STRENGTHENING MEMORY BY REPEATED LABILIZATION-RECONSOLIDATION EVENTS:
THE EFFECTS OF STRESS ON MEMORY RETRIEVAL

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Enmanuelle Pardilla-Delgado

________________________________________
Jessica D. Payne, Director
Graduate Program in Psychology
Notre Dame, Indiana
June 2017
STRENGTHENING MEMORY BY REPEATED LABILIZATION-RECONSIDERATION EVENTS:
THE EFFECTS OF STRESS ON MEMORY RETRIEVAL

Abstract
by
Enmanuelle Pardilla-Delgado

Memory consolidation was once thought to stabilize memories; that is, a memory was immune to alterations after the consolidation process had completed. Currently this view has changed, as reactivating a memory will make the memory labile once again, susceptible to being modified by amnesic and enhancing agents. The process that occurs after this reactivation-dependent labilization is called ‘reconsolidation’, and it is thought that when several labilization-reconsolidation events occur, a memory can be strengthened. The current study aimed to 1) replicate this memory strengthening effect and 2) test, in an ecologically valid way, the efficacy of this memory strengthening process by multiple reactivations. Stress has been robustly shown to be detrimental to memory retrieval. Therefore, we used the Trier Social Stress Test, a psychosocial stressor, at retrieval to probe how resistant memories become once strengthened by repeated reconsolidation. We partially replicated the strengthening
effect, as multiple (two but not four) reactivations resulted in better memory. However, no stress or strengthening effects were observed, likely due to low statistical power.
This is for everyone who ever believed in me, but mostly, and obviously, for my mother.
CONTENTS

Figures .................................................................................................................................. v

Preface .................................................................................................................................... viii

Acknowledgments ................................................................................................................... ix

Chapter 1: Introduction ........................................................................................................... 1
  1.1 Overview ......................................................................................................................... 1
    1.1.1 Memory consolidation ................................................................................................. 2
    1.1.2 Memory reconsolidation ............................................................................................. 5
      1.1.2.1 Memory reconsolidation in non-human animal models .................................. 7
      1.1.2.2 Memory reconsolidation in human models ......................................................... 12
        1.1.2.2.1 Reconsolidation of human episodic memory ............................................ 18
      1.1.2.3 Functions and theories of memory reconsolidation ......................................... 26
      1.1.2.4 Boundary conditions for memory reconsolidation .......................................... 30
    1.1.3 The current study ....................................................................................................... 35
      1.1.3.1 Stress and memory retrieval ............................................................................... 36
      1.1.3.2 Hypotheses of the current study ........................................................................ 39
        1.1.3.2.1 Experiment 1: Replication of Forcato et al., 2011 ..................................... 39
        1.1.3.2.2 Experiment 2: Testing the ‘memory strengthening by multiple reactivations’ effect against stress at retrieval ........................................ 41

Chapter 2: Method .................................................................................................................. 43
  2.1 Participants ....................................................................................................................... 43
  2.2 Materials .......................................................................................................................... 44
    2.2.1 Demographic Information Questionnaire ............................................................... 44
    2.2.2 Three-day Sleep Log ................................................................................................. 45
    2.2.3 Positive and Negative Affect Schedule (PANAS) .................................................. 45
    2.2.4 Self-Assessment Manikin (SAM) ............................................................................. 45
    2.2.5 State-Trait Anxiety Inventory (STAI) ........................................................................ 45
    2.2.6 Trier Social Stress Test ............................................................................................ 46
    2.2.7 Experimental Room ................................................................................................. 47
  2.3 Procedure ......................................................................................................................... 48
2.3.1 General procedure of experiments 1 and 2 ........................................ 48
2.3.2 Encoding (Day 1) ........................................................................ 52
2.3.3 Reactivations/ No reactivations (Day 2) ........................................ 54
2.3.4 Testing (Day 3) .......................................................................... 55
2.3.5 Memory measures and statistical analyses ................................. 56

Chapter 3: Results .................................................................................. 59
  3.1 Experiment 1: Replication study ................................................... 59
    3.1.1 Descriptive and control variables ........................................... 59
    3.1.2 Day 1 – Encoding ............................................................... 61
    3.1.3 Day 2 – Manipulation .......................................................... 63
    3.1.4 Day 3 – Memory test ............................................................ 64
    3.1.5 Change scores – Day 3 minus Day 1 ...................................... 66
      3.1.5.1 Change scores for specific types of errors – Day 3 minus
              Day 1 ............................................................................. 66
    3.1.6 End-of-encoding performance as a covariate ......................... 69
    3.1.7 Interim discussion for Experiment 1 ....................................... 70
  3.2 Experiment 2: Testing the strengthening effect against a stressor before
      memory retrieval ......................................................................... 72
    3.2.1 Descriptive and control variables .......................................... 73
    3.2.2 Day 1 – Encoding ............................................................... 74
    3.2.3 Day 2 – Reactivation manipulation ....................................... 77
    3.2.4 Day 3 – Stress manipulation and memory test ...................... 78
      3.2.4.1 Stress manipulation check ........................................... 78
      3.2.4.2 Memory test on Day 3 ................................................... 84
    3.2.5 Change scores – Day 3 minus Day 1 ...................................... 86
      3.2.5.1 Change scores for specific types of errors – Day 3 minus
              Day 1 ............................................................................. 87
    3.2.6 Interim discussion for Experiment 2 ....................................... 88
  3.3 Collapsed data for experiments 1 and 2 ......................................... 89

Chapter 4: Discussion ............................................................................ 91
  4.1 Summary of results and overview ................................................ 91
  4.2 Discussion of experiment 1 ......................................................... 92
    4.2.1 Mechanisms for reconsolidation, and its strengthening function .. 94
    4.2.2 Limitations of experiment 1 .................................................. 99
  4.3 Discussion of experiment 2 ......................................................... 101
    4.3.1 Limitations of experiment 2 .................................................. 105
  4.4 Conclusion .................................................................................... 106

Bibliography ....................................................................................... 108
Figure 1.1 A) Experimental procedure of Hupbach et al., 2007. B) Main findings of their original experiment: the memory updating effect is showed by intrusions of list 2 into the memory of list 1 (taken from Hupbach et al., 2007).

Figure 1.2 Graphical depictions of the hypotheses for A) experiment 1 and B,C) experiment 2. Note that the y-axis depicts mean number of errors, i.e., lower bars = better memory.

Figure 2.1 Graphical depiction of power analysis using G*Power

Figure 2.2 Graphical depiction of the experimental room

Figure 2.3 Graphical depiction of experimental protocol for experiment 2

Figure 2.4 Graphical depiction of the context period

Figure 2.5 Graphical depiction of the syllable period and the different possibilities of feedback a participant could receive

Figure 2.6 Graphical depiction of the timing of real and fake trials

Figure 2.7 Graphical depiction of the reactivation trial presented on Day 2

Figure 3.1 Encoding by trial. Error bars represent standard error of the mean (SEM).

Figure 3.2 Average of the last four encoding trials, by group. Groups 0 and four were significantly different by the end of encoding (t(35)=2.14, p=0.04). Error bars represent SEM.

Figure 3.3 Percentage of correct context responses during averaged through the 32 encoding trials. Error bars represent SEM.

Figure 3.4 Memory test at Day 3. Group R4 was significantly worse (i.e., more errors) at the first trial (T1), vs. Group R0 (p=0.07), vs. Group R1 (p=0.02), and vs. Group R2 (p=0.03). Error bars represent SEM.
Figure 3.5 Change score, constructed by subtracting the average of the last four encoding trials from the average of the four test trials, by group. Group R2 was the only group who performed significantly better at test (Day 3). Error bars represent SEM. ...................................................................................................... 67

Figure 3.6 Change score for void errors, constructed by subtracting the average of the last four encoding trials from the average of the four test trials, by group. Group R4 did significantly worse (i.e., more void errors than the other three groups (p’s<0.05). Error bars represent SEM. ................................................................. 70

Figure 3.7 Encoding by trial. Error bars represent SEM................................................. 75

Figure 3.8 Average of the last two trials of encoding. Error bars represent SEM. ........... 76

Figure 3.9 Percentage of correct context responses averaged through the 16 encoding trials. Error bars represent SEM................................................................. 78

Figure 3.10 Arousal ratings from the SAM. Control groups rated their arousal levels similarly before (pre) and after (post) the control-TSST (paired t’s n.s.). Arousal increased for the stress groups (paired t’s>2.69, p’s<0.03). Error bars represent SEM. ........................................................................................................ 80

