Fighting against fear

Novel approaches to understanding, modifying, and manipulating maladaptive memories

Elsey, J.W.B.

Publication date
2020

Document Version
Other version

License
Other

Chapter 7.

General Discussion

This chapter expands upon issues raised in two publications:

Elsey, J.W.B., & Kindt, M. (2017). Tackling maladaptive memories through reconsolidation: From neural to clinical science. *Neurobiology of Learning and Memory*, *142*, 108-117.

Elsey, J.W.B., & Kindt, M. (2017). Breaking boundaries: Optimizing reconsolidation-based interventions for strong and old memories. *Learning & Memory*, *24*(9), 472-479.

Fundamental research into memory formation, storage, and modification has opened up novel avenues for understanding and tackling the maladaptive memories believed to underpin a range of disorders. Though research in animal models is hugely informative, harnessing the potential of procedures based around the modification or manipulation of memory requires the translation of such findings into human subjects in clinically relevant paradigms, and ultimately efforts at clinical intervention. The chapters contained within this thesis aimed both to directly investigate memory modifying procedures in human subjects, and to provide a theoretical and ethical framework within which such research can be situated. The general discussion that follows first provides a summary of the findings and perspectives presented in the previous chapters. I then consider some possible explanations for the null findings of Chapters 2 and 4. Finally, I consider how the diverse lines of research covered in each chapter may be synthesized in an overarching conceptualization of the therapeutic goals of a range of therapies that appear to affect memory.

Summary of chapters

In Chapter 2, I presented findings from a study in which we investigated whether propranolol could influence memory for an emotional movie compilation, when administered either before or after the movie. Three competing hypotheses were compared: that there would be no effect of propranolol on later memory, that only pre-encoding propranolol would affect later memory (i.e., propranolol affects encoding or a rapid consolidation process), or that both pre- and post-encoding propranolol could affect memory (i.e., interference with consolidation). Although some amnestic effects of propranolol could not be completely ruled out, we found most evidence for no effects of propranolol on intrusions, self-reported impact of the experience, and declarative memory one week later (these null findings are considered in detail in the section on *Null Findings* below).

However, we did find that propranolol seemed to have some influence on the initial emotional responding to the trauma film. This could have implications for both research and clinical practice. Firstly, research in which propranolol is given before an emotional experience or reactivation may be subject to similar effects, which should be assessed and controlled for. From a clinical perspective, it is also conceivable that propranolol could be administered in advance of predictable stressful events (such as an operation, if not contraindicated), allowing one to leverage possible

stress-reducing effects alongside any effects on memory that there might be. Effects on emotional experience and distress might also partially explain some of propranolol's reported benefits when given in the aftermath of trauma (Vaiva et al., 2003). I would, however, caution against interpreting the results too strongly. As noted in the discussion of that Chapter, the trauma film paradigm remains far-removed from true traumas. Even for experimental studies, to more fully understand potential effects of propranolol on emotional responses, a more highly powered study with a range of time distances from the emotional event would ideally be used to delineate more precisely when such effects begin to arise, and their magnitude.

In Chapter 3, I considered the ethical implications of moving beyond interference with memory *con*solidation to the pharmacological disruption of *re*consolidation. This article makes the case that use of reconsolidation-based treatment procedures sidesteps a number of concerns raised about consolidation-based prevention approaches (President's Council on Bioethics, 2003; Kolber, 2006). Notably, by focusing on those who have actually developed a disorder as a consequence of trauma, concerns regarding general pathologizing of bad memories, excessive or unnecessary medication for those who do not need it, and the preclusion of opportunities for 'post-traumatic growth' can be considerably reduced if not nullified. Reconsolidation-based approaches also expand the applicability of memory-modifying procedures beyond traumatic memories. I then aimed to lead an empirically informed discussion of some further issues that might arise from this novel therapeutic approach, such as the possibility of interfering with the wrong memories or removing one's declarative memory for important events, which I argued are unlikely prospects given what we currently know regarding the effects of current reconsolidation-based interventions. Moreover, I considered several potential misuses and abuses of memory manipulation, making the case that these were either unlikely to succeed, or that alternative and easier means could and perhaps already are being used for those purposes. Some of the null effects observed in the present thesis, though disappointing, may also serve to partially validate the arguments I made: even when trying hard and for all the right reasons to produce memory-modifying effects, we may well fail. Accidental memory erasure or the pursuit of far-fetched memory manipulation scenarios seem increasingly unrealistic. Although there is some potential for misuse and abuse of memory modification, clear instances could be met

with social or sometimes even legal sanctions, while we retain a focus on developing therapeutic applications.

