The cognitive triad

One of Beck's best-known ideas is the **cognitive triad** of depression: a negative view of oneself ("I'm a failure"), of the world and one's experience ("Nobody cares about me") and of the future ("It'll always be like this") (documented fact; Beck, 1967 — as summarised by The Open

University, OpenLearn course "Exploring depression", section 3.3; Beck's books were not opened). The third part, hopelessness, is studied separately as a risk factor.

The content-specificity hypothesis

Beck went further and suggested that each disorder has its own "script" of thoughts **the author's own account; Beck & Haigh, 2014**:

- depression: loss, failure, worthlessness, self-blame;
- anxiety: threat, danger, "I won't cope";
- panic: catastrophic readings of bodily sensations (documented fact: this is how D. Clark's cognitive model describes panic; Clark, 1986, *Behaviour Research and Therapy, 24*(4), 461–470; covered in Episode 8);
- OCD: conditional beliefs about contamination or responsibility, such as "If I touch an unfamiliar object, I may be contaminated", which as the disorder worsens becomes the unconditional "I am contaminated" **the author's own account; Beck & Haigh, 2014, p. 7**.

What testing showed

In 2001 R. Beck and T. Perkins pooled 13 studies. Depressive thought content was more strongly linked to depression than to anxiety, so it was specific to some degree. But anxious content was linked to both anxiety and depression. The different measures of thought content were also highly correlated with each other. The authors concluded that the findings were "generally inconsistent with the cognitive content-specificity hypothesis" **proven by research; Beck & Perkins, 2001**. R. Beck, the meta-analysis author, is not Aaron Beck.

That's not surprising, given how often anxiety and depression go together. In 2014 Aaron Beck and E. Haigh proposed a compromise, the **generic cognitive model**: the processes are shared across disorders (attention, interpretation, memory, safety behaviours), while the content of beliefs differs **the author's own account; Beck & Haigh, 2014**. In the same paper they describe disorders as exaggerations of normal adaptation: the difference between adaptation and disorder, in their words, is "largely quantitative" **the author's own account; Beck & Haigh, 2014, p. 3**. A bias towards noticing the bad, on their account, is useful in evolutionary terms because it catches real danger, albeit "at the cost of many false alarms" **the author's own account; Beck & Haigh, 2014, p. 4**.

> Myth vs fact **Myth:** negative thoughts in your head mean depression. **Fact:** everyone has negative automatic thoughts. The model describes a continuum from normal functioning to disorder, not two separate worlds **the author's own account; Beck & Haigh, 2014**.

Why beliefs lie dormant, then wake up

If someone believes "I'm unlovable", why aren't they low all the time? Cognitive theory's answer: beliefs can be latent.

Vulnerability plus stress

The **diathesis–stress model** explains a disorder as a combination of vulnerability (diathesis) and stress. In cognitive theory, the vulnerability is dormant dysfunctional beliefs and schemas that wake up in response to a fitting event: a break-up might wake "I'm unlovable", a failure at work "I'm incompetent" **the author's own account; Beck & Haigh, 2014; our examples**. According to Beck and Haigh, even after symptoms have gone a belief can stay latent and be reactivated if it hasn't itself changed substantially **the author's own account; Beck & Haigh, 2014, p. 20**.

The sad music experiment

The clearest test of this idea was carried out in 2006 by Zindel Segal and colleagues **proven by research; Segal et al., 2006**:

1. 301 patients with major depressive disorder were randomly assigned to antidepressants (152) or CBT (149).
2. Of those who reached remission, 99 people (40 after medication, 59 after CBT) went through a "sad mood provocation": for ten minutes they listened through headphones to an excerpt from Prokofiev's "Alexander Nevsky" played at half speed while recalling a time they had felt sad.
3. Before and after, dysfunctional attitudes were measured with the Dysfunctional Attitude Scale (DAS), a questionnaire on rigid rules and attitudes.
4. They were then followed for 18 months to see whose depression came back.