Figure 3.11 Valence ratings from the SAM. Control groups rated their arousal levels similarly before (pre) and after (post) the control-TSST (paired t’s<1.92, p’s>0.10). Valence decreased (more negative) for the stress groups (paired t’s>2.83, p’s=0.02). Error bars represent SEM......................................................................................... 82

Figure 3.12 Salivary cortisol throughout the experiment for all groups (top) and collapsed by stress/control condition (bottom). Stress (R0S and R2S groups collapsed) increased cortisol immediately after the stressor (p = 0.01). Errors represent SEM.......................................................................................................................... 83

Figure 3.13 Memory test by trial. Error bars represent SEM. ........................................... 85

Figure 3.14 Change score from the end of encoding to test, by group. The Multiple Reactivations group performed significantly better than the 0 Reactivations group. Error bars represent SEM. .................................................................................................................. 90

Figure 4.1 Taken from Bonin and Konick, 2015 explaining Hu and Schacher, 2014. In Aplysia, reconsolidation was triggered only by stimuli that generated heterosynaptic but not homosynaptic plasticity. A) Initial stimulation causes LTP. B) The reactivation of the potentiated synapse alone (homosynaptic activation) does not necessarily trigger synaptic reconsolidation. C) Reactivation of several synapses, including the potentiated synapses (heterosynaptic activation), can trigger the depotentiation and repotentiation processes of synaptic
reconsolidation. D) Behaviorally, only stimulus that cause heterosynaptic activation will induce reconsolidation.
PREFACE

Memory is what makes humans, humans; like the Doctor said: “We’re all stories in the end”.

viii
ACKNOWLEDGMENTS

First, thank you to all those undergrads; RAs that helped me through the years and participants that made my research possible. Second, thank you Mike (AKA Dr. V) for virtually building the light box used in these experiments by yourself; I was just there looking over your shoulder. Similarly, thanks to Chuck Crowell for letting me use a whole room in their lab. Thanks to my committee, Chuck, Kathy, and G.A. for sharing your wisdom with me. I will try my best to make you all proud.

Finally, I would like to thank my advisor, Jess, for all the wonderful mentoring that she has provided me these past six years, also, for the wine and the scotch, and the beer, and the great food. She will be greatly missed (I will get you those two NN pubs!). I would like to thank my colleagues, my friends, and my brothers, Tony Cunningham and Steve Mattingly who experienced (suffered) grad school along my side. Sara Alger deserves probably the biggest admiration, as a wonderful mentor, mother, and woman. Although I’ve known the new kids, Little t and Skim, for less than a year, they are part of this warm family. Thank you for the great memories, I will greatly miss you all.
CHAPTER 1:

INTRODUCTION

1.1 Overview

The classical view of memory consolidation indicates that after consolidation is terminated, a memory trace is stabilized; that is, the memory is virtually unchangeable (Dudai, 2004; McGaugh, 2000; Squire & Alvarez, 1995). Reconsolidation theory, on the other hand, states that after retrieval or reactivation, a stable memory enters a labile phase once again, and it is required to go through a similar (but different) stabilization process (Alberini, 2005; Tronson & Taylor, 2007). During this re-stabilization phase, coined reconsolidation, a memory is susceptible to being disrupted (e.g., Nader, Schafe, & Le Doux, 2000), updated (e.g., Hupbach, Gomez, Hardt, & Nadel, 2007), or even enhanced (e.g., Rodriguez, Horne, & Padilla, 1999). It has been proposed that if multiple labilization-reconsolidation events are induced, a memory is strengthened (see Forcato, Fernandez, & Pedreira, 2014). This study aimed to test the efficacy of this multiple-reactivation strengthening using an ecologically-valid psychosocial stressor during the retrieval of an already strengthened memory.
1.1.1 Memory consolidation

There has been over a century of memory research dedicated to what Muller and Pilzecker (1900) named Konsolidierung in their seminal studies using nonsense-syllable pairs. In a series of experiments, they found that recall of target syllables improved a few minutes after encoding, but that if new stimuli were presented during this time period, recall of the target syllables would be impaired—a phenomenon they called retroactive interference (as summarized in Lechner, Squire, & Byrne, 1999). They concluded that, rather than information being stored immediately, there is a post-training interval in which new memories persist in a fragile state before they are eventually consolidated.

It was not until the second half of the century that memory consolidation research gained momentum, with systematic studies of animal models of human retrograde amnesia (Glickman, 1961; McGaugh, 1966; see Squire, Clark, & Knowlton, 2001), the dual-trace theory of Hebb (1949), and Lashley’s conclusion that memories were distributed throughout the cortex (Lashley, 1950). Scientists were in search of the biopsychological foundations of what are now called synaptic (‘fast’) and systems (‘slow’) consolidation (Dudai & Morris, 2000). Muller and Pilzecker’s aforementioned work is a good example of fast consolidation, whereas the temporally graded retrograde amnesia of brain lesion patients (discussed below) is a good example of slow consolidation failure.

Synaptic consolidation is completed within minutes to hours following training and it produces a cascade of intracellular events, including changes in protein synthesis,
transcription factors, and gene expression (e.g., Alberini, Ghirardl, Metz, & Kandel, 1994; Bailey & Chen, 1983; Davis & Squire, 1984; O’Leary, Byrne, & Cleary, 1995). These changes culminate in long-term modifications of synaptic excitability, neurotransmitter release, and synaptic remodeling/growth, which are argued to make the memory trace more stable (Dudai, 2002).

Systems consolidation is thought to take days, months, or even years. This slower consolidation process involves long-term reorganization of a memory trace in which the brain regions (or systems) that were recruited to encode the information (e.g., medial temporal lobe structures such as the hippocampus) change as a function of the consolidation process. Is thought that during this process, the memory trace ‘relocates’ or spreads to other brain regions (e.g., neocortical areas), making the trace independent of the hippocampus (McClelland, McNaughton, & O’Reilly, 1995; Squire & Alvarez, 1995). Some evidence that supports system consolidation is provided by research with amnesic patients, who largely show good retrieval of remote memories, and gradually worse retention of more recent events (Burnham, 1903; Squire & Alvarez, 1995; Squire, Chace, & Slater, 1975). This so-called temporally graded retrograde amnesia is often a product of lesions to the hippocampus and related medial temporal lobe regions (e.g., Bayley, Hopkins, & Squire, 2003; Scoville & Milner, 1957).

Alternatively, temporally ungraded/flat retrograde amnesia (i.e., impaired recollection of events regardless of memory age) is observed in patients with neocortical damage in the lateral and anterior temporal lobe (see Squire et al., 2001). These findings support the classic view that the hippocampus (and related structures) is required for the
formation of long-term declarative memory and its initial storage; however, the role of the hippocampus is time limited, and storage is eventually consigned to the neocortex. Another piece of evidence supporting system consolidation comes from patients like Henry Molaison (H.M), who had most of his medial temporal lobe (MTL) removed bilaterally and his very short term recall of verbal information was not impaired (Corkin, 2002). More recently, neuroimaging studies with healthy human subjects have shown that activity in the medial temporal lobe decreases, while neocortical activity increases, as the age of the memory progresses (e.g., C. N. Smith & Squire, 2009; J. F. Smith et al., 2010). This, together with the successful recall of remote memories, suggests that deficits in amnesia are not a problem of encoding, short-term memory, or overall storage, but rather with consolidation.

Although this view of systems consolidation was widely accepted, it was still inadequate to explain certain findings. The presentation of certain reminder cues is able to bring back a memory in amnesic animals (e.g., Bradley & Galal, 1988; Miller & Springer, 1972) and humans (e.g., Warrington & Weiskrantz, 1970). Additionally, the notion that the hippocampus slowly becomes unnecessary for the storage of long-term memories was refuted by the multiple trace theory (MTT), which states that the hippocampal formation is necessary for the entire life of some types of memory (Nadel & Moscovitch, 1997). In the MTT model, the hippocampus and then the neocortex encode the memory trace in a distributed manner. Eventually, through the repetitive reactivation of the original trace, multiple, related traces are created, which facilitate the extraction of related information and hence the integration of new information into
pre-existing networks. As more traces are created, a memory becomes stronger because it depends less and less on a single region, requiring more extensive lesions to impair older memories.

In addition, the schema assimilation model (Tse et al., 2007) theorizes that systems consolidation can be accomplished at a faster rate than the standard consolidation model would predict. This model states that faster systems consolidation is possible if the individual has a pre-existing body of related information (i.e. a schema) in which the new information could readily be integrated into.