Chapter 4 reflected an effort to extend pharmacological, reconsolidation-based approaches to a naturally occurring fear beyond spiders. Not only is public speaking anxiety a common and sometimes highly problematic fear among the general population (Bodie, 2010), but a successful new approach to tackling this fear might open the way to tackling more severe and generalized cases of social anxiety disorder, which can be extremely debilitating (Aderka et al., 2012). Practically, a standardized and replicable approach to tackling fear of public speaking would also open up the pool of subclinical participants with which we can investigate reconsolidation-based treatments, as the focus is currently on those with fear of spiders (many other fears we have investigated do not seem common enough to achieve a reasonable sample size). Participants who underwent our intervention did, on average, display reductions in speech-related distress during a public speaking task during the test session one week post-treatment, and also on questionnaire measures of public speaking anxiety three months after treatment. Unfortunately, anxiety reductions in those who received placebo were almost precisely the same as those who received propranolol, indicating a non-specific (i.e., not due to disrupting reconsolidation) effect of the intervention. These findings highlight the difficulties of clinical translation. I discuss some prospects and challenges related to these results in the section on *Null Findings* below.

In Chapter 5, my co-authors and I aimed to provide a comprehensive overview of the multiple lines of human experiments that come under the umbrella of reconsolidation-based research, spanning procedural, aversive/appetitive, and declarative memory types, and encompassing both pharmacological and behavioral interventions. Beyond reviewing the literature, we subjected each study to assessment regarding whether several criteria for consistency with reconsolidation had been tested and met, and considered various alternative explanations. We found that although research related to the concept of reconsolidation has been proliferating, relatively few studies included sufficient controls to provide the most convincing demonstrations that reconsolidation was a good candidate mechanism to explain the observed effects. Although this should not be taken to mean that reconsolidation does not underpin many of the effects reported in the reviewed studies, it does highlight that some of the enthusiasm for reconsolidation, and claims which seem to

increasingly extend the concept to encompass all manner of observed experimental and clinical effects (e.g., Gray & Liotta, 2012; Lee, Nader, & Schiller, 2017), may be premature. Although we suggest that researchers can unite around the goal of harnessing reconsolidation for clinical purposes, the clinical utility of the approach may ultimately depend on a precise understanding of the mechanisms underpinning any observed effects.

Finally, in Chapter 6, I presented findings from a series of experiments investigating the nature of irrational beliefs held by fearful individuals. We utilized a novel paradigm in which participants initially expressed their beliefs regarding the likelihood of several fear-relevant outcomes occurring, and then had to 'put their money where their mouth is' by betting on whether or not the events would actually occur. While highly fearful individuals expressed much higher initial beliefs in fear-related outcomes occurring than those with low fear, evidence for differences between high and low fear groups was much weaker when they were asked to bet on those outcomes occurring. Intriguingly, only among Low Fear participants did beliefs show a clear correlation with their later bets. For High Fear participants, beliefs were instead related to emotional responses. These findings were interpreted under a framework emphasizing competing adaptive and maladaptive representations of threat posed by feared stimuli. We suggest that fearful participants likely draw upon a highly emotive and easily retrievable negative representation of their feared stimulus. When motivated to consider what will *actually* happen to them during an exposure situation, many participants are quite capable of drawing on a more realistic representation of threat. However, this logical recognition may be largely divorced from their feelings, placing limitations on purely logical/verbal approaches to belief change.

Ongoing experiments aiming to further elucidate the nature of belief are currently underway in our lab. These involve assessing the possibility that fearful participants' bets may have been affected by providing an initial belief rating beforehand, and using a neutral betting task in which the objective probabilities of possible outcomes are controlled and clearly observable, which would allow us to determine what sort of underlying probabilities participants' bets might be taken to reflect. Further investigations that may also prove informative include having participants place bets on genuinely catastrophic outcomes occurring (the most catastrophic outcomes were typically not rated as the most likely, and so not selected

for bets). It would also be interesting to know when various adaptive and maladaptive representations might develop. Hence, investigating the developmental progression of irrational beliefs across various ages would be of considerable interest.

In the section below (*Integration of findings into a broader framework of memory competition*), I aim to synthesize our findings in relation to beliefs, and attempts at memory modification from our own and other research groups, in an overarching multiple representations framework, with a view to understanding therapeutic change across a range of therapies that appear to affect maladaptive memories. First, however, I discuss the absence of memory-modifying effects from propranolol in Chapters 2 and 4.

