Changing your Mind With Memory Reconsolidation

A Framework for Internal Debugging

When I'm debugging myself or others for unproductive behavioral, emotional, or mental patterns (hereby referred to as patterns) , I use this quasi-linear framework to make sure that I don't miss any critical steps and the issue stays solved. I've used this framework successfully with clients, housemates at Gentle Mesa, and hotel mates at The EA Hotel, as well as in debugging my own internal conflicts. I'm currently creating a GuidedTrack course that takes people through search of these steps and their associated techniques.

This model makes some assumptions about the underlying structure of the internal issue. It generally assumes that our patterns have internal reasons behind them, that those reasons are often coded as subconscious beliefs about the world or yourself, and that those beliefs sometimes cluster together to be understood as what we often refer to as "parts" or "subagents".

Finally, it assumes that when we have a pattern that we consider negative, one way to understand this is a tension between the part that is causing the pattern and the part that sees that pattern as negative (hereby referred to as anti-patterns). A successful debugging is one in which that tension is dissolved, rearranged, or in rare cases, accepted.

1. Awareness

The first step to solving an internal conflict is recognizing that you have it. In this stage, you use awareness of self-talk, mental imagery, and bodily sensations to notice that there's something "off".

Associated Techniques

- Mindfulness
- CFAR's Bug List Prompts
- CFAR's MurphyJitsu

2. Introspection

In this stage, you work to understand the cause of the pattern. You're trying to make explicitly conscious the beliefs that underlie the pattern. On the flipside, you're working to understand the cause of the anti-pattern, and the beliefs that cause the pattern to be seen as problem.

It's often useful here to not just understand the first level need being met (e.g. "I don't want to procrastinate because it lets me be more productive") but to find the "deeper cut" - The deeper need or fear that is being served (e.g. "If I'm not productive than I don't deserve anybody's love")

Finally, in this step, it's important to check for "second order constructions" - that is, do I have a belief that this tension needs to exist in order to meet another need. Do I

have a belief that if either of these beliefs are dissolved something bad will happen (beyond not meeting the original need the belief was obstensibly there for).

For example "I have to believe that my father is a nice person, because if I didn't the identity based on my childhood would crumble and that would be scary," or"It's important to strike a balance between being productive and relaxing so that I can see myself as well rounded. If I dissolve this tension, I'll stop being well-rounded."

Associated Techniques

- Gendlin's Focusing
- Leverage's Belief Reporting
- Pj Eby's Relaxed Mental Inquiry
- Coherence Therapy's Radical Inquiry

3. Acceptance

In the acceptance stage, you're working to truly integrate that you've been choosing this pattern over an alternative in order to get your needs met (based on the beliefs you introspected in the last step). You're also working to accept that this pattern is not malicious or something to be destroyed, but instead honor that it's based on it's wisdom and beliefs. On the flip side, you're looking to accept the deeper cut behind the anti-pattern, and the deep reasons that drove you to try to solve the problem in the first place.

There are two different flavors of acceptance. A more 'yin' acceptance in which you're accepting the beliefs you had as valid and embracing them with love, which can be hard for the anti-pattern. And a 'yang' acceptance in which you're taking ownership that you've caused the negative pattern in your life and chosen to continue it up to now, against the protests of the anti-pattern. Note that if you merely flipped which part was the pattern and the anti-pattern, the yin and yang methods flip as well.

A complete acceptance that doesn't cause problems in the future should include both types of acceptance, which as pointed out is symmetrical to an acceptance of both parts (the pattern and the anti-pattern).

Associated Techniques

- Coherence Therapy's Practice of Verbalizing Beliefs
- Acceptance and Commitment Therapy's Half Smile, Open Hands
- Neuro-Linguistic Programming's Future Projection
- Sedona Method's Three Questions

4. Memory Reconsolidation

In the memory reconsolidation stage, the goal is to find a way to get all needs met. This can involve the pattern and the anti-pattern both dissolving to create a third, new pattern. Or it can involve either the pattern or the anti-pattern dissolving in a realization they're not required to get their need met. Sometimes, it can involve an

add on belief that even though the pattern or the anti-pattern is not getting their need met, that's ok temporarily.

There are dozens of techniques I've cataloged for working different types of reconsolidation on different types of patterns encoded in different modalities. In fact, the alignment part of the process itself a process of going through a series of different types of techniques based on the properties of the patterns until you find one that fits.

A useful distinction to have is between techniques that use reconditioning (that is, slowly building up the new belief neuron pathways over time, so they it fire instead of the old one) with techniques that use reintegration (that is, they use the process of emotional memory reintegration to change the original pathway such that they match the new belief). Reintegration techniques should be a go-to as they are much faster and more permanent, but it's good to have both tools in your toolbox. I'll be writing a post up about this distinction and the book I learned it from (Unlocking the Emotional Brain) at some point.

Associated Techniques

- CFAR's Internal Double Crux
- Cognitive Behavioral Therapy's Thought Countering
- Bryon Katie's The Work
- Mark Lippman's Folding
- My Dialectical Reprocessing Technique

5. Completeness and Creation

This is really two steps, but it's rolled into one for the purposes of keeping the process linear. Now that you've made a new pattern through reconsolidation that aims to get all needs met, you need to make sure that you have everything you need to actually meet that need.

For instance, you started with a pattern that said you needed to be serious all the time to get work done, and a pattern that didn't like being serious all the time because it didn't help you socially. Your new pattern says that you can work in an easy, carefree joking way. However, when you use some of the "completeness checks" below, you realize that you're missing some of the internal and external resources needed to actually act easy, carefree and joking.

- You realize that even though you no longer have a self-image of yourself as serious, you still don't have an image of yourself as carefree.
- You realize that even though you imagine yourself joking around with this new belief, when you get higher resolution you don't actually have the skill to make jokes.

So, you use the creation step to install these new supporting beliefs, habits, and skills needed to actually get your needs met.

Associated Completeness Check Techniques

- CFAR's MurphyJitsu

- Neuro-Linguistic Programming's Future Projection.
- TRIZ's 9 Window Technique

Associated Creation Techniques

- CFAR's Trigger-Action Plans
- Steve Andreas's Self-Image Installation
- My Three Books Technique
- Neuro-Linguistic Programming's Swish Technique

6. Ecology

Now that you've made many changes internally, a final check of your whole life and environment is done to ensure that you haven't created conflicts and tension elsewhere in the process of resolving your old pattern and anti-pattern. You'll use a number of techniques to see consequences of the changes you made, and then you'll use your awareness techniques to notice any new conflicts that arise. If any do, you'll repeat the rest of the process.