The results:

- After antidepressants, attitudes worsened markedly under sadness; after CBT they did not.
- The more thinking "sagged" under sadness, the higher the risk of relapse, meaning the return of depression.
- Among those whose attitudes worsened markedly after the sad mood, 69% relapsed. Among those whose attitudes barely changed, 30% did.
- Overall relapse rates after CBT (39%) and after medication (47.5%) did not differ significantly.

The point: after recovery, what matters more is not how someone thinks on a good day but how fast their thinking deteriorates under ordinary sadness. This is called **cognitive reactivity**. One limitation: the sample is small, 99 people **documented fact; Segal et al., 2006**.

Mindfulness-based cognitive therapy (MBCT) builds heavily on ideas about cognitive reactivity; it's covered in Episode 13. The UK's NICE guidelines recommend group CBT or MBCT to

prevent depressive relapse clinical consensus; NICE NG222, 2022, as summarised by Kendrick et al., 2022.

Important The data on mechanisms in this chapter and the next come almost entirely from depression. For anxiety disorders, OCD and PTSD the picture is different, and the conclusions can't simply be carried over.

Does CBT work because it changes thoughts?

The logic looks airtight: thoughts drive feelings, so therapy works by changing thoughts. Testing that took more than thirty years, and the answer is more interesting than "yes".

First, three terms. A **mechanism of change** is the reason a treatment helps. A **mediator** is what the treatment works through to affect the outcome, such as a change in attitudes. To count something as a mechanism, methodologists require a lot: a strong and specific association, consistent findings, the mediator changing before the symptoms, a dose–response link (more change, more effect), and ideally an experiment in which the mediator is deliberately manipulated clinical consensus; Kazdin, 2007, as described by Lemmens et al., 2016.

Evidence for: thinking changes, and that's linked to improvement

In DeRubeis and colleagues' 1990 study, 64 depressed patients received either cognitive therapy or medication alone. In the cognitive therapy group, changes by mid-treatment on three scales (attributional style, dysfunctional attitudes, hopelessness) predicted further reductions in depression; in the medication group they did not. The authors concluded that cognitive change plays a mediating role, but that the data do not support treating it as a "sufficient" mediator proven by research; DeRubeis et al., 1990.

Taking therapy apart

In 1996 Neil Jacobson's group ran a **component (dismantling) study**, comparing the full therapy with its parts proven by research; Jacobson et al., 1996. 150 outpatients with major depression were randomly assigned to one of three conditions:

1. **Behavioural activation**: getting back to activities that bring pleasure and a sense of meaning, with no direct work on thoughts;
2. behavioural activation plus work on automatic thoughts;
3. full cognitive therapy, including work on core beliefs.

The full version did no better, either at the end of treatment or at six months. What's more, behavioural activation and the automatic-thoughts version changed negative thinking and attributional style as much as full cognitive therapy did proven by research; Jacobson et al., 1996. At two years, among 137 participants, full therapy was no better than its components at

preventing relapse either proven by research; Gortner et al., 1998. All treatment was delivered by four experienced cognitive therapists, and the authors themselves note "a clear bias favoring CT", so any bias would have worked in its favour documented fact; Jacobson et al., 1996, abstract.

Later trials pointed the same way:

- **Dimidjian and colleagues, 2006:** 241 adults with major depression. Among the more severely depressed, behavioural activation did about as well as antidepressants, and both did better than cognitive therapy proven by research; Dimidjian et al., 2006.
- **COBRA, 2016:** 440 patients in the UK. Behavioural activation was delivered by junior mental health workers with no professional training in psychological therapy; CBT by psychological therapists. At 12 months behavioural activation was not inferior to CBT (a difference of 0.1 points on the PHQ-9 depression scale) and cost less. Many participants were also taking antidepressants proven by research; Richards et al., 2016; Richards et al., 2017.
- **A 2021 network meta-analysis** (45 trials, 3,382 participants): cognitive restructuring, behavioural activation and their combination in CBT did not differ in effectiveness for depression; all three beat usual care and waiting lists. The evidence on "pure" cognitive restructuring is the thinnest proven by research; Ciharova et al., 2021.