Null findings

As is now increasingly recognized in scientific circles (Ioannidis, 2005), and perhaps especially among psychological scientists (Francis, 2012), 'null' findings are important to publish and can be just as informative as reports of clear and striking effects. Still, in the absence of an effect, the researcher is often led to wonder whether their findings represent a true demonstration that something does not work, or rather that some choices they made or failure of experimental control/design might at least partially explain that absence. In the following sections I aim to reflect openly upon some ways in which the consolidation and reconsolidation interference studies (Chapters 2 and 4, respectively) might be improved.

Trauma film study

As discussed in the introduction and discussion section of Chapter 2, there are reasons why post-learning propranolol might *not* be expected to interfere with initial memory consolidation. Perhaps most notably, Debieç and LeDoux (2004) found that propranolol administered after fear conditioning – a laboratory analogue of a traumatic learning experience – did not interfere with subsequent memory. Using the same paradigm, Bush and colleagues (2010) found that pre-learning propranolol could interfere with subsequent memory, as it affected biological processes during acquisition, but that beta-adrenergic blockade after learning had no such effects. For these reasons, we were not necessarily expecting there to be a clear effect of post-learning propranolol on memory in our study. However, there is obviously not an absolute parallel between fear conditioning in rodents, real traumas in humans, or the trauma film paradigm, and so we cannot assume that post-learning consolidation blockade in humans is not possible. Researchers such as Pitman (Pitman, 1989) have

also argued that after a traumatic experience, repeated recollection of the event could result in further stress hormone/neurotransmitter release and serve to further cement the trauma in memory. This process might at least be diminished by post-trauma propranolol. In addition, researchers outside of the NYU-lab found that post-learning propranolol *could* affect subsequent memory expression, but selectively affected those animals that learned the task well (Cahill, Pham, & Setlow, 2000). The key outcome variables of the trauma film paradigm (intrusions, impact of events) do not lend themselves well to an immediate post-learning test, so we also remain ignorant as to whether our participants might be described as 'good learners', or what this might mean in the context of the trauma film paradigm.

Around the same time that our trauma film paper was released, another study using a very similar design found quite a convincing and unambiguous effect of post-learning propranolol on subsequent intrusions (Kamboj et al., 2019). Several differences with the study I have presented here may be noteworthy: Kamboj and colleagues only used women, and these women were all using hormonal contraceptives. There may be hormonal or sex differences in the impact of propranolol or of the traumatic event, and doses might be more likely to affect women than men given average differences in body mass. We have been assured by a consulting pharmacologist that body mass should not affect propranolol's effects in our studies and that a constant dose across participants is appropriate, but we have experienced conflicting views when speaking with other researchers. While we used a 40mg oral dose, Kamboj and colleagues used 80mg, which may also impact the results – irrespective of body mass.

Finally and perhaps most importantly, though I would again stress that the number of intrusions in our study did not seem arbitrary or unrelated to the participants' emotional experience of the trauma film, the participants of Kamboj and colleagues had a much greater number of overall intrusions, precluding floor effects. The film used by Kamboj and colleagues was a particularly graphic and disturbing rape scene, and may be especially unpleasant for young women to view (the victim is a young woman). The film used in our study was a compilation of several scenes and, though it has been frequently used to produce large numbers of intrusions in previous studies (James et al., 2016), it may be that it has become outdated, or that is does not resonate with international audiences. I have since heard from multiple researchers that they similarly failed to induce many intrusions at all using this standard

compilation of clips. A more emotionally engaging movie may be needed to produce the sort of effects seen by Kamboj and colleagues.

If I were to do this study again without restrictions, then I would consider several changes. Alongside generic improvements such as increasing the sample size, it would be informative to run full-sized groups of men and women separately, ideally controlling for circulating hormone levels (as argued convincingly by Shanksy 2019, this may be equally important among males and females). For consistency with Kamboj and colleagues, a dose of 80mg could be used instead of 40mg. While I would not wish to use the same scene as Kamboj and colleagues (even with consent and careful screening for victims of abuse, I can imagine this video would really affect how people feel in their daily lives, such as walking through a subway/underpass similar to the movie scene, for a long time after the study), I am sure that a more engaging film than the typical trauma film could be utilized instead. Finally, I would use a means of measurement other than a paper diary, such as time-stamped digital software. Despite these possible improvements and the alternative results of Kamboj and colleagues, it should be noted that diminished intrusions following propranolol in their study were already apparent on the first day of the manipulation, which might well suggest that the effect is not – or not only – one of blocking consolidation, as such effects would usually be disclosed after a delay.