Associated Techniques

- CFAR's Inner-Sim
- Neuro-Linguistic Programming's Perceptual Positions
- Timeline Therapy's Timeline Check

7. Integration

In the integration step, you'll take the new patterns you've created in the reconsolidation and creation steps. and integrate them in two ways. Firstly, you'll want to tie the new neuron pattern for the pathway to as many other paths as possible, lest you forget it. Second, you'll want to make the new pattern as deeply as possible tied to your most fundamental values, such that it becomes core to who you are.

Associated Techniques

- Cognitive-Behavioral Therapy's Mental Rehearsal
- Neuro-Linguistic Programming's Swish Technique
- Connirae Andreas's Parental Timeline Reimprinting
- Connirae Andreas's Core Transformation
- My Modality Integration Technique

Conclusion

This is a process that has been invaluable to me in helping myself and others overcome bothersome patterns. It gives you a framework to hang many techniques on such that you know when they're useful. And it shows you where things may not be

working for you if you're skipping or missing steps. I hope others can get some use from it.

Book summary: Unlocking the Emotional Brain

If the thesis in _Unlocking the Emotional Brain_ (UtEB) is even half-right, it may be one of the most important books that I have read. Written by the psychotherapists Bruce Ecker, Robin Ticic and Laurel Hulley, it claims to offer a neuroscience-grounded, comprehensive model of how effective therapy works. In so doing, it also happens to formulate its theory in terms of belief updating, helping explain how the brain models the world and what kinds of techniques allow us to actually change our minds. Furthermore, if UtEB is correct, it also explains why rationalist techniques such as Internal Double Crux [1 2 3] work.

UtEB's premise is that much if not most of our behavior is driven by emotional learning. Intense emotions generate unconscious predictive models of how the world functions and what caused those emotions to occur. The brain then uses those models to guide our future behavior. Emotional issues and seemingly irrational behaviors are generated from implicit world-models (schemas) which have been formed in response to various external challenges. Each schema contains memories relating to times when the challenge has been encountered and mental structures describing both the problem and a solution to it.

According to the authors, the key for updating such schemas involves a process of memory reconsolidation, originally identified in neuroscience. The emotional brain's learnings are usually locked and not modifiable. However, once an emotional schema is activated, it is possible to simultaneously bring into awareness knowledge contradicting the active schema. When this happens, the information contained in the schema can be overwritten by the new knowledge.

While I am not convinced that the authors are _entirely_ right, many of the book's claims definitely feel like they are pointing in the right direction. I will discuss some of my caveats and reservations after summarizing some of the book's claims in general. I also consider its model in the light of an issue of a psychology/cognitive science journal devoted to discussing a very similar hypothesis.

Emotional learning

In UtEB's model, emotional learning forms the foundation of much of our behavior. It sets our basic understanding about what situations are safe or unsafe, desirable or undesirable. The authors do not quite say it explicitly, but the general feeling I get is that the subcortical emotional processes set many of the priorities for what we want to achieve, with higher cognitive functions then trying to figure out _how_ to achieve it - often remaining unaware of what exactly they are doing.

UtEB's first detailed example of an emotional schema comes from the case study of a man in his thirties they call Richard. He had been consistently successful and admired at work, but still suffered from serious self-doubt and low confidence at his job. On occasions such as daily technical meetings, when he considered saying something, he experienced thoughts including "Who am I to think I know what's right?", "This could be wrong" and "Watch out - don't go out on a limb". These prevented him from expressing any opinions.

From the point of view of the authors, these thoughts have a definite cause - Richard has "emotional learnings according to which it is adaptively *necessary* to go into negative thoughts and feelings towards [himself]." The self-doubts are a *strategy* which his emotional brain has generated for solving some particular problem.

Richard's therapist guided Richard to imagine what it would feel like if he was at one of his work meetings, made useful comments, and felt confident in his knowledge while doing so. This was intended to elicit information about what Richard's emotional brain predicted would happen if it failed to maintain the strategy of self-doubt. The book includes the following transcript of what happened after Richard started imagining the scene as instructed:

> **Richard:** Now I'm feeling really uncomfortable, but-it's in a different way.

> **Therapist:** OK, let yourself feel it - this different discomfort. *[Pause.]* See if any words come along with this uncomfortable feeling.

> **Richard:** *[Pause.]* Now they hate me.

> **Therapist:** "Now they hate me." Good. Keep going: See if this really uncomfortable feeling can also tell you *why* they hate you now.

> **Richard:** *[Pause.]* Hnh. Wow. It's because... now I'm... an arrogant asshole... like my father... a totally self-centered, totally insensitive know-it-all.

> **Therapist:** Do you mean that having a feeling of confidence as you speak turns you into an arrogant asshole, like Dad?

> **Richard:** Yeah, exactly. Wow.

> **Therapist:** And how do you *feel* about being like him in this way?

> **Richard:** It's horrible! It's what I've always vowed *not* to be!

Richard had experienced his father as being assertive as well as obnoxious and hated. His emotional brain had identified this as a failure mode to be avoided: if you are assertive, then you are obnoxious and will be hated. The solution was to generate feelings of doubt so as to stop him from being too confident. This caused him suffering, but the prediction of his emotional brain was that acting otherwise would produce even *worse* suffering, as being hated would be a terrible fate.

UtEB describes Richard as having had the following kind of unconscious schema:

> **Perceptual, emotional and somatic memory of original experiences:** *his suffering from his father's heavily dominating, hyper-confident self-expression, plus related suffering from unmet needs for fatherly expressions of love, acceptance, understanding, validation. (This is the "raw data"; matching features in current situations are triggers of the whole schema.)*

> **A mental model or set of linked, learned constructs operating as living knowledge of a problem and a solution:**

> **The problem: knowledge of a vulnerability to a specific suffering.**
> *Confident assertiveness in any degree inflicts crushing oppression on others and is hated by them. I would be horrible like Dad and hated by others, as he is, if I*

asserted my own knowledge or wishes confidently. (This is a model of how the world is, and current situations that appear relevant to this model are triggers of the whole schema.)

The solution: knowledge of an urgent broad strategy and concrete tactic(s) for avoiding that suffering. *Never express any confident assertiveness, to avoid being horrible and hated (general strategy and pro-symptom purpose), by vigilantly noticing any definite knowledge or opinions forming in myself and blocking them from expression by generating potently self-doubting, self-invalidating thoughts (concrete tactic and manifested symptom).*

Emotional schemas can be brought to light during a variety of ways, including [Focusing](), [IFS](), and imagining yourself doing something and seeing what you expect to happen as a result.