There is abundant evidence that supports memory consolidation, regardless of what theoretical viewpoint is taken: a new memory trace is transformed from a short-term (more labile form) into a long-term (more stable) unit, not immediately, but rather after a post-learning time interval (Hebb, 1949). However, the idea that this kind of process occurs only once has been strongly contested over the last 15 years (although some refuting evidence exists long before, e.g., Misanin, Miller, & Lewis, 1968): it is not so much the time after encoding that determines the stability of a memory, but whether the trace is active (retrieved or reactivated) or inactive (stored). A reactivated trace becomes once again labile and susceptible to modifications, requiring a re-consolidation (Sara, 2000b).

1.1.2 Memory reconsolidation

Memory reconsolidation is the process by which a fairly stable (i.e., previously consolidated) memory trace is re-stabilized after being made labile by retrieval or
reactivation. Although the term reconsolidation research is somewhat new, the first evidence for a reconsolidation-type process was reported in the 1960s by Lewis and colleagues (see Lewis, 1979). In their laboratory, rats were water deprived and learned to lick a water bottle; after this behavior was deeply-rooted, the rats were presented with a tone+footshock combination during licking, resulting in a cessation of licking when hearing the tone. The next day, the rats were presented with the tone only, as a “reminder”, and then the application of electroconvulsive shock through ear clips, which is a well-known amnestic agent and consolidation disruptor. Only those rats that received the “reminder” before the electroconvulsive shock showed decreased behavioral responding, i.e., less freezing to the tone (Misanin et al., 1968). Importantly, similar results were found in a similar passive-avoidance task, but this time the reminder+shock combination was administered six hours after training, by which time the memory was known to be resistant to a shock only (i.e., the memory trace was considered stable) (Schneider & Sherman, 1968). In another study, supporting evidence was found with systemic injections of the protein synthesis inhibitor anisomycin after a reminder, and only after a reminder (Judge & Quartermain, 1982), showing that protein synthesis is required for reconsolidation (as it is for consolidation). However, Dawson & McGaugh (1969) were unable to replicate the reactivation disruption effect using electroconvulsive shock in a paradigm virtually identical to Misanin et al., 1968. There were inconsistencies in the findings of different laboratories (e.g., Dawson & McGaugh, 1969; Mactutus, Riccio, & Ferek, 1979), which could help explain why there was not much more research in this topic until 2000.
1.1.2.1 Memory reconsolidation in non-human animal models

Research on reconsolidation began receiving attention again at the start of the 21st century and, as was the case for consolidation research, most of the early evidence for reconsolidation came from studies that disrupted this process. This was most commonly achieved by showing a memory impairment as a result of a disruptive manipulation after memory reactivation. Nader and colleagues (2000) trained rats on classical conditioning to associate a tone with a footshock, and after several tone+shock pairings, the animals responded to the tone with freezing behavior (a fear behavior in most mammals). Similar to the work with electroconvulsive shocks mentioned above, the application of anisomycin directly into the amygdala after fear conditioning training resulted in consolidation blockade. Nader et al. found that amygdala infusions of anisomycin after retrieval of the fear-conditioned memory resulted in freezing behavior reduction, indicating that protein synthesis in the amygdala is crucial for fear memory reconsolidation. Of critical importance is that anisomycin in the amygdala has no effect when infused without the reactivation of the fear memory. This report catapulted memory reconsolidation research into the field that is today.

Debiec, LeDoux, & Nader (2002) found similar results with contextual fear conditioning. This time, instead of microinfusions of anisomycin into the amygdala, the rats received these infusions in the hippocampus, which is necessary for the consolidation of this type of memory. Rats that received anisomycin after memory reactivation showed less freezing to the conditioned stimulus (CS, i.e., the context), showing that the hippocampus is necessary for reconsolidation of a contextual fear
memory. Interestingly, this effect was shown even when reactivation+anisomycin occurred 45 days later, when the memory is supposedly hippocampus-independent; this suggests that reactivation of the memory, and therefore reconsolidation, makes the memory trace hippocampus-dependent once again.

Reconsolidation is not simply triggered by simple retrieval of the consolidated memory. It has been suggested that a lack of reinforcement, or prediction error, is necessary for memory destabilization and subsequent reconsolidation to occur. Pedreira, Pérez-Cuesta, & Maldonado (2004) used a contextual conditioning paradigm with the crab *Chasmagnathus* and the protein synthesis inhibitor cycloheximide (CHX) to show this prediction-error requirement for reconsolidation. In their first experiment, crabs were placed in a learning context (CS) in which half were presented with a visually dangerous stimulus (unconditioned stimulus, US) every ~3 min (for ~50 min); the other half were placed in the learning context (controls) for the same amount of time with no US presentation. In this paradigm, the animal’s early response to the US is an escape response, whereas the animal’s late response is freezing. That is, if the CS-US association is well learned, then a good memory is represented by lack of movement. On day 2, the memory was reactivated by placing all animals in the learning context (CS) for 5 minutes (no US presentation). Two hours later, half of the animals were administered CHX and half were administered saline (SAL), resulting in 4 groups: trained-CHX, trained-SAL, control-CHX, and control-SAL. On day 3, memory was tested with a single presentation of the visual dangerous stimulus (US). No differences in performance were found between the control-CHX and control-SAL animals. Trained-
SAL animals showed a decrease in movement, showing effective retrieval of the CS-US associative memory, but trained-CHX animals showed no such decrease, suggesting a reconsolidation-disruption effect. In a second experiment, the procedure was similar with the exception that the reactivation treatment on day 2 included a CS-US presentation (i.e., placing the animals in the context for 5 minutes and a single visual danger stimulus presentation at the end of the 5 minutes), as opposed to a CS only presentation. CHX had no effect on memory, as shown by a decrease in movement as response to US presentation during the day-3 test. The authors argue that, in this second experiment, the treatment did not include a prediction error; that is, the animals were expecting an US presentation paired with the learning context (CS), and that is what they received. In the first experiment, the treatment included a mismatch between what the animals predicted and what was actually presented, resulting in labilization/reconsolidation of the CS-US associative memory.

The findings by Pedreira, Perez-Cuesta, and Maldonado (2004) are in agreement with the classic notion that learning (of Pavlovian conditioning) only occurs when the US event is ‘surprising’ to the animal (Kamin, 1968). Rescorla and Wagner (1972) provided a more formal model of associative learning that is well explained in their own words: “...organisms only learn when events violate their expectations. [...] expectations initiated by that [stimulus] complex and its component stimuli are then only modified when consequent events disagree with the composite expectation” (p.75). In other words, it is not the simple co-occurrence of a CS and an US, but the divergence between what has been already learned and what more can be learned on a given trial, that
guide prediction error and therefore associative learning. Similar findings in memory reconsolidation research that support the importance of this prediction error component have also been found in appetitive conditioning paradigms (e.g., Takahashi et al., 2009) and instrumental learning (e.g., Exton-McGuinness, Patton, Sacco, & Lee, 2014).

In a study by Exton-McGuinness and colleagues (2014), rats were trained to press a lever for food on a fixed ratio of 1-to-1 for 10 days (1 press=1 pellet). When a NMDA receptor antagonist was administered just before a contingency change to a variable ratio of 20-to-1 (in which a mean of 20 presses results in 1 pellet), reconsolidation disruption was observed. Interestingly, a contingency change to a fixed ratio of 20-to-1 (20 presses=1 pellet) did not result in any behavioral change, suggesting that it is the unpredictability of the variable ratio that triggered reconsolidation. Exton-McGuinness, Lee, & Reichelt (2015) argue, however, that these findings cannot be fully explained by the Rescorla-Wagner model, since the mean reinforcement in both the fixed and variable ratio contingencies is the same (i.e., 20 presses = 1 pellet), which should generate equal prediction error. The relevance of this expectancy/prediction error component is also debated in human memory research (e.g., Forcato et al., 2007; Sevenster, Beckers, & Kindt, 2013) and is discussed below.

By now, reconsolidation has been demonstrated in a variety of different species, ranging from the honeybee (e.g., Stollhoff, Menzel, & Eisenhardt, 2005) to humans (e.g., Walker, Brakefield, Allan Hobson, & Stickgold, 2003), although most research has been conducted with rodents (e.g., Nader et al., 2000). In animal models, the range of tasks is
also quite wide: from auditory and contextual fear conditioning to drug-related memories, to odor discrimination and beyond (for review, see Besnard, Caboche, & Laroche, 2012).