Public speaking study

Regarding the failure to find an effect of administering propranolol on fear of public speaking, one could again consider tighter controls for participants' sex or hormonal status, as well as alternative drug doses or timing of administration. However, these factors probably do not explain what seems to be a clear null finding: A previous study using a similar approach in spider fearful participants, again with a predominantly female sample, without hormonal controls, and using 40mg of propranolol, found a large effect of a reconsolidation-based intervention (Soeter & Kindt, 2015a). One further possibility is that fear of public speaking may simply be unamenable to reconsolidation-based interventions. Though certainly possible, we should not immediately subscribe to this point of view, given that this was our first attempt at tackling this fear, and there are many ways in which reactivation might be varied so as to better target reconsolidation (Elsey & Kindt, 2017a). For discussion of this study, I will concentrate on the consideration of 'prediction error', it's possible

absence in the study, and the use of this concept in understanding reconsolidation-based research and its clinical translation.

As has been discussed (e.g., Chapter 5), numerous studies indicate that some form of prediction error – that is, the violation of learned expectations regarding feared outcomes – may be necessary for the induction of reconsolidation. Although prediction error was not explicitly controlled in Soeter and Kindt's (2015a) spider study, it can be imagined that being exposed to such a frightening spider and nothing objectively bad happening could provide participants with a clear prediction error. Common expectations among spider fearful individuals are that the spider will jump on or even attack them, and it might be speculated that the violation of these expectations could be a cause for the induction of reconsolidation in that study. Similarly to how prediction error was not explicitly incorporated into Soeter and Kindt (2015a), we reasoned that speech lengths of varying durations could provide ample opportunity for prediction errors to occur, as the speeches would rarely be catastrophically bad, and participants' expectations of how bad they would perform would still likely be violated. For example, even in the worst case scenario of 'blanking', there could be a long and quite awkward pause, but the panel would then simply ask some questions. In retrospect, it could instead be that participants' fears were partially or even largely realized. Firstly, the panel remained neutral which, in a social situation, might well be perceived as negative or judgmental. In addition, while never terribly bad, participants' speeches were rarely exceptionally good, and high self-criticism could make participants feel like they totally failed. The speech situation might also be too dissimilar to usual public speaking settings to properly reactivate their fears or generate their usual expectations. Perhaps, as suggested in the chapter, we could design an alternative reactivation in which there is a clear prediction error.

When considering initial designs for the study, we considered providing participants with some form of false feedback. The speech panel could at some point give unambiguously positive reactions, or praise the participant. However, this might corrupt participants' ratings of their stress/anxiety levels, as ending the speech on a clearly positive note could affect their retrospective assessment of earlier points in the speech before the feedback. To counter this, we could first have participants provide their own ratings for their performance, and then give them tailored feedback that suggests the judges rated them as having performed better than they believed. This

might be pursued in the future, but we felt it could also increase the chances of changes in participants' ratings at test that are not related to propranolol administration but rather just an updated sense of how harshly they would be judged by the panel. Although it did not appear necessary in Soeter and Kindt (2015a), another alternative would be to take account of participants' individual expectations and feared outcomes, and vary the precise procedure around that. Ultimately, I thought that it would be best to begin with the least complicated design that might work in principle, and add more complex elements if the approach did not work.

While I believe all these considerations are warranted and that, if we draw insights from laboratory studies of reconsolidation, prediction error really does appear to be important, we should also be careful that we do not mislead ourselves into thinking we know more about why particular reconsolidation-based interventions did or did not work than we really do. We do not know with certainty whether or what sort of prediction error participants in Soeter and Kindt (2015a) experienced, for example. Thinking of counterfactuals might also help convey how much we really understand prediction error in clinical settings. Let us imagine that the public speaking intervention *did* work. We could retrospectively imagine any number of violations of expectations that could hypothetically have triggered reconsolidation. Participants did not have panic attacks, only a few had clear periods where they totally ran out of things to say, no one fainted from fear, no one laughed at the speaker. Where participants' fears might have been partially validated, the magnitude of the outcome can also be a source of prediction error. They may have trembled, or been afraid, or stumbled over their words, but not as much as they anticipated. Prediction error can also be conceptualized as including continued learning when learning has not yet reached asymptote, and so novel situations in which something further may be learned could also be construed as providing a form of prediction error. All number of possibilities can be imagined, and it is not a great exaggeration to say that for almost any intervention design, whether it does or does not work, we could appeal to prediction error or its absence to explain the effects, or lack thereof.