Everyone's thinking changes

In 2015 Cristea and colleagues pooled 26 randomised trials (2,002 patients). CBT moderately reduces dysfunctional thinking compared with control groups (g = 0.50; g is an effect size measure where 0 means no effect and 0.5 a medium effect), and the effect holds at follow-up (g = 0.46). But compared with other psychotherapies there was no significant difference (g = 0.17), nor compared with medication (g = 0.04; only four studies). The authors allow that dysfunctional thoughts may be "one more symptom" that shifts along with treatment proven by research; Cristea et al., 2015.

In a 2008 Canadian study, 130 depressed patients received CBT, interpersonal therapy or medication. When CBT was compared with interpersonal therapy, the data fitted a model in which change in attitudes explained CBT's effect. When CBT was compared with medication, the reverse model fitted better: symptoms fall first, and attitudes change afterwards proven by research; Quilty et al., 2008 — based on the abstract.

Improvement comes in jumps

In 1999 Tony Tang and Robert DeRubeis noticed that improvement often came not gradually but in a leap. A **sudden gain** is a sharp, lasting drop in symptoms between two consecutive sessions: at least 7 points on the Beck Depression Inventory and at least 25% documented fact; criteria as given by Norton et al., 2010. In archival data on 61 clients they found 29 such gains in 24 people. A single gain accounted on average for 51% of all the improvement in therapy, and most often came around session 5 of a 16-week course. In the session just before the gain, clients showed substantial changes in thinking

proven by research; Tang & DeRubeis, 1999, as described by Norton et al., 2010. For supporters of the cognitive model this was a strong argument.

The twist came later. A 2020 meta-analysis (50 studies, 6,355 participants) found that sudden gains predict treatment outcome (g = 0.68 at the end of therapy and 0.61 at follow-up) and are not confined to CBT. The authors call them "a ubiquitous phenomenon in psychotherapy" proven by research; Shalom & Aderka, 2020. So sudden gains support the idea that a shift in thinking can trigger improvement, but not that this route is unique to CBT (interpretation).

What reviews of mechanisms say

A 2016 review by Lemmens and colleagues examined 35 studies of mediators in psychotherapy for depression. Dysfunctional attitudes and negative automatic thoughts each turned out to be significant mediators of cognitive therapy in 4 of 7 studies. Only 12 of the 35 studies (34%) measured the proposed mechanism at least three times during treatment. None manipulated the proposed mechanism experimentally, and "not one study meets all criteria" proven by research; Lemmens et al., 2016. As early as 2007 Alan Kazdin used cognitive therapy for depression as an example: even for well-studied therapies, we can't yet explain with evidence how they work (interpretation; Kazdin, 2007).

The "do we need to challenge thoughts?" debate

In 2007 R. Longmore and M. Worrell put the question in their title: "Do we need to challenge thoughts in cognitive behavior therapy?" Their conclusion: component studies offer little evidence that cognitive techniques add to effectiveness (interpretation; Longmore & Worrell, 2007). Stefan Hofmann replied that equal effectiveness of behavioural and cognitive methods doesn't show that thoughts don't matter: behavioural methods may change thoughts too, by a different route (interpretation; Hofmann, 2008). Longmore and Worrell answered that Hofmann had misrepresented their main argument (interpretation; Worrell & Longmore, 2008). The debate remains open.

Beck and Haigh stick to their account: behavioural activation, in their view, also works through beliefs, because it gives people chances to test negative beliefs about themselves and the world the author's own account; Beck & Haigh, 2014, p. 18. Critics respond that if every improvement is explained by a change in beliefs, the model becomes hard to disprove (interpretation; Longmore & Worrell, 2007). Steven Hayes proposed another path: change not the content of a thought but one's relationship to it the author's own account; Hayes, 2004; that's Episode 13.