But suppose that you do manage to bring up a schema which seems wrong to you. What do you do then?

Memory reconsolidation: updating the emotional learning

The formation of memory traces involves *consolidation*, when the memory is first laid out in the brain; *deconsolidation*, when an established memory is "opened" and becomes available for changes; and *reconsolidation*, when a deconsolidated memory (along with possible changes) is stored and becomes frozen again. The term "reconsolidation" is also used to refer to the general process from deconsolidation to reconsolidation; UtEB generally applies the term to mean the entire process. Unless the context indicates otherwise, I do the same.

UtEB reviews some of the history of memory research. Until 1997, neuroscientists believed that past emotional learning became permanently locked in the brain, so that memories could only consolidate, never de- or reconsolidate. More recent research has indicated that once a memory becomes activated, it is temporarily unlocked, allowing it to be changed or erased.

Starting from 2004, new studies suggested that activation alone is not sufficient to deconsolidate the memory. The memories are used to predict that things will occur in a similar fashion as they did previously. Besides just activation, there has to be a significant *mismatch* between what one experiences and what the memory suggests is about to happen. The violation of expectation can be qualitative (the predicted outcome not occurring at all) or quantitative (the magnitude of the outcome not being fully predicted). In either case, it is this prediction error which triggers the deconsolidation and subsequent reconsolidation.

The memory erasure seems to be specific to the interpretation from which the prediction was produced. For example, someone who has had an experience of being disliked may later experience being liked. This may erase the emotional generalization "I am inherently dislikable", but it will not erase the memory of the person also having been disliked.

Applied reconsolidation: an example of the schema update process

So, assuming that the model outlined above is correct, how does one apply it in practice?

From what we have discussed so far, the essential steps of erasing a learned belief (including an emotional schema) involves identifying it, activating it, and then finding a mismatch between its prediction and reality.

The first difficulty is that the beliefs involved with the schema are not necessarily consciously available at first. Richard knew that he suffered from a lack of self-esteem, but he was not aware of its reason. The process started from him describing in concrete details how this manifested: as skeptical self-talk during a daily meeting.

As he was guided to imagine what would happen if he didn't have those thoughts and acted confidently, his therapist was seeking to retrieve the implicit schema and bring it into consciousness so that its contents would become available for access. Once it had been retrieved, the therapist and Richard worked together to express the belief in the schema in maximally emotional language:

> "Feeling *any* confidence means I'm arrogant, self-centered, and totally insensitive like Dad, and people will hate me for it, so I've got to *never* feel confident, ever."

The authors have developed a therapeutic approach called Coherence Therapy, whose steps closely follow the steps of the memory reconsolidation process. The example of Richard is from this school of therapy.

In Coherence Therapy (as well as related approaches, such as Internal Family Systems), one initially avoids any impulse to argue with or disprove the retrieved schema. This would risk it being pushed away *before* it has become sufficiently activated to allow for reconsolidation.

Instead, one stays with it. Richard was given a card with the above phrase and instructed to review it every day until the next therapy session, just feeling the extent to which it felt true to him. This served to further integrate access to the schema in question, making it better consciously available.

Two weeks later, Richard had frequently noticed his self-doubt, used it as a prompt for reading the card, and experienced its description ringing true as a reason for his thoughts. When speaking with his therapist, he mentioned a particular event which had stuck in his memory. In a recent meeting, he had thought of a solution to a particular problem, but then kept quiet about it. A moment later, another person had spoken up and suggested the same solution in a confident manner. Looking around, Richard had seen the person's solution and confidence being received positively by the others. Richard had been struck by how that reaction differed from what his schema predicted would happen if *he* had made the same suggestion in that tone.

Because Richard had made the implicit assumptions in his schema explicit, he was able to consciously notice a situation which seemed to violate those assumptions: a prediction mismatch. His therapist recognized this as a piece of contradictory knowledge which could be used to update the old schema. The therapist then guided

Richard through a process intended to activate the old schema while bringing the contradictory information into awareness, triggering a reconsolidation process.

The therapist first instructed Richard to mentally bring himself back to the situation where he had just thought of the solution, but held it back. To properly activate the schema, the therapist guided Richard's attention to the purpose behind his reluctance and Richard's certainty of any confidence making him disliked. Next, the therapist told Richard to re-live what happened next: the other person making the same suggestion, and the other people in the room looking pleased rather than angry.

The book then has a transcript of the therapist guiding Richard to repeat this juxtaposition of the old schema and the disconfirming experience (italicized brackets in the original):

> **Therapist:** Stay with that. Stay with being surprised at what you're seeing—surprised because in your life, you've had such a definite knowing that saying something confidently to people will always come across like Dad, like an obnoxious know-it-all, and people will hate that. That's what you know, yet at the same time, here you're seeing that saying something confidently *isn't* always like Dad, and then people are fine with it. And it's quite a surprise to know that. *[That was an explicit prompting of another side-by-side experiencing of the two incompatible knowings, with the therapist expressing empathy for both, with no indication of any favoring of one knowing over the other. The therapist paused for several seconds, then asked:]* Does it feel true to describe it like that? Your old knowing right alongside this other new knowing that's so different?
>
> **Richard:** *[Quietly, seeming absorbed in the experience.]* Yeah.
>
> **Therapist:** *[Softly.]* All along, it seemed to you that saying something confidently could be done only in Dad's dominating way of doing it, and now suddenly you're seeing that saying something confidently can be done very differently, and it feels fine to people. *[This was another deliberate repetition of the same juxtaposition experience.]*
>
> **Richard:** Yeah.
>
> **Therapist:** Mm-hm. *[Silence for about 20 seconds.]* So, how is it for you be in touch with both of these knowings, the old one telling you that anything said with confidence means being like Dad, and the new one that knows you can be confident in a way that feels okay to people? *[Asking this question repeated the juxtaposition experience yet again, and, in addition, the "how is it" portion of the question prompted Richard to view the experience with mindful or metacognitive awareness, while remaining in the experience.]*
>
> **Richard:** It's sort of weird. It's like there's this part of the world that I didn't notice before, even though it's been right there.
>
> **Therapist:** I'm intrigued by how you put that. Sounds like a significant shift for you.
>
> **Richard:** Yeah, it is. Huh.
>
> **Therapist:** You're seeing both now, the old part of the world and this other part of the world that's new, even though it was right there all along. *[That cued the juxtaposition experience for a fourth time, followed by silence for about 30*

seconds.] So, keep seeing both, the old part and the new part, when you open your eyes in a few seconds and come back into the room with me. *[Richard soon opens his eyes and blinks a few times.]* Can you keep seeing both?