A good explanation is hard to vary (Deutsch, 1997). As currently understood, prediction error is a very elastic explanation. By appealing to prediction error, the possibility of reconsolidation-based treatments becomes almost unfalsifiable, as we can always imagine an alternative procedure that we suggest should produce the desired effects. For the clinical utility and scientific validity of the concept of

reconsolidation and prediction error, therefore, it will pay dividends to get a more precise sense of what is meant by prediction error in clinical settings. A one-size-fits-all definition may be out of reach, but prediction error could be more strictly operationalized in clinical studies. For example, if a generic type of prediction error that would be expected to apply to most people is considered important, a condition with or without this could be tested. Alternatively, participants could provide idiosyncratic expectations, and be randomly assigned to reactivations in which these are or are not violated.

This may itself be difficult, as the research presented in Chapter 6 indicates that exactly how such questions are asked of participants might affect their expression of belief. Should the researcher focus on events that are catastrophic but not necessarily likely, or outcomes that participants rate as most likely, but which are less extreme? What of the sort of beliefs that participants express strong belief in but simultaneously recognize are not realistic? Finally, certain experimental designs geared towards testing these very explicit and high level 'cognitive' prediction errors can begin to sound increasingly like the kind of brief behavioral experiments advocated for in contemporary cognitive and behavioral therapies (Bennett-Levy et al., 2004; Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; McManus, Sacadura, & Clark, 2008). These are precisely the interventions that are proposed to work through the generation of new memory traces (Brewin, 2006), and not via reconsolidation. Highly explicit prediction errors such as one might expect in a behavioral experiment could also lead to changes in expectancies and even fear expression during reactivation, which are proposed to be boundary conditions on reconsolidation (Faliagkas, Rao-Ruiz, & Kindt, 2018). If some form of behavioral experiment is pursued, then placebo controls will also be increasingly important, as disproving a patient's core worries might produce strong treatment effects even without propranolol.

I have presented quite a dizzying number of tough questions and provided little in the way of solution to them here. A major improvement when relying on the idea of prediction error would, however, be to more explicitly incorporate it into designs as an experimental manipulation. While this may not capture everything that could be encompassed by prediction error and might not always extend to other fears or designs, it could at least be said whether or not a clearly operationalized definition of the concept led to one sort of effect or another in a particular experiment, rather

than relying on later speculation. Such studies could help narrow down which kind of prediction error related factors are actually important, refine our clinical intuitions, and perhaps incrementally get us closer to a more general understanding of the concept. Other members of the lab are working towards a more precise understanding of prediction error in laboratory studies, and I hope they can extend their insights to really explicitly consider the proper definition and role of prediction error in the clinic.

Integration of findings into a broader framework of memory competition

The studies and theoretical considerations presented in this thesis have developed out of groundbreaking research in animal models. While there remains considerable debate regarding whether reconsolidation is a legitimate phenomenon (Riccio et al., 2006), and what exactly is the neurobiological basis of memory (Gold, 2008, 2017), reconsolidation-based research has led to a dramatic shift in how the malleability of memory is understood. From the outset, researchers were also aware that their fundamental insights from model organisms into this malleability could have implications for treatment (Nader et al., 2000; Przybyslawski et al., 1999), and many contemporary researchers maintain this ambition (Exton-Mcguinness & Milton, 2018; Milton, 2019). The realization of these implications could lead to significant breakthroughs in the treatment of certain mental illnesses (Elsey & Kindt, 2017b). To reiterate a best-case scenario, reconsolidation-based treatments could allow for dramatic and long-lasting improvements in symptoms resulting from one therapeutic session, with the single administration of a commonly used drug. Although the practicalities of this approach are radically different to typical pharmacotherapy or psychotherapy, I would like to consider how reconsolidation-based approaches, as well as the idea of multiple representations discussed in Chapter 6, may be integrated into a broader understanding of the goals of therapy across multiple treatment modalities.

Retrieval competition among adaptive and maladaptive representations/memories has been put forward by researchers and clinicians as an overarching framework within which to understand the effects of several therapies (Bouton, 2002; Brewin, 2006; Craske et al., 2014). According to this perspective, cognitive or behavioral therapies do not primarily operate through the direct modification of maladaptive representations, but rather create and/or strengthen adaptive representations that can then compete with or inhibit the maladaptive ones

(for the following discussion, I will interchangeably use the terms memory and representation. Representation may be a preferable term, as it avoids confusion as to whether we are only talking about very precise memories for particular events, or all kinds of stored information more generally). Findings from the field of reconsolidation suggest that we may be able to directly affect maladaptive representations, although I did not succeed in doing so with public speaking anxiety. Nevertheless, the outcome of a successful reconsolidation-based treatment can be understood under a framework emphasizing memory competition. Figure 1 provides a (certainly simplified) schematic representation of how different therapies might affect adaptive and maladaptive representations.