Most animal research has been conducted with the purpose of comparing the molecular and neuroanatomical differences and similarities of reconsolidation and consolidation (see Alberini, 2005; Tronson & Taylor, 2007). As mentioned above, it has been well established that protein synthesis is necessary for both consolidation and reconsolidation processes; this has been shown in studies with systematic injections of protein synthesis inhibitors (e.g., Pedreira, Perez-Cuesta, and Maldonado, 2004) and with localized infusions (e.g., Debiec et al., 2000). On the other hand, other studies have shown that consolidation and reconsolidation might require different brain regions and/or different molecular mechanisms. Taubenfeld, Milekic, Monti, & Alberini (2001) were able to transiently block the expression of the transcription factor CCAAT enhancer binding protein B (C/EBPB), important for memory consolidation (Alberini et al., 1994), in the dorsal hippocampus after inhibitory avoidance training, or after its reactivation (i.e., during reconsolidation). They found that expression of C/EBPB in the dorsal hippocampus is necessary for inhibitory avoidance consolidation but not for its reconsolidation. Other studies have shown similar effects for distinct brain regions in other learning tasks (e.g., Hernández, Zainos, & Romo, 2002; Salinska, Bourne, & Rose, 2004). Additionally, there is some evidence that supports the notion of shared anatomical function but different molecular processes. Lee, Everitt, & Thomas (2004) found that if hippocampal expression of brain-derived neurotrophic factor is blocked,
consolidation, but not reconsolidation, of contextual fear is impaired. Instead, if
hippocampal expression of the transcription factor zif268 is blocked, reconsolidation,
but not consolidation, of contextual fear is impaired. Therefore, consolidation and
reconsolidation are distinct processes, with some degree of molecular overlap and
sharing of some, but not all, brain areas and circuits.

1.1.2.2 Memory reconsolidation in human models

The first report to show evidence of reconsolidation in humans used a
procedural finger-tapping task. Walker et al., (2003) trained volunteers to press four
numeric keys using the fingers of the non-dominant hand, repeating a five-element
sequence as quickly and as accurately as possible, for a trial period of 30 seconds,
followed by 30 seconds of rest. They found that learning a second sequence just after
retrieval of the first sequence resulted in memory impairment for the first sequence.
Specifically, participants learned sequence 1 on day 1, and were tested for sequence 1
on day 2 (at this point, accuracy and speed for sequence 1 has improved, arguably due
to sleep-dependent consolidation during this 24hr time interval). Immediately after this
1st test, they learned sequence 2. In a 2nd test on day 3, accuracy of sequence 1
plummeted to levels below those found at 1st test (during day 2), showing that
reconsolidation (triggered by 1st test at day 2) disruption by sequence 2 learning actually
eliminated the sleep-dependent enhancement, bringing the memory strength back to
acquisition levels.
Some research has also been conducted using fear conditioning in humans. In a direct translation from rodents (i.e., Monfils, Cowansage, Klann, & LeDoux, 2009) to humans, Schiller, Monfils, Raio, Johnson, Ledoux, and Phelps (2010) found that extinction training can practically erase a conditioned fear memory, measured in this study by skin conductance responses to the conditioned stimulus (CS). In this study, participants learned an associative fear memory on day 1, pairing a colored square (CS) with a wrist shock, the unconditioned stimulus (US). On day 2, they received a single presentation of the colored square, the CS, as memory reactivation. Then, after 10 minutes or 6 hours, they received repeated presentations of the CS, as extinction training. At 6 hours the reconsolidation window has arguably closed (Duvarci & Nader, 2004), i.e., no reconsolidation blockade should be observed in this group. Spontaneous recovery was assessed on day 3 with another session of extinction. Return of fear was observed only in the group that received extinction 6 hours after reactivation; that is, extinction training during the reconsolidation window (10 min after reactivation) completely erased the CS-US fear memory. One year later, this effect was still present. In an additional experiment, participants learned two CS-US associations, but only one was reactivated on day 2. The reconsolidation blockade by extinction training was only observed with the stimulus that was reactivated on day 2, showing that this effect is dependent on the reactivation (reconsolidation) of the memory. Similar results have been reported by other laboratories (Agren, Furmark, Eriksson, & Fredrikson, 2012; Agren et al., 2012) and with aversive sounds instead of electric shocks as US (Oyarzún et
al., 2012; but for null results see Golkar, Bellander, Olsson, & Öhman, 2012; Kindt & Soeter, 2013).

The reconsolidation of fear memories has also been disrupted by pharmacological interventions. Brunet, Orr, Tremblay, Robertson, Nader, and Pitman (2008) used propranolol, a beta-adrenergic receptor antagonist known to decrease physiological responses of fear (Pitman et al., 2002), to disrupt the reconsolidation of traumatic experiences. Individuals with post-traumatic stress disorder (PTSD) described the traumatic event that caused their disease with the purpose of reactivating the traumatic memories. After this, participants were given oral propranolol or placebo in a randomized, double-blind manner. Personalized recorded scripts of the trauma memories were developed by the research team, and a week later these were played back to the respective individuals while physiological responses were being measured. Participants that received propranolol a week before had a lower physiological response to their mental imagery of their traumatic experience, as measured by lowered SCR and heart rate (but not electromyography).

Other studies have shown similar patterns in fear conditioning with humans. It is noteworthy to point out that this type of learning is directly dependent on amygdala functioning (LeDoux, 2000), specifically on adrenergic activity of the amygdala, as administration of propranolol has been shown to disrupt fear memory consolidation (see McGaugh, 2004) and reconsolidation (Dębiec & Ledoux, 2004; Kindt, Soeter, & Vervliet, 2009). Kindt and colleagues (2009) had participants learn that a picture of a spider (CS) was associated with a mild electrical shock to the wrist (US) on day 1. At Day
2, the fear memory was reactivated by presenting the CS only after propranolol was administered. On day 3, memory was tested with extinction training (i.e., multiple CS-only presentations) followed by reinstatement of the fear memory, i.e., presentation of the US should bring back the fear memory after extinction. Only the combination of propranolol+reactivation resulted in a complete erasure of fear memory, as represented by no fear expression (measured by eyeblink startle response) throughout extinction training and, more impressively, no return of fear by reinstatement. This effect has been shown when propranolol is given 90 min before reactivation (to assure peak plasma levels at reactivation; Kindt et al., 2009), when given after reactivation (Soeter & Kindt, 2012b), and when testing occurred 28 days after propranolol+reactivation (Soeter & Kindt, 2012a), showing long-term disruption of the memory through reconsolidation blockade. Interestingly, the same effect has been shown for an instructed fear memory. In this study, participants were told that a certain CS would be followed by a shock (US). Participants learned the CS-US association, as expressed by a heightened startle response, even though an actual shock was never administered. At Day 2, the memory was reactivated and propranolol (or placebo) was administered. At Day 3, fear expression as well as subjective distress was diminished by propranolol+reactivation.

In all these studies (Kindt et al., 2009; Soeter & Kindt, 2010, 2012a), the explicit knowledge of the CS-US association was measured either online (i.e., during encoding) or retrospectively by asking participants if they were expecting the shock (US) in any given trial. While fear measured by the startle response and subjective distress was decreased by reactivation+propranolol, the explicit knowledge of the CS-US association
was not reduced in any of the experiments. The authors argued that US expectancy reflects the declarative (hippocampus-dependent) system of fear conditioning/extinction learning. Similarly, in other human fear conditioning studies, a double dissociation between hippocampus and amygdala has been suggested. Whereas amygdala lesion patients are able to verbally report the CS-US association without showing a conditioned response (LaBar, LeDoux, Spencer, & Phelps, 1995), hippocampal lesion patients are able to acquire conditioned fear but cannot recall the appropriate CS-US relationship (Bechara, Tranel, Damasio, Adolphs, & al, 1995). Reconsolidation may not affect all aspects of a memory in parallel; while the cognitive/declarative and emotional systems are highly related, they can also operate independently (LaBar & Cabeza, 2006; Phelps, 2004). Therefore, theoretically speaking, different memory systems could be differentially impacted by post-reactivation manipulations (i.e., during reconsolidation).