Richard: Yeah.

Therapist: What's it like to see both and feel both now? *[With the transformation sequence complete, this question begins the next step of verification— Step V— because it probes for whether the target learning still exists as an emotional experience.]*

Richard: *[Pause, then sudden, gleeful laughter.]* It's kind of funny! Like, what? How could I think that? *[This is an initial marker indicating that the pro-symptom schema may have been successfully disconfirmed, depotentiated, and dissolved by the transformation sequence.]*

Therapist: Do you mean, how could you think that simply saying what you know, or mentioning some good idea that you've had, would make you seem arrogant, insensitive and dominating like Dad and be hated for it?

Richard: *[Laughing again.]* Yeah!

Afterwards, the therapist and Richard wrote a new card together, which Richard was told to review daily:

All along it's been so clear that if I confidently say what I know, I will always come across as arrogant, insensitive, and dominating like Dad, and be hated for it. And it's so weird, looking around the room and seeing that it doesn't come across like that.

The purpose of the card was to provide additional juxtaposition experiences between the old schema and the new knowledge. While the original transformation sequence might have been enough to eliminate the old schema, the schema might also have been stored in the context of many different situations and contained in several memory systems. In such a situation, further juxtapositions would have helped deal with it.

In a follow-up meeting, Richard reported having lost the feelings of self-doubt, and that speaking up no longer felt like it was any big deal. To verify that the old schema really had lost its power, the therapist tried deliberately provoking his old fears again:

Dropping his voice to a quieter tone, the therapist added, "But tell me, when you have something to say and just say it, what about the danger of coming across as a know-it-all, like Dad, and being hated for that? What about your fear of that and how urgent it is to protect yourself from that?" […]

Richard took in the question, gazed at the therapist in silence for a few seconds, and then replied, "Well, I don't know what to tell you. All I can say is, that doesn't trouble me any more. And hearing you say it, it seems a little strange that it ever did—like, what was my problem?"

Applied reconsolidation: the schema update process in general

Now that we have looked at a specific example, we can look at a more general version of the process.

Accessing sequence

In Coherence Therapy, the *accessing sequence* is the preliminary phase of making both a person's implicit schema and some disconfirming knowledge accessible, so that they can be used in the juxtaposition process:

1. *Symptom identification.* Establishing which specific symptoms the person regards as problematic, and when and where they manifest. In Richard's case, the general symptom was a lack of confidence, which specifically manifested as negative self-talk in meetings.
2. *Retrieval of target learning.* Bringing into explicit awareness the purpose behind the symptoms. This can then be used to guide the search for disconfirming knowledge, as well as accessing the original schema in order to reconsolidate it. In Richard's case, the purpose was to avoid expressing confidence in a way that would make people hate him.
3. *Identification of disconfirming knowledge.* Identifying some past or present experience which directly contradicts the original learning. This knowledge does not necessarily need to feel "better" or "more positive" than the old one, just as long as it is mutually exclusive with the old one. In Richard's case, the disconfirming knowledge was the experience of his co-worker confidently proposing a solution and being well-received.

Erasure sequence

Once both the target schema and the disconfirming knowledge are known, the erasure steps can be applied to update the learning:

1. *Reactivation of the target schema.* Tapping into the felt truth of the original learning, experiencing it as vividly as possible.
2. *Activation of disconfirming knowledge, mismatching the target schema.* Activating, at the same time, the contradictory belief and having the experience of simultaneously believing in two different things which cannot both be true.
3. *Repetitions of the target-disconfirmation pairing.*

Something that the authors emphasize is that when the target schema is activated, there should be no attempt to explicitly argue against it or disprove it, as this risks pushing it down. Rather, the belief update happens when one experiences their old schema as vividly true, while *also* experiencing an entirely opposite belief as vividly true. It is the juxtaposition of believing X and not-X at the same time, which triggers an inbuilt contradiction-detection mechanism in the brain and forces a restructuring of one's belief system to eliminate the inconsistency.

The book notes that this distinguishes Coherence Therapy from approaches such as Cognitive Behavioral Therapy, which is premised on treating some beliefs as intrinsically irrational and then seeking to disprove them. While UtEB does not go further into the comparison, I note that this is a common complaint that I have heard of CBT: that by defaulting to negative emotions being caused by belief distortions, CBT risks belittling those negative emotions which are actually produced by *correct* evaluations of the world.

I would personally add that not only does treating all of your beliefs - including emotional ones - as provisionally valid seem to be a requirement for actually updating them, this approach is also good rationality practice. After all, you can only seek evidence to test a theory, not confirm it.

If you notice different parts of your mind having conflicting models of how the world works, the correct epistemic stance should be that you are trying to figure out which one is true - not privileging one of them as "more rational" and trying to disprove the other. Otherwise it will be unavoidable that your preconception will cause you to dismiss as false beliefs which are actually true. (Of course, you can still reasonably *anticipate* the belief update going a particular way - but you need to take seriously at least the *possibility* that you will be shown wrong.)

This can actually be a relief. Trying to stack the deck towards receiving favorable evidence would just also sabotage the brain's belief update process. So you might as well give up trying to do so, relax, and just let the evidence come in.

I speculate that this limitation might also be in place in part to help avoid the error where you decide which one of two models is more correct, and then discard the other model entirely. Simultaneously running two contradictory schemas at the same time allows both of them to be properly evaluated and merged rather than one of them being thrown away outright. I suspect that in Richard's case, the resulting process didn't cause him to entirely discard the notion that some behaviors will make him hated like his dad was - it just removed the overgeneralization which had been produced by having too little training data as the basis of the schema.

Of course, this means that there does need to be some contradictory information available which *could* be used to disprove the original schema. One might have a schema for which no disconfirmation is available because it is correct, or a schema which might or might not be correct but which is making things worse and cannot easily be disconfirmed. UtEB mentions the example of a man, "Tómas", who had a desire to be understood and validated by someone important in his life. Tómas remarked that a professional therapist who was being paid for his empathy could never fulfill that role. The update contradicting the schema that nobody in his life really understood him, would have to come from someone actually in his life.