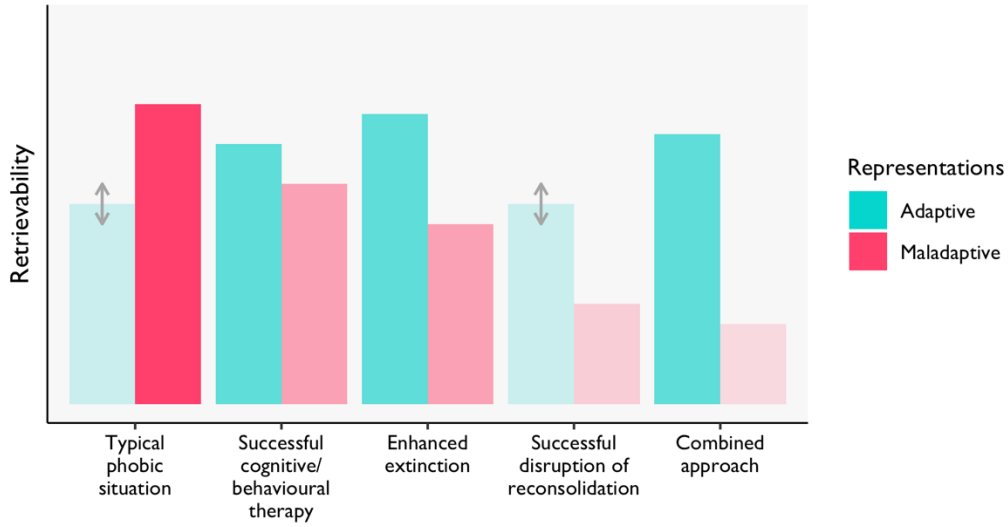

Figure 1. Memory competition in various treatment modalities. Color intensity reflects motivational value/salience of the respective representations.

Memory competition accounts typically emphasize *retrieval* competition and the relative retrievability of memories. Rather than retrieval *per se* as an end point, I would more explicitly highlight a representation's control over behavior as a further key factor of interest. I suggest that this is determined not only by that representation's retrievability, but also its motivational salience or value. This salience may reflect both certain qualities inherent in the representation (e.g., the feelings it elicits) as well as a person's metacognitive stance (e.g., have they been trained to pause and consider the validity of their emotional reactions?). Retrievability and motivational salience are almost certainly not orthogonal to one another, but I

hope to show that highlighting them separately can yield more insight than thinking solely in terms of retrieval competition.

For a phobic person encountering their phobic object or situation, maladaptive representations are highly retrievable, often even arising automatically. These representations may involve catastrophic cognitions, an intangible sense of imminent threat, or simply the expectation of a highly unpleasant emotional reaction, which provide a strong impetus for the person to avoid the situation. Individuals vary in the degree to which they harbor adaptive representations of the feared stimulus and their abilities to respond to it. For example, in our study of beliefs and bets, fearful participants varied quite substantially in the bets they placed on negative outcomes occurring, with some clearly recognizing that the outcome was unlikely. Yet, even if adaptive representations are accessible, this logical recognition alone may carry little weight in guiding a person's fear-related behavior when confronted with their fear - as Stott (2007) and others (Barnard & Teasdale, 1991; Bennett-Levy et al., 2004) have recognized. In my experience, some patients talk as if simultaneously accessing both an adaptive and maladaptive representation of the feared stimulus, with the logic of the adaptive one holding little sway over the urgency of their intense feelings, which provide the opposite message. Hence, not only the retrieval of a representation, but also the patient's inclination to respond to it, are key.

In typical cognitive therapy, the therapist and patient work together to reinforce existing adaptive representations of the stimulus, as well as to create new ones. Socratic questioning and mnemonic devices analogous to the 'reality cheque' may help patients retrieve or generate new, more realistic or otherwise adaptive representations. Through behavioral experiments and practice, these adaptive representations can become increasingly retrievable, and can outcompete or inhibit maladaptive ones. Owing both to the high personal relevance of evidence from behavioral experiments, but perhaps also because of enhancement of these 'success' memories caused by the emotionality of first-hand encounters, adaptive representations may also take on more motivational salience than the logical recognition of safety alone. Furthermore, the patient can be trained to reconsider what internal thoughts and feelings they *should* pay attention to and act upon, with encouragement to recognize emotional reasoning and focus on listening to their more reasoned deliberations. Relaxation and metacognitive strategies such as giving oneself a cool-down period or engaging in mindfulness when intense emotions are elicited

may further increase the distance between immediate reactions and maladaptive behavior. Hence, a cognitive approach may be seen to aim at increasing the amount and retrievability of adaptive memories and accordingly decrease the retrievability of maladaptive ones. Additionally, strategies can be employed to heighten or lessen the motivational value of adaptive and maladaptive representations, respectively, when they are retrieved.