Sevenster, Beckers, and Kindt (2013) used these differential effects of propranolol on US expectancy vs. fear startle responses (i.e., reactivation+propranolol erases fear startle responses while having no effect on US expectancy) to investigate if prediction error (PE) played a role in this phenomenon. They changed (or not) the CS-US contingency at reactivation to induce prediction error (or not). Participants, on day 1, learned that a specific CS (e.g., spider picture, CS+) was paired with an electric shock while another CS (e.g., gun picture, CS-) was not. Online US expectancy ratings were taken during each trial. For the first group, participants received a shock with all the CS+ (i.e., 100% contingency). On day 2, the fear memory was reactivated with a CS+ only (no
shock) trial. This group was called the negative-PE group, as it entailed a decrease of reinforcement to reactivate the fear memory. The second group also received 100% contingency during CS-US training on day 1, but received a similar CS-US trial as reactivation on day 2. This group was called the no-PE group, as it entailed a reactivation consistent with training (i.e., reinforcement of the 100% contingency). The third group received a 33% contingency during CS-US training (i.e., in only 1 out 3 trials was CS+ paired with the US) and a CS-US trial as reactivation. This group was called the positive-PE group because it entailed an increase in learning compared to the 33% contingency received on day 1. All participants received propranolol immediately after their reactivation sessions on day 2 and extinction learning (CS only trials) on day 3 to probe CS-US contingency learning (by US expectancy ratings) and fear memory (by startle responses to CS+). The authors’ PE assumptions were confirmed by the different patterns of the US expectancy ratings in the three groups. The negative-PE group had decreased US expectancy ratings from the end of training (day 1) to the beginning of extinction (day 3), while the positive-PE group had increased ratings. Accordingly, the no-PE group showed no changes in ratings. That is, for the PE groups, but not for the no-PE group, changes in expectancy occurred from initial learning to reactivation. Reactivation+propranolol caused erasure of the fear memory (as showed by lack of differential response to CS+ compared to CS-) in both PE groups, but not in the no-PE group. These findings support the notion that it is not retrieval per se (as even in the no-PE group the memory was retrieved during reactivation) that triggers
labilization/reconsolidation processes, but the inconsistency (i.e., prediction error) at retrieval/reactivation.

1.1.2.2.1 Reconsolidation of human episodic memory

Two simultaneous reports (Forcato et al., 2007; Hupbach et al., 2007) were the first to show reconsolidation in human episodic memory, i.e., memory for autobiographical events, contexts, emotions, etc. (Tulving & Donaldson, 1972). Hupbach and colleagues used a 3-day experiment (Figure 1.1A), not unlike the fear conditioning studies discussed above, but with 48hr intervals between experimental days (instead of 24h in the fear conditioning studies). Participants learned a set of objects (to a certain criterion) that were presented one by one and then placed in a blue basket (day 1). On day 2 one group had the memory for the objects learned on day 1 reactivated by a simple reminder: the same experimenter in the same room asked the participants what they did with the blue basket; no explicit recall of items was permitted (i.e., participants were immediately stopped if they started to talk about the items they saw on day 1). This was followed by learning of a new set of objects, which were presented all at the same time (different encoding procedure than day 1 objects).
This group was compared to a group that learned the new set of objects without reactivation (in a different context: new room with a new experimenter) and to another group that completely omitted day 2. Another 48 hours later (experimental day 3), memory for the objects presented at day 1 was tested with a free recall test. Those in the reactivation/new learning group showed more intrusions of day 2 objects in their recall of day 1 objects, compared to the other two groups. This suggests that because
new learning occurred within the reconsolidation window, the original memory was updated. The important comparison here is that this ‘updating’ effect was dependent on the reminder of the day-1 task (Figure 1.1B). Actual memory for day 1 items (i.e., number of day 1 objects remembered) was not significantly different between the groups. Follow-up experiments investigated factors that directly affect this memory updating effect. The effect is not shown if memory is tested immediately after learning of set 2 objects (experiment 2), suggesting that reconsolidation is a time-dependent process (Nader, 2003), in this case showing that reconsolidation has to end before memory for set 2 can intrude memory for set 1. In a third experiment, they showed that if memory for day 2 items was tested in day 3, no differences in the mean percentage of day 2 items were found between the groups. The authors argue that these findings provide evidence against a source-monitoring account (Johnson, Hashtroudi, & Stephen, 1993). From a source monitoring perspective, the more features two events have in common, the more likely are errors to occur, and since for the reminder group, learning of set 1 and set 2 occurred in the same context, one could assume that the intrusions of set 2 on recall of set 1 (experiment 1) could be due to source monitoring. This is not an adequate explanation because one would expect a similar pattern for intrusions of set 1 on recall of set 2 (experiment 3), and that was not the case.

Furthermore, this research group has showed that context may be a primary factor necessary for reconsolidation to occur. That is, in an experiment where only one component of the original reminder was presented (i.e., experimenter, room, or
reminder question), only the room alone (the spatial context) resulted in the memory updating effect (Hupbach, Hardt, Gomez, & Nadel, 2008).

The studies performed by Hupbach and colleagues are argued to show a memory updating effect (including items from set 2 into set 1 recall), but not an actual enhancement or disruption of the original memory (i.e., memory for set 1 was not significantly different between the groups). Forcato and colleagues (2007) showed impairment of a declarative memory by disrupting reconsolidation with new learning of related material (after reactivation). They decided to use retrieval-induced forgetting (RIF; Anderson, Bjork, & Bjork, 1994; MacLeod & Macrae, 2001) of the second memory as the method of assessing impairment (or not) of the first memory. RIF refers to the fact that remembering (memory A) can temporarily block subsequent retrieval of another related memory (memory B). Their logic came from the notion that while retrieval deficits (of memory A) at testing could be due to errors with storage and/or simultaneous retrieval of related information (Mayes & Downes, 1997), RIF of memory B could only occur if memory A is intact. Accordingly, an absence of RIF (for memory B) is taken as memory impairment (for memory A) (Levy & Anderson, 2002). In this case, memory A was the first list of syllable pairs (L1; learned in day 1) and memory B was the second list of syllable pairs (L2; learned in another day). In their task, participants learned a list (L1) of five nonsense syllable pairs (e.g., ITE-OBN) on day 1 within a learning context that consisted of background music, a colored light, and a picture displayed on a screen. The learning context persisted throughout the presentation of the five syllable pairs. On day 2, the memory for L1 was reactivated (or not) with a ‘fake
retrieval test': the learning context was presented (colored light+music+picture), and then the first cue syllable (ITE) was presented for 2 secs before the trial was abruptly ended, prohibiting the participant from answering with the first target syllable (OBN). Five minutes after reactivation, participants learned (or not) a second list (L2) of syllable pairs (with a new learning context). On day 3, participants were tested for both lists. In this first experiment, they found that reactivation+L2 learning, but not reactivation only or L2 learning only, impaired L1 memory, as measured by an absence of RIF of L2. In two follow-up experiments, the interval between L1 learning and reactivation was manipulated. Reconsolidation disruption (i.e., no RIF) was observed if the interval was 6 hours but not if it was 10 hours, suggesting that reconsolidation processes in this task last somewhere between 6-10 hours (Schneider & Sherman, 1968; but see Duvarci & Nader, 2004; Schiller et al., 2010). The authors also showed that this L1 memory impairment lasted for at least 48 hours by testing the memory on day 4. Forcato and colleagues have also showed the importance of the structure of the reminder (Forcato, Argibay, Pedreira, & Maldonado, 2009). In this experiment, the reactivation (i.e., the reminder) was manipulated. They demonstrated that presenting the learning context only did not impair L1 memory (as measured by presence of L2 RIF). Additionally, allowing the participant to type in a target syllable in response to the cue syllables did not impair L1 memory. Only the learning context + presentation of cue syllable with no chance to answer resulted in L1 memory impairment. This is in agreement with the notion that for labilization-reconsolidation to occur, there needs to be a mismatch in the expectancy of the subject (i.e., a prediction error) at the moment of reactivation (Exton-
McGuinness et al., 2015; Morris et al., 2006; Pedreira et al., 2004; Sevenster et al., 2013), which in this case is the unavailability of responding with the target syllable (OBN) when a cue syllable (ITE) is presented.

Schwabe and Wolf (2009a) were able to disturb autobiographical memory by disrupting reconsolidation mechanisms. Participants reactivated memories of their life by recalling positive, negative, and neutral events that occurred within the last week (writing them down in detail for 4 minutes). Immediately after, they were instructed to memorize a fictional story (interference). Memory for the events recalled in day 1 was tested a week later by asking them to provide as many details as possible for each memory. The reactivation/interference group remembered significantly fewer details of the neutral events. The authors suggest that emotional memories are more resistant to change, via amygdala-dependent enhancement of consolidation processes. This effect was not found when they just reactivated the memory or read the fictional story, suggesting that this is dependent on reactivation/reconsolidation processes. In another experiment (Wichert, Wolf, & Schwabe, 2013b), participants learned a set of negative and neutral pictures, one week later reactivated the memory by thinking about the pictures, and either were presented a new set of pictures one time or three times. The purpose of this study was to assess whether strength of the new encoding (pictures at experimental day 2) affected the reconsolidation-based disruption of the original memory. Indeed, the authors found that strong (3x) encoding, but not weak (1x) encoding, lead to a decrease of the original memory for those who learned the new set of pictures after reactivating the original memory, indicating that strength of the
new/interfering memory may be a necessary component to disrupt reconsolidation. Interestingly, in this and similar studies by this group (Wichert, Wolf, & Schwabe, 2011, 2013a), no differences between neutral and emotional material was found (i.e., reconsolidation disruption affected memory regardless of valence). Similar effects of memory disruption by new learning after reactivation have also been found in post-event misinformation paradigms (Chan & LaPaglia, 2013; St Jacques, Olm, & Schacter, 2013; St Jacques & Schacter, 2013).