The case for cognitive work: durability

Hollon and colleagues' 2005 study looked at what happened after treatment in 104 people who had responded well. Depression returned in 30.8% of those who had completed cognitive therapy, in 76.2% of those switched from medication to placebo, and in 47.2% of those who stayed on medication. The difference between cognitive therapy and continued

medication was not statistically significant. The authors concluded that cognitive therapy "has an enduring effect that extends beyond the end of treatment" **proven by research; Hollon et al., 2005**. This is a finding about relapse prevention, not a recommendation to stop medication: decisions about medication are made with a doctor.

But here too the data disagree: in the long-term follow-up of Jacobson's group, full cognitive therapy protected against relapse no better than its behavioural component **proven by research; Gortner et al., 1998**.

Myth vs fact **Myth:** since behavioural activation works just as well, thoughts don't matter. **Fact:** behavioural activation changed negative thinking as much as full cognitive therapy did **proven by research; Jacobson et al., 1996**. Thoughts may change through behaviour. Two methods producing equal outcomes doesn't rule out that change runs through thoughts (interpretation; Hofmann, 2008).

So what follows: what's proven and what isn't

One table in place of long conclusions. Almost all the data come from depression.

Claim	What is known	Key studies
Interpretation is linked to emotion and behaviour	The basis of the model and of CBT protocols; fits clinical data well	Beck & Haigh, 2014
Depression has its own thought content	Partly supported	Beck & Perkins, 2001
Anxiety has its own thought content	Not supported: anxious content is linked to depression too	Beck & Perkins, 2001
Cognitive reactivity to sadness predicts relapse	Supported in one trial with a small sample	Segal et al., 2006
CBT changes dysfunctional thinking	Supported by meta-analysis	Cristea et al., 2015
This change is unique to CBT	Not supported: other therapies and medication change thinking comparably	Cristea et al., 2015; Quilty et al., 2008
Changing thoughts is the cause of improvement	Not proven: mixed data, no study meets all the criteria	Lemmens et al., 2016; DeRubeis et al., 1990
The cognitive component adds to the behavioural one (depression)	Not supported	Jacobson et al., 1996; Dimidjian et al., 2006; Ciharova et al., 2021
The effect of cognitive therapy lasts after it ends	Supported in a trial; some contradicting data	Hollon et al., 2005; Gortner et al., 1998
Sudden gains are linked to treatment outcome	Supported by meta-analysis, across many kinds of therapy	Shalom & Aderka, 2020

The most honest summary: "CBT works because it changes thoughts" has not been proven, but neither has "there's no need to change thoughts". What has been shown well is that thinking changes with successful treatment, and that one strong way to change a belief is to test it through experience, not just argue with it in your head (interpretation).

Myths about the cognitive model: a quick check

The myths from these notes and a few other common ones, in one table.

Myth	What the sources say	Label
CBT is positive thinking	CBT looks for an accurate, useful thought, not a pleasant one; an alternative thought can be unpleasant but realistic	**Clinical consensus**
You create your own feelings and are to blame for depression	The model includes predisposition, early experience and stress	**The author's own account (Beck & Haigh, 2014)**
"I feel like a failure" is a feeling	It's a thought, a judgement about oneself; sadness or shame may lie behind it	**Interpretation, standard CBT practice**
You can find a core belief in one exercise and "let it go"	Core beliefs are almost never changed in one session; discussing them can temporarily increase distress	**Clinical consensus (Wenzel, 2012)**
People come in types: unlovable, helpless, worthless	These are working categories, not types; the data show four, not three	**Proven by research (Osmo et al., 2018)**
It's proven that CBT works by changing thoughts	Not proven: no study of mechanisms meets all the criteria	**Proven by research (Lemmens et al., 2016)**
Since behavioural activation works as well, thoughts don't matter	Behavioural activation changes thoughts too; the question is open	**Interpretation (Hofmann, 2008)**
Beck was the first to say thoughts affect feelings	The idea goes back to the Stoics, and Ellis developed it in parallel; Beck's contribution was an empirical programme of testing (see Episode 1)	**Interpretation**

Try it: four questions

A one-minute exercise. It's a first step in self-observation, not therapy, and it doesn't replace working with a professional.