A simple extension of this would be novel approaches to 'extinction enhancement', such as the administration of d-cycloserine (DCS) as a means of augmenting cognitive behavioral or exposure therapy (Davis, Ressler, Rothbaum, & Richardson, 2006; Rosenfield et al., 2019). DCS is proposed to enhance the consolidation of new, adaptive memories generated during exposure (Davis et al., 2006), and can thereby increase the odds that these memories are retrieved during a potentially fear-provoking encounter.

We next move on to reconsolidation-based interventions, should they prove successful. Firstly, disruption of reconsolidation may in some way ablate, scramble, or otherwise interfere with neuronal ensembles that represent the memory at the level of the brain, rendering it directly less retrievable. Alternatively, or in addition, the representation may remain retrievable but have in some way been denuded of its affective valence and concomitant motivational value. This would align with studies from our lab indicating that fear memory reactivation + propranolol can selectively neutralize the affect associated with a memory (Soeter & Kindt, 2012a), and the near-reflexive threat responses that memory would usually elicit, while leaving the explicit knowledge of what was learned intact (Kindt et al., 2009; Soeter & Kindt, 2010).

Cogan, Shapses, Robinson, and Tronson (2019) provided a striking demonstration of this possibility in relation to appetitive memories in rodents. Rats underwent a 'Pavlovian conditioned approach task', in which illumination of a lever predicted the arrival of a food pellet. Although engagement with the lever did not actually produce the reward, some rats ('sign trackers') appeared to imbue the lever with motivational properties and learned to approach and engage with it vigorously. Cogan et al. (2019) found that propranolol + reactivation seemed to selectively diminish the *motivation* of sign-tracking rodents to engage with the reward-related lever. These rats apparently retained a memory for the lever being related to reward, as they would still attend to it and sometimes even approach and lightly engage with it, but their behavior was much less energetic than those treated with saline

(importantly, such effects were not apparent without reactivation, or with a betablocker that does not cross the blood-brain barrier). While keeping in mind all the caveats regarding anecdotal evidence and selective recall in mind, a striking parallel can be seen in patients successfully treated with a reconsolidation-based approach, who sometimes make statements such as: "I still imagined that [the spider jumping] might happen, but it didn't bother me". A maladaptive representation may come to the patient's mind, but it is no longer linked with the motivational value that would make them act upon it. Such instances could be analogous to the initially weak adaptive representations that phobic patients hold.

Following on from a successful reduction in a patient's fear responses, the patient may, over time (and with experience of the absence of fear upon exposure), become increasingly confident in their lack of fear responses and newfound abilities to cope with their phobic stimulus. Adaptive self-representations might thus be developed or heightened to further reinforce positive effects. In addition, one highly important maintaining factor in avoidance – the knowledge that one will have highly unpleasant and even exhausting stress responses upon exposure – could be invalidated, with a new representation formed on the basis of new experience.

Rather than cognitive and reconsolidation-based approaches being opposed to one another, there are clear advantages to a combined approach. If a person were to have a successful reconsolidation-based treatment, then one could try to capitalize on the initial gains by conducting behavioral experiments and tests that might not have been possible without substantial time investment from patient and therapist under ordinary conditions. One patient, for example, was particularly afraid of having a mouse run across her feet. We reactivated the patient's fear by leading her to believe that she would have to enter an enclosure containing a mouse, in her bare feet (which did not ultimately occur) (Elsey & Kindt, 2017a). The patient was in tears at this prospect when standing outside the enclosure, but was then led away and given propranolol. Following this reconsolidation-based treatment, we were able not only to show that mice will usually actively avoid the patient's feet, but also what happens when the mouse is placed directly on them, and even in her hair. All this could be done in approximately half an hour, alongside many other tests, to solidify the patient's confidence. Given that it is unlikely that a patient's fear is entirely neutralized, and that we cannot be sure that it could not begin to resurface, it may also prove valuable to educate and train patients treated with such an approach in various

metacognitive strategies that can help keep residual fears in check, rather than feeling like they are relying upon some pharmacological trick.