Memory strengthening through reconsolidation has also been shown in declarative memories more recently (e.g., Forcato, Fernandez, & Pedreira, 2013; Forcato, Rodríguez, & Pedreira, 2011; St Jacques et al., 2013; St Jacques & Schacter, 2013). Forcato and colleagues used a similar version of their task (discussed above) to show that multiple labilization-reconsolidation processes result in memory strengthening. On day 1, participants learned a list of nonsense syllable pairs. On day 2, participants received 1, 2, or 4 reactivations (using the specific reminder of their previous studies, i.e., learning context + one cue syllable presentation without the chance to answer) within 5 minutes of each other. Memory for the syllable pairs was tested on day 3. They showed (experiment 1) that 2 and 4 reminders resulted in equal memory enhancement (as showed by fewer errors), when compared to the single reminder condition. In a second experiment, they showed that 1 cue-only reminder (i.e., not allowing participants to answer) yielded similar performance as 2 retrieval reminders (i.e., allowing participants to answer). The authors argue that these findings support the idea that a retrieval reminder does not produce memory strengthening.
because it does not trigger labilization/reconsolidation of the memory. In a third experiment, they showed that if tested immediately after reactivations (i.e., short-term memory test) there was no memory strengthening by multiple reactivations (1 reminder vs. 2 reminders), further supporting the idea that this was a reconsolidation-dependent effect, since the process has to be culminated for effects to be observed. In their next experiments, they investigated the time-dependency of the multiple reactivations. If reactivation #2 occurred two hours after reactivation #1, then performance was comparable to when they were separated by 5 minutes (i.e., memory strengthening was observed), in agreement with the idea that reconsolidation is not completed in 2 hours. On the contrary, if reactivation #2 occurred 24 hours after reactivation #1, no beneficial effect of multiple reactivations was observed, suggesting that subsequent reactivations need to occur inside the reconsolidation window of the previous reactivations. In a later experiment, Forcato et al. (2013) found that multiple reactivations also made the original memory more resistant to interference. Using their original paradigm (L1 learning at day 1 and L2 learning at day 2), they found that if L1 memory was reactivated twice before learning L2 (at day 2), memory (for L1) on day 3 was better than if it was reactivated only once before learning L2. The authors also showed that memory enhancement by multiple reactivations seems to be impaired in older memories (7 days), probably in part due to natural forgetting of the original memory, even when the data is re-analyzed to only include good learners (>85% performance; Forcato, Fernandez, & Pedreira, 2014).
Wichert, Wolf, and Schwabe (2013a) were not able to find a memory enhancement (or protection) effect with multiple reactivations. Participants learned a set of negative and neutral pictures on day 1. One week later, memory was reactivated via a free recall test, which was followed by a new set of pictures. Memory for the first set of pictures was tested on day 8. In additional groups, participants received a single reactivation or 3 reactivations (of set 1), between day 1 and day 8 (only applying 1 reactivation per day). They found that, in all conditions, new learning after reactivation disrupted the original memory equally; that is, multiple reactivations on different days did not protect the original memory from reactivation+interference disruptions. This is consistent with Forcato et al. (2011), in which 2 reminders separated by 24 hours had no memory strengthening effect, further suggesting that repeated reactivations have to occur within the time window of the initial reactivation.

Other reports have also found memory strengthening, not by multiple reactivations, but by other manipulations during the reconsolidation window, such as benzodiazepine treatment (Rodríguez et al., 2013), stress (Bos, Schuijer, Lodestijn, Beckers, & Kindt, 2014; Coccoz, Maldonado, & Delorenzi, 2011; but see Dongaonkar, Hupbach, Gomez, & Nadel, 2013) and transcranial magnetic stimulation in the prefrontal cortex (Sandrini, Censor, Mishoe, & Cohen, 2013).

1.1.2.3 Functions and theories of memory reconsolidation

As suggested by the evidence discussed above, memory reconsolidation is hypothesized to serve two adaptive functions: memory updating and memory
stabilization. Memory updating refers to the function of incorporating new information into old memory traces, as observed in animal models (Morris et al., 2006; Rodriguez-Ortiz, Cruz, Gutiérrez, & Bermudez-Rattoni, 2005), and in human models like the Hupbach et al. studies discussed above (Hupbach et al., 2007; Hupbach, Gomez, & Nadel, 2011; Hupbach, Gomez, Bootzin, & Nadel, 2009; Hupbach, Gomez, & Nadel, 2009). Nonetheless, others have argued that the ‘updating effect’ of the Hupbach group can be explained outside of a reconsolidation view, with the temporal context model. This model states that items are encoded simultaneously with the context in which those items were encoded. According to the model, participants that receive the reminder at day 2 activate the context of day 1, so that items learned on day 2 are also associated with this context. This could explain the unidirectional effect of day 2 items intruding in recall of day 1 items (Sederberg, Gershman, Polyn, & Norman, 2011). In the Hupbach studies the ‘updating effect’ was not observed in a short-term memory test (i.e., before reconsolidation ended), but according to this model, this reflects the idea that consolidation (and not reconsolidation) of day 2 items must end before the intrusion effect can be observed.

Early evidence of reactivation-induced strengthening comes from animal models (Inda, Muravieva, & Alberini, 2011; Lee, 2008; Rohrbaugh & Riccio, 1970; Rossato, Bevilaqua, Medina, Izquierdo, & Cammarota, 2006). Exposure to the context (the CS in contextual fear conditioning) after consolidation has ended can result in the decrease of fear due to forgetting (Rohrbaugh & Riccio, 1970) or reverse the decrease of a spatial memory due to partial extinction (Rossato et al., 2006).
Research with human subjects has also shown memory strengthening via memory reactivation/reconsolidation by applying stress (e.g., Coccoz et al., 2011; but see Tollenaar, Elzinga, Spinhoven, & Everaerd, 2009 for conflicting results) or glucose during reconsolidation (Coccoz, Sandoval, Stehberg, & Delorenzi, 2013) and, as discussed above, by multiple subsequent reactivations (Forcato et al., 2013, 2011).

A question that is older than the term ‘reconsolidation’ itself is still very much applicable to the studies of reconsolidation processes: Do the memory impairments of reconsolidation disruption really reflect the loss of the memory (storage deficit) or, rather, do they reflect a failure to correctly retrieve the memory (retrieval deficit). Research that points out the molecular differences (and similarities) of memory consolidation and reconsolidation (e.g., Lee et al., 2004) seems to support the storage account. This account challenges the standard model of consolidation, such that the standard model proposes a single consolidation mechanism that occurs during the post-encoding interval (McClelland et al., 1995; Squire & Alvarez, 1995). The lingering consolidation hypothesis tries to propose a model to (start to) answer the storage vs. retrieval question (Dudai & Eisenberg, 2004). According to this model, reactivation of a memory destabilizes the long-term memory and a continuing phase of consolidation occurs. Reconsolidation is not a new process, but a continuation of consolidation; that is, this model assumes that consolidation is a much slower process than most studies adopt. What some assume is “cellular/synaptic consolidation” (arguably a minutes/hours-long process) does not reflect the completion of memory consolidation per se. Evidence showing long-term effects of reconsolidation manipulation (e.g., Debiec
et al., 2002; Soeter & Kindt, 2010; but see Wichert et al., 2011) would support a storage account for reconsolidation, but the lingering consolidation model would argue that these effects are still part of the long (original) consolidation process.

Differences in the reconsolidation effects on new and old memories (e.g., Forcato et al., 2013) can also be interpreted under the mechanisms of the standard model (McClelland et al., 1995; Squire & Alvarez, 1995) and the multiple trace theory of consolidation (Nadel & Moscovitch, 1997). Under the standard model of consolidation, if it is assumed that reconsolidation is dependent on the brain region of encoding (and initial storage), like the hippocampus, then labilization/reconsolidation through reactivation is only possible when the memory is still dependent on this region. Therefore, for an old memory, reactivation/reconsolidation may be harder (if possible) to be triggered because the memory is arguably hippocampus-independent. Under the multiple trace theory, for recent memories, reconsolidation is conceivable because there are a limited number of traces that represent a recent memory in the hippocampal-neocortical system. For older memories, there are more and distributed memory traces, which might result in a higher difficulty to initiate memory reconsolidation because this process would have to be triggered in all traces.