Another issue that may pop up with the erasure sequence is that there is another schema which predicts that, for whatever reason, running this transformation may produce adverse effects. In that case, one needs to address the objecting schema first, essentially be carrying out the entire process on it before returning to the original steps. (This is similar to the phenomenon in e.g. Internal Family Systems, where objecting parts may show up and have their concerns addressed before work on the original part can proceed.)

Verification step

Finally, after the erasure sequence has been run, one seeks to verify that lasting change has indeed happened and that the target schema has been transformed. UtEB offers the following behavioral markers as signs that a learning which has previously generated emotional responses has in fact been erased:

- "A specific emotional reaction abruptly can no longer by reactivated by cues and triggers that formerly did so or by other stressful situations."

- "Symptoms of behavior, emotion, somatics, or thought that were expressions of that emotional reaction also disappear permanently."
- "Non-recurrence of the emotional reaction and symptoms continues effortlessly and without counteractive or preventive measures of any kind."

The authors interpret current neuroscience to say that only memory reconsolidation can produce these kinds of markers. They cannot be produced by counteractive or competitive processes, such as trying to learn an opposite habit to replace a neurotic behavior. Counteractive processes are generally fragile and susceptible to relapse. When these markers are observed in clinical work, UtEB argues that one may infer that reconsolidation has led to the original learning being replaced.

Further examples

For additional examples of the schema update process, I recommend reading the book, which contains several more case studies of issues which were dealt with using this approach. Here's a brief summary of the most detailed ones (note that some of these examples are actually more detailed and include additional complications, such as more than one symptom-producing schema; I have only summarized the most prominent ones to give a taste of them):

- "Charlotte". *Issue*: obsessive attachment to a former lover. *Schema:* "It would be much better if I was merged with my lover". *Contradictory knowledge:* The harm caused by not having boundaries.
- "Ted". *Issue:* an inability to hold a steady job and a general lack of success in life. *Schema:* "If my life is a mess, my father will be forced to admit how badly he screwed up as a parent." *Contradictory knowledge:* Realizing that Ted's father would never admit failure, no matter what.
- "Brenda". *Issue:* stage fright when having a leading role in an upcoming play. *Schema:* being on the stage in front of an audience means being unable to get off, causing helplessness similar to when Brenda was in a car with her alcoholic father and couldn't get off. *Contradictory knowledge:* re-imagining the scene and the way how Brenda *could* actually have gotten out of the car.
- "Travis". *Issue:* inability to experience intimate emotional closeness in relationships. *Schema:* "Nobody will pay attention to how I feel or give me understanding for how I'm hurting. I don't matter, and I'm all on my own." *Contradictory knowledge:* the therapist's empathic presence and listening.
- "Regina". *Issue:* strong anxiety and panic during/after interacting with other people. *Schema:* "I'm acceptable and lovable only if I do everything *perfectly*." *Contradictory knowledge:* Regina's Uncle Theo loves her regardless of what imperfections she might have.
- "Carol". *Issue:* wanting to avoid sex with her husband despite feeling emotionally close to him. *Schema:* engaging in any sexuality means being overtly sexual and harming Carol's daughter, in the way that Carol was harmed by her mother's overt sexuality. *Contradictory knowledge:* once the schema was made conscious, it activated the brain's spontaneous mismatch detection mechanisms and started to feel silly.

As the last item suggests, sometimes just making a schema explicit is enough to start to dismantle it. The authors suggest that the brain has a built-in detection system which compares any consciously experienced beliefs for inconsistencies with other things that a person knows, and can spontaneously create juxtaposition experiences by bringing up such inconsistent information. They suggest that therapies which are

based on digging up previously unconscious material, but which do not have an explicit juxtaposition step, work to the extent that the uncovered material happens to trigger this spontaneous mismatch detection. (We already saw this happening with Richard - once his underlying schema had been made conscious, he was startled to later notice what seemed like a contradiction.)

One may note the connection to [the model in *Consciousness and the Brain*](#) that when some subsystem in the brain manages to elevate a mental object into the content of consciousness, multiple subsystems will synchronize their processing around that object. If the object is an explicit belief, then any subsystem which is paying attention to that object may presumably detect inconsistencies with that subsystem's own models.

Besides these case studies from Coherence Therapy, the authors also analyze published case studies from Accelerated Experiental Dynamic Psychotherapy, Emotion-Focused Therapy, Eye-Movement Desensitization and Reprocessing, and Interpersonal Neurobiology. They try to show how these cases also carried out a juxtaposition process, even if the theoretical frameworks of those therapies did not explicitly realize it. It is the claim of the authors that any therapy which causes lasting emotional change does it through reconsolidation. Finally, the book contains four essays from other therapists (using Coherence Therapy and EMDR), who analyze some of their own case studies.

Evaluating the book's plausibility

Now that we have looked at the book's claims, let's look at whether we should believe in them.

It is unclear to me how reliable the neuroscience results are; the authors cite a number of studies, but each individual claim only references a relatively small number of them.

On a brief look, I could not find any reviews or papers that would have directly made a critical assessment of the book's model. However, I found something that might be even better.

Behavioral and Brain Sciences is a respected journal covering subject areas across the cognitive sciences. BBS publishes "target articles" which present some kind of a thesis or review about a particular topic, together with tens of brief commentaries which respond to the target article, and a final response by the target article's authors to the commentaries.

In 2015, four prestigious (with a total of 500 published research articles between them) psychologists published a BBS target article, *[Memory reconsolidation, emotional arousal, and the process of change in psychotherapy: New insights from brain science](#)* (Lane et al. 2015). While the exact model that they outline has a number of differences from the UtEB model, the core idea is the same: that therapeutic change from a wide variety of therapeutic approaches, "including behavioral therapy, cognitive-behavioral therapy, emotion-focused therapy, and psychodynamic psychotherapy, results from the updating of prior emotional memories through a process of reconsolidation that incorporates new emotional experiences."