Note that this way of understanding memory competition is not invalidated by models of retrieval-induced amnesia that differ from reconsolidation. For example, just like the reconsolidation hypothesis, the modified state-dependent retrieval (MSDR) account of retrieval-induced amnesia equally suggests that reactivation + an amnesic agent dramatically affects the retrievability of the memory (Gisquet-Verrier & Riccio, 2018). The key difference between MSDR and reconsolidation is that this deficit in memory retrievability is believed to be caused by later retrieval of the memory becoming highly contingent upon the presence of the amnestic agent as a retrieval cue. The diagram would simply need to incorporate an additional pair of bars for when the recipient is re-administered the amnesic agent, which ought to make the maladaptive memory trace highly retrievable again, similarly to how we understand that a change of context can affect the retrievability of an extinction memory (Bouton, 2002). Such alternatives might change our understanding of the precise mechanisms underpinning retrieval-induced amnesia, but we need not adapt the overarching model of how these techniques effect a therapeutic outcome to account for this.

Additionally, though the diagram and discussion has focused on phobias, it should be evident how this framework can apply to other disorders in which maladaptive memories are proposed to play a key role, such as in PTSD. In this case, as well as later modification of competing memories, successful interference with consolidation would hope to decrease the initial emotional/stress-induced enhancement of memory storage, thereby preventing the development of a highly retrievable and salient maladaptive representation in the first place.

Conclusion

Memories remain valuable and viable therapeutic targets. Although the precise approaches may vary considerably, numerous therapies – including the novel approaches examined in this thesis – may operate through their effects on adaptive and maladaptive representations stored in the brain, which compete for control over behavior. The key issue is how we can most effectively bolster adaptive representations, or attenuate maladaptive ones. Pharmacological interference with consolidation and (perhaps especially) reconsolidation hold the promise of being highly potent and rapid means of diminishing the negative impact of maladaptive memories. As my supervisor and I stated (Elsey & Kindt, 2017b, p.116), if these

promises are realized, and "if reconsolidation-based procedures become a viable treatment option, then they would be one of the first mental health treatments to have been derived directly from the translation of neural to clinical science. This would surely be a triumph for the scientific study of mind and brain". We are evidently quite far from achieving this goal, with the most convincing demonstrations of reconsolidation being in experimental models (i.e., fear conditioning) and impressive but difficult to generalize interventions for specific fears (e.g., Soeter & Kindt, 2015a). Yet, we also stated that "given the intricacy of the problem, researchers should not be overly discouraged if fledgling attempts at reconsolidation-based treatment are unsuccessful" (Elsey & Kindt, 2017b, p.115-116). Even seemingly 'simple' phobias are vastly more complex than a typical fear conditioning paradigm. We have many open questions, yet we do not have a production line of patients with whom we can rapidly get answers to them. It is therefore not surprising that we have yet to develop a reconsolidation-based intervention that can be readily extended to a host of disorders.

We have nevertheless gathered a wealth of experience in treating patients with a whole host of fears, and gained some insights from controlled studies that we hope can help us design more fruitful reconsolidation-based trials in the future. A greater appreciation of the syntheses that might be achieved through combining the various approaches to tackling maladaptive representations pursued in different therapies also paves the way for further research possibilities, and potentially improved clinical outcomes. Alongside more time-consuming trials of reconsolidation-based interventions, analogue studies in fearful individuals – such as in Chapter 6 – may not only help us gain insights into the nature of fear-related beliefs (and therefore perhaps prediction error), but could also reveal novel mnemonic strategies that could enrich reconsolidation-based and typical cognitive approaches alike. Fortunately, we are also not the only ones aiming to leverage the power of memory manipulation for the treatment of mental health disturbances. Efforts to move from the lab to the clinic are now being undertaken by several other research groups using a range of approaches. Some, for example, are using cognitive tasks (Iyadurai et al., 2018), others are focusing more specifically on PTSD with alternative timing and dosages of propranolol (Brunet et al., 2018), and still others are considering the use of completely different pharmacological agents for tackling reward-related maladaptive memories (Das et al., 2019). Views of researchers using these different approaches

frequently conflict, and resolving both technical and ethical issues as to the optimal ways of understanding, preventing, and treating mental disorders is vital. Nevertheless, we must remember that the real battle we are trying to win is that between the adaptive and maladaptive representations held by the many people suffering from mental health problems across the world. Concerted and clinically relevant efforts to understand how we can bolster adaptive memories, and attenuate maladaptive ones, will be the key to winning this fight.