In the current study, the notion that reconsolidation is a distinct process from consolidation, and that is in itself a storage mechanism, is adopted, because much of the research supports this view; that is, multiple labilization/reconsolidation cycles strengthen the storage, not the retrieval success, of the reactivated memory.
Nonetheless, it is of importance to note that, with behavior only, it is virtually impossible to separate storage deficits from retrieval deficits.

1.1.2.4 Boundary conditions for memory reconsolidation

As can be inferred from the literature discussed above, memory reconsolidation does not occur in an absolute manner every time a memory is retrieved or reactivated. Boundary conditions for labilization/reconsolidation have been presented and are summarized here (see Besnard et al., 2012; Dudai, 2012):

1) *Trace competition*: This boundary condition is largely based on fear conditioning studies, in which the reactivation is normally a presentation of the unpaired CS or unpaired US (i.e., tone only or shock only presentation). When this occurs, two mutually exclusive processes can be triggered (Merlo, Milton, Goozée, Theobald, & Everitt, 2014): reconsolidation or extinction. If the reminder (e.g., single CS) presentation is short, reconsolidation is triggered; if the presentation is long, or if there are several presentations, then extinction learning occurs (e.g., Pedreira et al., 2004). Similarly, if training is robust and extinction-resistant, then *reconsolidation* is affected by amnesic agents, but if training is weak, then *extinction learning* is affected by amnesic agents (Eisenberg, Kobilo, Berman, & Dudai, 2003).

2) *Expectancy/prediction error*: This boundary condition refers to the predictability of the reactivation stimulus and the availability of new information during reactivation (Pedreira et al., 2004). This can be observed in several animal studies (e.g., Morris et al., 2006; Winters, Tucci, & DaCosta-Furtado, 2009) and in human studies (e.g.,
Sevenster et al., 2013). It is important to note, that no distinction has been made in the research to date regarding whether procedural and content prediction error are functionally different in their role to trigger reconsolidation. In the literature discussed above, the reminder structure of the Forcato et al. task does not trigger reconsolidation if the participant is able to respond; that is, if there is no expectancy/prediction error between what the participant expects and what actually occurs. It is important to note that there have been studies that show (declarative) memory reconsolidation even when no prediction error is assumed (e.g., Strange, Kroes, Fan, & Dolan, 2010).

There is some evidence that suggests a ‘switch’ mechanism, regarding the role of reconsolidation, for this mismatch component: if there is a mismatch, updating occurs; if there is no mismatch, strengthening (or precision-maintenance) occurs (De Oliveira Alvares et al., 2013). In a series of elegant experiments, De Oliveira Alvares and colleagues (2013) used a contextual fear conditioning paradigm in rodents to show how different reactivation conditions result in different roles of reconsolidation. In their first set of experiments, rats were placed in context A (CS) for 4 minutes, in which they received 4 footshocks (US). Two days later, as a reactivation, they were exposed (or not) to a hybrid AB context for 90 seconds. Two days after, memory was tested in context B (a context that had some of the details of context AB). Rats that received the reactivation session, compared to those who did not, showed higher freezing to the new context, arguably showing a memory updating effect (i.e., reconsolidation at day 2 modified the memory for context A to integrate context AB). This effect was not shown
if rats were tested on a completely novel context, showing that this was not a simple effect of accelerated fear generalization (i.e., all contexts leading to a fear response).

To show that this memory updating effect was due to reconsolidation, De Oliveira Alvares and colleagues used the L-type voltage-gated calcium channel antagonist nimodipine to block memory destabilization. Nimodipine has been shown to prevent the reconsolidation impairment caused by protein synthesis inhibitors, arguably by blocking memory destabilization (Suzuki, Mukawa, Tsukagoshi, Frankland, & Kida, 2008). De Oliveira Alvares and colleagues injected the rats with nimodipine before the reactivation session, which resulted in the reduction of freezing to context B (at test) in the nimodipine+reactivation group. The authors suggest that this memory updating by reconsolidation is governed by the mismatch between the training (in context A) and the reactivation session (in context AB).

In another set of experiments, De Oliveira Alvares and colleagues (2013) showed that memory reactivation could also lead to maintenance of memory precision. Rats were trained in context A. Two weeks later, they received a reactivation session (or not) in context A (note the lack of mismatch). Twenty-eight days later, they were tested in either context A or a novel context. Under normal conditions, rats are able to discriminate between a conditioning context and a novel conditioning shortly after training; that is, rats only show a fear response (i.e., freezing) in the learning context. In contrast, four weeks after training, rats will show a generalized fear response; that is, they will freeze in any (fairly similar) context. Non-reactivated rats had similar freezing levels in both the learning (A) and the novel context, showing fear generalization. Rats
that received the reactivation session at day 14 showed significantly lower freezing in the novel context, showing that reconsolidation maintained the contextual details (i.e., memory precision) of the learning context. Nimodipine before reactivation resulted in similar freezing levels in both contexts, i.e., generalization of fear still occurred, arguably because the drug blocked memory destabilization/reconsolidation.

De Oliveira Alvares and colleagues also showed that reconsolidation could lead to memory strengthening in this contextual fear paradigm. Rats were trained, reactivated (or not), and tested for 4 minutes in context A (note the lack of mismatch). These 3 sessions were separated by 2 weeks each. The first test (at day 28) showed no differential effect on performance (freezing) between the groups. Because freezing levels were fairly high in this first test, the researchers gave the rats another (longer) test the following day (day 29). This longer test (12 min) was done in order to detect possible late signs of extinction. Interestingly, no differences between the groups were shown in the first 6 minutes of the test. In contrast, in the last 6 minutes, the non-reactivated group and the nimodipine+reactivated group showed signs of extinction (i.e., reduced freezing), whereas the reactivated group showed higher levels of freezing, comparable to those during the first test (at day 28). In other words, reactivation/reconsolidation resulted in a stronger memory, as depicted by resistance to extinction.

These series of studies hint at the idea that the mismatch/prediction error component (or lack of), at least in contextual fear conditioning paradigms, could be considered a ‘switch’ for the biological function of memory reconsolidation. Under
consistent (no mismatch) conditions, in which no new information is present (like training and reactivation in the same context A), memory strengthening (or at least precision maintenance) is triggered. Under inconsistent (mismatch) conditions, like training in context A and reactivation in context AB, memory updating is then triggered.

3) Age and strength of the memory: Under this condition it is understood that newer memories can be labilized/reconsolidated more easily (Suzuki et al., 2004). Forcato et al. (2013) found no memory strengthening via multiple reactivations when the reactivations came 7 days after initial encoding. Only in participants that were good learners (>85%) did they observe a reactivation effect at day 8, and even then, there was no memory strengthening by additional reactivations (i.e., no differences between 1, 2, or 4 reactivations).

This boundary condition has been suggested to be related to the expectancy/prediction error condition (e.g., Díaz-Mataix, Ruiz Martinez, Schafe, LeDoux, & Doyère, 2013; see Exton-McGuinness et al., 2015; Sevenster et al., 2013). When fear conditioning training is asymptotic (i.e., freezing is at its maximum by the end of training), reactivating the memory with an additional CS-US trial does not elicit labilization/reconsolidation; that is, there is zero prediction error at reactivation. In contrast, if the reactivation session includes new learning (i.e., a prediction error), like the change in the timing between the CS and US, then reconsolidation (and memory updating) is triggered (Díaz-Mataix et al., 2013).
1.1.3 The current study

The specific aim of the current research is to test, in an ecologically valid way, the efficacy of the memory strengthening process caused by multiple subsequent reactivations (following the design of Forcato et al., 2011). To date, the only test of this memory strengthening had been against new (interference) learning (Forcato, Fernandez, and Pedreira, 2013) and performed by the same research group. Therefore, replication and extension of this research in another laboratory was necessary, and additional research was required to assess the new strength of a (multiple) reactivated memory. The first experiment attempted to replicate the original findings of experiment 1 and 2 in Forcato and colleagues (2011); that is, the general finding that multiple reactivations (2 or 4 reactivations within 5 minutes of each other) result in a stronger memory than a single or no reactivation.