One interesting difference was that Lane et al. describe emotional schemas somewhat differently. In their model, the schemas form memory structures with three mutually integrated components: emotional responses, episodic/autobiographical memories, and semantic structures (e.g. abstract beliefs which generalize over the various incidents, such as the claim that "people are untrustworthy"). Any of these components can be used as an entry point to the memory structure, and can potentially update the other components through reconsolidation. They hypothesize that different forms of therapy work by accessing different types of components: e.g. behavior therapy and emotion-focused therapy access emotional responses, conventional psychoanalysis uses access to biographical memories, and cognitive behavioral therapy accesses semantic structures.
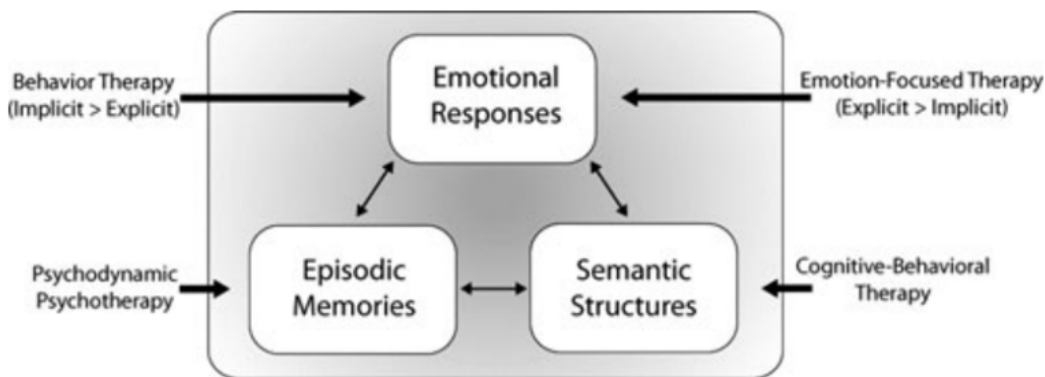
I read the target article, all the commentaries, and the responses. Given the similarities between Lane et al.'s model and the UtEB model, I think we can consider the responses to Lane et al. to generally offer a useful evaluation of the UtEB model as well.

One significant difference which needs to be noted is that Lane et al.'s model of memory reconsolidation does not mention the requirement for a prediction mismatch before reconsolidation can happen. This was remarked on in the response from UtEB's authors. In their counter-response, Lane et al. noted UtEB's model to be highly compatible with theirs, and remarked that further research is needed to nail down the conditions which make reconsolidation the most effective.

The other responses to Lane et al. were mostly from psychologists, psychiatrists, and neuroscientists, but also included the occasional economist, philosopher, philologist and folklorist. Several of the responses were generally positive and mostly wanted to contribute additional details or point out future research directions.

However, there were also a number of skeptical responses. A common theme which emerged from several concerned the limitations of the current neuroscience research on memory reconsolidation. In particular, most of the studies so far have been carried out on rats, and specifically testing the elimination of a fear response to electric shocks. As one of the responses points out, "neither the stimuli nor the subjects are generalizable to the kind of rich autobiographical memories involved in therapy." A number also raised the question whether all therapeutic change really involves reconsolidation, as opposed to some related mechanism, such as creating new memory structures which compete with the original as opposed to replacing it.

My non-expert reading is that the critical responses are right in that a gap remains between the clinical and behavioral findings on the other hand, and the neuroscience

findings on the other. There are various patterns which can be derived from psychological research and clinical therapy experience, and a small number of neuroscience findings which establish the existence of something that *could* explain those patterns. However, the neuroscience findings have only been established in a rather narrow and limited context; the connection between them and the higher-level patterns is a plausible link, but it remains speculative nonetheless.

Personally I consider the book's model tentatively promising, because it seems to explain many observations which I had independently arrived at *before* reading it. For example, I had noticed an interesting thing with anxieties, where I let e.g. a sensation of social anxiety stay active in my mind, neither accepting it as truth *nor* pushing it away while I went to do social things. This would then [cause the anxiety to update](), making me feel less anxious if it was indeed the case that the social interaction was harmless. This fits nicely together with the framework of an activated memory structure becoming open to reconsolidation and then being updated by a prediction mismatch (the situation not being as bad as expected).

Likewise, in my post [Integrating disagreeing subagents](), I reviewed a variety of rationality and therapeutic techniques, and suggested that they mostly worked either by merging or combining two existing models that a person's brain already had, or augmenting the existing models by collecting additional information.

In particular, I considered an example from cognitive behavioral therapy, where a man named Walter was feeling like he was impossible to be in a relationship with after he had broken up with his boyfriend. At the same time, he did not think that someone *else* breaking up with their partner was an indication of them being impossible to be in a relationship with. He and his therapist role-played an interaction where the therapist pretended to be a friend who had recently broken up, and Walter explained why this did not make the friend a relationship failure. In the process of doing so, Walter suddenly realized that he wasn't a failure, either.

I commented:

> Walter was asked whether he'd say something harsh to a friend, and he said no, but that alone wasn't enough to improve his condition. What did help was putting him in a position where he had to really think through the arguments for why this is irrational in order to convince his friend, and then, after having formulated the arguments once himself, get convinced by them himself.
>
> In terms of our framework, we might say that a part of Walter's mind contained a model which output a harsh judgment of himself, while another part contained a model which would output a much less harsher judgment of someone else who was in otherwise identical circumstances. Just bringing up the existence of this contradiction wasn't enough to change it: it caused the contradiction to be noticed, but didn't activate the relevant models extensively enough for their contents to be reprocessed.
>
> But when Walter had to role-play a situation where he thought of himself as actually talking with a depressed friend, that required him to more fully activate the non-judgmental model and apply it to the relevant situation. This caused him to blend with the model, taking its perspective as the truth. When that perspective was then propagated to the self-critical model, the easiest way for the mind to resolve the conflict was simply to alter the model producing the self-critical thoughts.

This seems like a straightforward instance of belief juxtaposition, and one where I ended up independently deriving something like UtEB's memory reconsolidation model: I too noted that the relevant belief structures need to be simultaneously activated in the right way to allow for the brain to revise one of them after noticing the contradiction. In general, UtEB's model of how things work rings true in my experience, making me inclined to believe that its description of how therapy works is correct, and that its model of how it is connected to neuroscience might also be.

UtEB and the subagent model

As many readers know, I have been writing [a sequence of posts](#) on multi-agent models of mind. In [Building up to an Internal Family Systems model](#), I suggested that the human mind might contain something like subagents which try to ensure that past catastrophes do not repeat. In [subagents, coherence, and akrasia in humans](#), I suggested that behaviors such as procrastination, indecision, and seemingly inconsistent behavior result from different subagents having disagreements over what to do.

As I already mentioned, my post on [integrating disagreeing subagents](#) took the model in the direction of interpreting disagreeing subagents as conflicting beliefs or models within a person's brain. [Subagents, trauma and rationality](#) further suggested that the appearance of drastically different personalities within a single person might result from unintegrated memory networks, which resist integration due to various traumatic experiences.