Example **One situation, four questions.** Think of a moment this week when your mood noticeably dropped. Not the worst one, just a noticeable one. Ask yourself:

1. What was happening?
2. What went through your mind, word for word, the way you'd say it to yourself? Or was it a picture?
3. What did you feel, in one word, and how strongly, from 0 to 100?
4. What did you do next?

Don't argue with the thought and don't dig deeper. Just notice that between the event and the feeling there was a thought.

This is where psychoeducation in CBT begins: the therapist takes a recent, specific situation, asks what went through the person's mind, clarifies the emotion and its strength, bodily sensations and behaviour, and then draws the chain together with the client. The next steps (testing thoughts against facts, Socratic dialogue, thought records) are covered in Episode 5, and cognitive distortions in Episode 3.

Important **What not to do.** Don't use the exercise to diagnose yourself or others, and don't try to dig down to a core belief on your own. If the thoughts you find are heavy and won't let go, or if thoughts come up that life isn't worth living, don't carry on alone; seek help: see If you need help now.

Check yourself

1. **How does an automatic thought differ from a core belief?** An automatic thought is specific and tied to a situation ("I can't manage this report"); a core belief is absolute and general ("I'm incompetent").
2. **Is "I feel nobody values me" a thought or an emotion?** A thought: it's a judgement that can be checked against facts. The emotion behind it might be hurt or sadness.
3. **Name the three kinds of intermediate belief.** Rule ("I must..."), attitude ("Making mistakes is terrible"), assumption ("If..., then...").
4. **What did Segal's sad music experiment show?** Recovered people whose thinking worsened markedly under mild sadness relapsed more often: 69% versus 30% over 18

months.

5. **Why does Jacobson's 1996 study matter so much for the debate on mechanisms?** Full cognitive therapy did no better than its behavioural component, and behavioural activation changed negative thinking as much as the full therapy did.

Key takeaways

- Between an event and a feeling there is an interpretation; the reaction is made of emotion, bodily sensations and behaviour.
- A thought is a statement with content; an emotion can usually be named in one word; a thought can also be an image.
- Under automatic thoughts lie intermediate beliefs (rules, attitudes, assumptions) and core beliefs about oneself, others and the world.
- Safety behaviours and compensatory strategies bring relief but stop the belief from being tested.
- Beliefs can lie dormant and wake up under stress, even under mild sadness.
- Thinking changes with any successful treatment for depression; what exactly makes CBT effective is not yet proven.

Sources

1. Beck, A. T. (1963). Thinking and depression: Idiosyncratic content and cognitive distortions. *Archives of General Psychiatry, 9*, 324–333. (Cited via Beck & Haigh, 2014; primary source not opened.)
2. Beck, A. T. (1967). *Depression: Clinical, experimental and theoretical aspects*. (Not opened.)
3. Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press. (Not opened.)
4. Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology, 10*, 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734 · PDF
5. Beck, J. S. (2020). *Cognitive behavior therapy: Basics and beyond* (3rd ed.). Guilford Press. Book page
6. Beck Institute (2021). Grossman, H. Automatic thoughts in CBT (Part 2); Ways to elicit automatic thoughts. Beck Institute blog. Part 2 · Eliciting thoughts
7. Beck, R., & Perkins, T. S. (2001). Cognitive content-specificity for anxiety and depression: A meta-analysis. *Cognitive Therapy and Research, 25*, 651–663. https://doi.org/10.1023/A:1012911104891