This post has discussed UtEB's model of conflicting emotional schemas in a way which further equates "subagents" with beliefs - in this case, the various schemas seem closely related to what e.g. Internal Family Systems calls "parts". In many situations, it is probably fair to say that this is what subagents are.

That said, I think that while this covers a very important subset of subagents, not *everything* which I have been referring to as a subagent falls straightforwardly under the belief-schema model. In [subagents and neural Turing machines](#) as well as [Against "System 1" and "System 2"](#), I also covered subagents in a more general way, as also including e.g. the kinds of subsystems which carry out object recognition and are used to carry out tasks like arithmetic. This was also the lens through which I looked at subagents in [my summary of Consciousness and the Brain](#). Which kind of view is the most useful, depends on exactly what phenomenon we are trying to understand.

A Practical Theory of Memory Reconsolidation

Memory Reconsolidation is one candidate for a scientific of theory of "How to Actually Change Your Mind." In this post, I'll give a few fake frameworks about how memory reconsolidation works, in order to provide intuition pumps for the rest of the sequence.

Schemas as Belief Clusters

For the purposes of this sequence, we'll describe a schema as a particular cluster of felt Aliefs that point towards a goal or set of goals.

I often find it useful to think of beliefs as arranged in a hierarchy of goals, as theorized in Perceptual Control Theory, Connection Theory, and Predictive Processing Theory.

Schemas, or Parts, then represent clusters of these beliefs trying to achieve some specific need.

In this framework, schema's don't represent any sort of natural boundary, and we can expand them and contract them at will based on what we're working on. Schemas themselves can overlap without natural boundaries, and can even contain their own internal conflicts.

3 Steps to Reconsolidate Schemas

In [Unlocking the Emotional Brain](#), the authors state three steps needed for memory re-consolidation:

1. Activate the Old Schema

2. Challenge the Old Schema

3. Learn a Replacement Schema

All of the techniques covered in this sequence implicitly or explicitly cover the three steps. However, step #3 is also augmented later on in the Debugging Process through the "Integration" step.

# The Tradeoff Between Activation and Challenge

There's an inherent tradeoff between activating the old schema and challenging it. The stronger and more perceptible the challenge is, the more likely it is to be strong enough to trigger the reconsolidation. However, this strength comes with two downsides:

1. Challenging yourself in a very strong way doesn't feel good. It feels internally violent.

2. The stronger the challenge, the more likely you are to deactivate the schema you're challenging, creating resistance and making reconsolidation impossible.

For this reason, we start with the techniques that provide the weakest challenges, and gradually work our way up to stronger and stronger challenge, waiting till we feel a shift towards a more nuanced and accurate schema.

4 Schema Access Points

In Kaj's Post on Unlocking the Emotional Brain, he points to a paper by Lane et al that give 3 ways to access a schema:
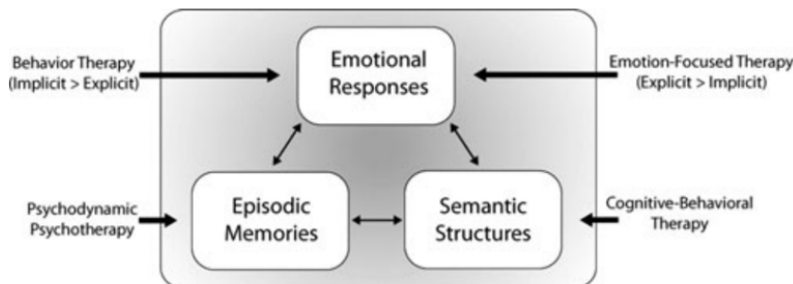
Because I think the names for these 3 methods to access a schema are overly long and technical, I'm going to call them Felt Sense (Emotional Responses), Belief (Semantic Structures) and Evidence (Episodic Memories). Based on my own experience with changework, I'll also add a fourth category, Metaphor (metaphorical representations of a schema).

Felt Sense involves looking at the physiological or indescribable "Feelings" evoked by a schema. An example is the tightness in your throat you get when trying to express

yourself.

Belief involves the semantic content of a schema as expressed through language. An example is "If I express myself, then I will be ridiculed."

Evidence represents our internal memories of how we learned the schema. An example is blurting something out in 4th grade and being laughed at.

Metaphor represents some novel/new way of representing the schema, as an object, location, situation, or story. An example is imagining that you're wearing a mask at all times, and if you take it off people will see the silly clown face underneath and laugh at you.

Practical Guidelines for Memory Reconsolidation

This post details a set of guidelines for working with the memory reconsolidation tools in the rest of the sequence. Use it to get the most out of your memory reconsolidation procedure.

Start with the More Cognitively Fused Schema

For every belief schema you're working with, there's (at least) two belief schema's at play. There's the side that believes a particular thing, and then there's a side that wants you to question the belief in that thing. As a general rule, you should always start with the side that's more [cognitively fused](#).

As an example, I was working with someone who was having issues going to bed on time, and wanted to change that. Before we started looking at the schema of "I should avoid ruminating by staying up late," We first examined the schema of "I should get more sleep."

By starting with the schema that you're more cognitively fused with, you avoid confirmation bias and end up with more accurate beliefs at the end.

The Resistance is the Way

If at any point, you encounter resistance to working on a particular technique with a particular schema, what you've found is a "Meta-schema" that believes changing this belief would be harmful. Rather than push through this resistance, loop back to the beginning of the Debugging process, and work with this new schema.

As an example, I found myself trying to change the schema that "I should avoid failure". I kept getting resistance, looped back, and found the schema "Most people should like me.", only once I worked on reconsolidating that schema was I able to return to the original schema.

Reverse Your Fusion

For any given technique, there's two ways you can approach it. You can work with the schema from "Inside" experiencing it as who you are, or you can work with it from the "Outside", putting some distance between yourself and the schema. As a general rule, I recommend reversing whatever your default is. If you frequently cognitively fuse with a schema, I recommend creating some distance/dissociation from it. If you frequently distance/dissociate from a schema, when doing the technique try as much as you can to fuse with it. This shift in perspective often goes a long way in providing new perspectives on the schema and allowing the re-consolidation to take place.

Go With the Most Salient Access Point

In contrast to the last point, you want to use whatever your natural inclination is as far as the schema access point. If a memory is coming up, use the evidence access point, if semantic content is coming up, use the belief access point, etc. Being able to fluidly switch access points to the schema as different things come up for you is key to quick reconsolidation.

Experiencing a Schema as True Allows for Updating

Via Kaj Satola

Something that been useful to me recently has been remembering that according to memory reconsolidation principles, experiencing an incorrect emotional belief as true is actually n*ecessary* for revising it. Then, when I get an impulse to push the wrong-feeling belief out of my mind, I instead take the objecting part or otherwise look for counterevidence and let the counterbelief feel s*imultaneously* true as well. That has caused rapid updates the way Unlocking the Emotional Brain describes.

I think that basically the same kind of thing (don't push any part out of your mind without giving it a say) has already been suggested in IDC, IFS etc.; but in those, I've felt like the framing has been more along the lines of "consider that the irrational-seeming belief may still have an important point", which has felt hard to apply in cases where I feel very strongly that one of the beliefs is actually just false. Thinking in terms of "even if this belief i*s* false, letting myself experience it as true allows it to be revised" has been useful for those situation

The Hierarchy of Memory Reconsolidation Techniques

As discussed in the previous post, the aim is to start with the techniques that least challenge existing schema, and then work up the hierarchy to techniques that more strongly challenge the existing schema.

Experience involves simply sitting with the schema as it is, and seeing if change comes from prolonged viewing of the schema.

Question involves posing questions about the schema, to draw attention to ways that it may be wrong.

Challenge involves directly providing counters to the schema, and seeing how it reacts.

Dissonance involves holding both the counter to the schema and the schema itself at once.

Change involves attempting to directly change how you represent the schema to take your counter into account.

Reconsolidation Through Experience

At this level, you're simply sitting with the schema. Oftentimes, this alone can produce a dramatic shift through memory reconsolidation simply as a matter of re-evaluating the old schema when you get a look at it. I also include non-memory reconsolidation techniques in this category that work through extinction of the old association.

Experiencing Evidence

Memory Collection

Memory collection involves finding and allowing yourself to experience every time you can remember when you felt or acted through a particular belief schema. For each piece of evidence, you allow yourself to think and feel through the schema in that particular memory, then move on to the next one. By simply allowing ourselves to see all the evidence for a schema with our current perspective, we can often come to a more nuanced understanding of the belief. I first learned about this technique through PJ Eby, who calls it the Feeling Elimination Technique. The book, thinking things done, that it was included in, is no longer available as far as I'm aware.

Imaginal Exposure

Imaginal exposure is a process of taking a particular salient memory/fear/trauma etc, and simply running through the experience over and over. This can often trigger memory reconsolidation, and, even if not, can change your schema through simple exhaustion induced extinction. **EMDR** is a version of imaginal exposure that adds the modality of moving your eyes back and forth as you run through the memory. My current best view of the evidence is that EMDR has no long term benefit over imaginal exposure, but the data is confusing.

Experiencing Beliefs

The Acceptance Statement

An acceptance statement is a tool from Coherence Therapy, that involves taking a belief, and the choices you've made about a belief and saying them out loud. For instance I could state out loud "I should not express myself. If I do, I'll be ostracized. I've decided it's more important to not be ostracized than to be fully seen." Saying this out loud is often enough to trigger the schema, and hearing out out loud is often enough to automatically trigger the questioning of the schema.

You can enhance this process by (in your imagination or in real life), saying the statement outloud to someone who has been affected by the choices you've made. I could imagine saying to my partner: "I should not express myself. If I do, you will

ostracize me. I've decided it's more important to not be ostracized than to be fully seen by you."

Emotional Freedom Technique is a process that combines saying acceptance statements out loud with tapping various parts of your body. It's unclear to me based on the evidence if the tapping adds anything outside of simply stating the beliefs.

Experiencing Felt Senses

Sitting With

A common recommendation is simply to spend some time sitting with your feelings. By simply sitting with the feeling, it gives you time to process it, allowing you to reconsolidate the schema with your new processed understanding.

Expressing

Another way to experience felt senses is to express them. **Focusing** is one process that allows you to express your felt senses, as are common forms of expressive therapy like Art Therapy.

Experiencing Metaphors

Clean Language

Clean Language is a series of questions you can ask yourself to start exploring your own metaphors. Simply by exploring our metaphors, we can often get a new perspective and reconsolidate them.

Reconsolidation Through Questioning

At this level, you're actively asking yourself questions about the correctness of the schema. You're not looking for any particular answers to these questions, or trying to get any result, you're simply holding the questions and seeing what comes up.

When coaching and teaching workshops, I find that questioning techniques are the most consistently successful at creating memory reconsolidation. They seem to strike the optimal balance of challenge and non-judgement.

They seem to work by actively directing our attention towards areas where the schemas may not match up with reality, without provoking any resistance by actively suggesting the schemas are wrong. For this reason, it's quite important that you don't actively try to "find answers" to these questions, as this starts to move into countering territory, and sticking at the questioning level can often work to change schemas that active countering cannot.

Questioning Evidence

The Lefkoe Belief Questions

The Lefkoe Belief Process is a process for finding different meanings for our evidence than the ones currently in our schema. Although the lefkoe belief process actually involves actively challenging the meaning, I've reworked it into a series of questions that simply allow you to question the evidence and draw your attention to ways that it might be interpreted differently.

The questions are:

1. What is this memory evidence of?

2. Is it possible that there are other interpretations of this memory?

3. What are some other possible interpretations of this memory?

4. How would my belief change if this memory no longer counted as evidence?

Remember, the goal is not to look for any specific answers, rather to simply hold these questions one by one in relation to the schema and see what comes up for you.

Questioning Beliefs

The Work of Byron Katie

The Work of Byron Katie is a method for questioning semantic beliefs, especially those related to "shoulds". The first part of the method involves a series of four questions related to your belief:

1. Is this true?

2. Can I be absolutely sure its' true?

3. How do I react, what happens when I believe that thought?

4. Who or what would I be without this thought?

Remember, the goal is not to look for any specific answers, rather to simply hold these questions one by one in relation to the schema and see what comes up for you.

Questioning Felt Senses

The Sedona Method

[The Sedona Method](#) is a 3 step process for questioning the need to hold on to a particular felt sense. Because the first step is asking one of 3 questions, you can repeat steps 2 and 3 for each question, for a total of 9 questions.

1. Ask yourself one of the following questions:

1a. Could I let this feeling go?

1b. Could I allow this feeling to be here?

1c. Could I welcome this feeling?

2. Would I? (AKA, Am I willing to let this go, allow it to be here, or welcome it).

3. When? (When am I willing to let go, allow it to be here, or welcome it).

Questioning Metaphors

I actually haven't found a good existing process for questioning metaphors, nor have I developed my own process. If you have done either of these things, let me know in the comments, and I'll edit the post!