8. Ciharova, M., Furukawa, T. A., Efthimiou, O., et al. (2021). Cognitive restructuring, behavioral activation and cognitive-behavioral therapy in the treatment of adult depression: A network meta-analysis. *Journal of Consulting and Clinical Psychology, 89*(6), 563–574. Abstract
9. Cristea, I. A., Huibers, M. J. H., David, D., Hollon, S. D., Andersson, G., & Cuijpers, P. (2015). The effects of cognitive behavior therapy for adult depression on dysfunctional thinking: A meta-analysis. *Clinical Psychology Review, 42*, 62–71. https://doi.org/10.1016/j.cpr.2015.08.003
10. DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology, 58*(6), 862–869.
11. Dimidjian, S., Hollon, S. D., Dobson, K. S., et al. (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. Abstract
12. Gortner, E. T., Gollan, J. K., Dobson, K. S., & Jacobson, N. S. (1998). Cognitive-behavioral treatment for depression: Relapse prevention. *Journal of Consulting and Clinical Psychology, 66*(2), 377–384.
13. 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.
14. Hofmann, S. G. (2008). Common misconceptions about cognitive mediation of treatment change: A commentary to Longmore and Worrell (2007). *Clinical Psychology Review, 28*, 67–70.
15. Hollon, S. D., DeRubeis, R. J., Shelton, R. C., et al. (2005). Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. *Archives of General Psychiatry, 62*(4), 417–422.
16. Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., Gortner, E., & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of Consulting and Clinical Psychology, 64*(2), 295–304. https://doi.org/10.1037/0022-006X.64.2.295
17. Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1–27.
18. Kendrick, T., Pilling, S., Mavranezouli, I., et al. (2022). Management of depression in adults: Summary of updated NICE guidance. *BMJ*. https://doi.org/10.1136/bmj.o1557
19. Lemmens, L. H. J. M., Müller, V. N. L. S., Arntz, A., & Huibers, M. J. H. (2016). Mechanisms of change in psychotherapy for depression: An empirical update and evaluation of research aimed at identifying psychological mediators. *Clinical Psychology Review, 50*, 95–107. PDF
20. Longmore, R. J., & Worrell, M. (2007). Do we need to challenge thoughts in cognitive behavior therapy? *Clinical Psychology Review, 27*(2), 173–187.

21. Manhattan CBT. How to use the downward arrow technique in CBT. https://manhattancbt.com/downward-arrow-technique/ (a clinic's practical resource, not a scientific source).
22. Norton, P. J., Klenck, S. C., & Barrera, T. L. (2010). Sudden gains during cognitive behavioral group therapy for anxiety disorders. *Journal of Anxiety Disorders, 24*(8), 887–892. Full text
23. Osmo, F., Duran, V., Wenzel, A., de Oliveira, I. R., Nepomuceno, S., Madeira, M., & Menezes, I. (2018). The Negative Core Beliefs Inventory (NCBI): Development and psychometric properties. *Journal of Cognitive Psychotherapy, 32*(1), 67–84. Article
24. Quilty, L. C., McBride, C., & Bagby, R. M. (2008). Evidence for the cognitive mediational model of cognitive behavioural therapy for depression. *Psychological Medicine, 38*(11). https://doi.org/10.1017/S0033291708003772
25. Richards, D. A., Ekers, D., McMillan, D., et al. (2016). Cost and outcome of behavioural activation versus cognitive behavioural therapy for depression (COBRA): A randomised, controlled, non-inferiority trial. *The Lancet*. Full report: Richards, D. A., et al. (2017). *Health Technology Assessment, 21*(46). Report
26. Segal, Z. V., Kennedy, S., Gemar, M., Hood, K., Pedersen, R., & Buis, T. (2006). Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Archives of General Psychiatry, 63*(7), 749–755. PDF
27. Shalom, J. G., & Aderka, I. M. (2020). A meta-analysis of sudden gains in psychotherapy: Outcome and moderators. *Clinical Psychology Review, 76*, 101827. Abstract
28. Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology, 67*(6), 894–904.
29. Wenzel, A. (2012). Modification of core beliefs in cognitive therapy. In I. R. de Oliveira (Ed.), *Standard and innovative strategies in cognitive behavior therapy*. InTech. CC BY 3.0. PDF
30. Worrell, M., & Longmore, R. J. (2008). Challenging Hofmann's negative thoughts: A rebuttal. *Clinical Psychology Review, 28*, 71–74.

If you need help now

If your life is in danger, call your local emergency number: 112 in most of Europe, 911 in the US and Canada, 999 in the UK, 000 in Australia. For support lines by country, see **Help now**.

Image credits

This page uses no third-party images.