The second experiment aimed to test the effectiveness of memory strengthening via multiple reactivations using a stressor before retrieval. Stress, and the stress hormone cortisol, has been robustly shown to be detrimental to memory retrieval (see Section 1.1.3.1 below). High levels of stress and glucocorticoids during retrieval have been shown to decrease memory performance in rodents (e.g., de Quervain, Roozendaal, & McGaugh, 1998) and humans (e.g., de Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000; Tollenaar, Elzinga, Spinhoven, & Everaerd, 2008; see Wolf, 2009). Therefore, in the second experiment we used an ecologically-pertinent, psychosocial stressor (the Trier Social Stress Test, TSST), which has been shown to robustly increase stress-induced physiological responses (Kirschbaum, Pirke, &
Hellhammer, 1993), immediately before retrieval (at day 3) to test its ability to impair a memory that has been strengthened by multiple reactivations/reconsolidation cycles.

1.1.3.1 Stress and memory retrieval

A stressor triggers the acute release of epinephrine and norepinephrine from the sympathetic nervous system, but more importantly the activation of the hypothalamic-pituitary-adrenal (HPA) axis. The HPA axis activation engages secretion of corticotropin-releasing hormone (CRH) from the paraventricular nucleus of the hypothalamus, then adrenocorticotropin hormone (ACTH) from the anterior lobe of the pituitary gland, and finally glucocorticoids from the adrenal cortex into the blood stream. The release of glucocorticoids (cortisol in humans) activates mineralocorticoid (MRs) and glucocorticoid receptors (GRs) which are expressed throughout the brain, but mostly in the amygdala, hippocampus, and prefrontal cortex, which are key regions for memory formation (Mcewen, Weiss, & Schwartz, 1968; see Finsterwald & Alberini, 2014).

The effects of stress and glucocorticoids on memory have been well researched in psychology (see Calabrese, 2008). In general, stress (or cortisol), at medium levels during or shortly after learning results in memory enhancement (e.g., Abercrombie, Kalin, Thurow, Rosenkranz, & Davidson, 2003; Buchanan & Lovallo, 2001), and increases in cortisol during learning are positively correlated with memory strength (e.g., Abercrombie et al., 2003; Smeets et al., 2008). The beneficial effects of cortisol during consolidation are believed to follow a dose-response pattern in the shape of an inverted-U; that is, too low or too high cortisol levels do not produce a memory
enhancement (see Diamond, Campbell, Park, Halonen, & Zoladz, 2007). In contrast, stress (or cortisol) shortly before retrieval is generally known to cause memory impairment (e.g., de Quervain et al., 2000; Kuhlmann, Piel, & Wolf, 2005; see Roozendaal, 2002; Tollenaar et al., 2009). Furthermore, there are few studies that have administered stress during reconsolidation (i.e., after memory reactivation) with mixed results. One study has shown that psychosocial stress after reactivation decreased reconsolidation-mediated memory updating, i.e., fewer list 2 intrusions in list 1 recall (Dongaonkar et al., 2013). Other studies have shown that psychosocial or physical stress after memory reactivation can increase memory for a slideshow story (Marin, Pilgrim, & Lupien, 2010), nonsense syllable pairs (Coccoz et al., 2011, 2013) and emotionally neutral, positive, and negative words (Bos et al., 2014). Because the current study was not interested in affecting memory consolidation or reconsolidation via stress, but rather using stress as a tool to impair memory, we will discuss further only the research of stress at the moment of memory retrieval.

De Quervain and colleagues (2000) were the first to assess the effects of cortisol specifically during retrieval. In a within-subject design, participants were given a 25mg pill of cortisone (which resulted in stressor-like levels of cortisol in saliva) or placebo in one of three groups: pill administration before encoding, after encoding (targeting consolidation), or before delayed retrieval (24 hr. after encoding). Cortisone before retrieval resulted in decreased delayed recall when compared to the placebo condition. No effects were found in the other two conditions. This finding supported similar results in rats, in which stress before retrieval in a water maze resulted in impaired memory (de
Quervain et al., 1998). In a similar experiment, psychosocial stress (instead of a cortisone pill) before delayed retrieval was found to impair memory for a word list, particularly of emotionally positive and negative words (Kuhlmann, Piel, et al., 2005).

Recent evidence has pointed to opposite roles of the MRs and GRs in memory retrieval in humans (Rimmele, Besedovsky, Lange, & Born, 2013). Blocking MRs using spironolactone (Aldactone) impaired memory retrieval, while blocking GRs using mifepristone (Mifeprex) enhanced memory retrieval. Participants learned emotionally negative and neutral texts and pictures followed by an immediate free recall test for both texts and pictures (to acquire baseline memory performance). In the first experiment, participants received the MR antagonist (or placebo; within subjects, double-blind) two days after encoding, during the night (11PM) and early morning (4AM) before retrieving the texts and pictures in the morning (~8PM). A similar procedure was done in the second experiment, in which a GR antagonist or placebo (mifepristone) were administered. Plasma cortisol was increased during the retrieval period in both experiments, confirming that MRs and GRs were successfully blocked during the retrieval period. MR blocking reduced recall (increased forgetting) of the negative, but not neutral, texts, pictures, and picture details. Interestingly, GR blocking had no effect on any type of text, but increased recall (reduced forgetting) of the neutral pictures only. Nonetheless, there was also an increase in recall for both negative and neutral picture details. The authors argue that MR activation is crucial for successful memory retrieval but do not rule out the possibility that the impaired retrieval (caused by MR blocking) could have also occurred as a result of predominant GR activation. It
has been showed that increased GR activation leads to retrieval impairment (e.g., de Quervain et al., 1998; Kuhlmann, Piel, et al., 2005). Accordingly, the authors argued that high GR activation is detrimental for memory retrieval, as blocking GRs resulted in increased recall. This increased recall can also be an effect of highly activated MRs. In short, MR activation is beneficial, while GR activation is detrimental for memory retrieval. This is consistent with the idea that cortisol’s effects on memory follow an inverted-U pattern; whereas medium levels of cortisol (in which MRs are virtually saturated) are good for memory, low levels (in which MRs are not activated) and high levels (in which GRs are activated) of cortisol impair memory.

1.1.3.2 Hypotheses of the current study

1.1.3.2.1 Experiment 1: Replication of Forcato et al., 2011

Experiment 1 aimed to replicate the main finding of Forcato and colleagues (2011), that is, that 2 or 4 reactivations lead to memory strengthening, by using a similar version of their task. This replication experiment included four groups:

R0: A group that received no reactivation on day 2.
R1: A group that received a single reactivation on day 2.
R2: A group that received 2 reactivations (separated by 5 minutes) on day 2.
R4: A group that received 4 reactivations (separated by 5 minutes) on day 2.

Figure 1.2A shows the graphical representation of the hypotheses. Note that the graphs depict ‘mean number of errors’, i.e., lower bars = better memory. We expected that the
Figure 1.2 Graphical depictions of the hypotheses for A) experiment 1 and B,C) experiment 2. Note that the y-axis depicts mean number of errors, i.e., lower bars = better memory.
R2 and R4 group would have better performance (fewer errors) than the R0 and R1 groups. No difference in performance should be observed between the R2 and R4 groups (experiment 1 of Forcato et al., 2011). No difference in performance was expected between the R0 and R1 groups (experiment 2 of Forcato et al., 2011).

1.1.3.2.2 Experiment 2: Testing the ‘memory strengthening by multiple reactivations’ effect against stress at retrieval

Experiment 2 used a similar behavioral paradigm as Forcato and colleagues (2011; 2013) and included four groups (Figure 1.2B,C):

R0-Control (R0C): A group that received no reactivations on day 2, and that received the control version of the TSST on day 3 before testing.

R0-Stress (R0S): A group that received no reactivations on day 2, and that received the regular version of the TSST on day 3 before testing.

R2-Control (R2C): A group that received 2 reactivations on day 2 (shown to be enough for memory strengthening (Forcato et al., 2011), and that received the control version of the TSST on day 3 before testing.

R2-Stress (R2S): A group that received 2 reactivations on day 2, and that received the regular version of the TSST on day 3 before testing.

Figure 1.2B shows the graphical representation of the hypotheses. In general, we expected an overall detrimental effect of stress on memory retrieval, measured by a higher number of errors committed at testing (as in Forcato et al., 2011, 2013); that is, a main effect of stress.
More importantly, multiple reactivations could strengthen a memory in such a way that performance after retrieval under stress (R2-STRESS) could be better even than performance with no stress, but with no reactivation (R0-CONTROL; Fig. 1.2B); that is, memory strengthening would be so effective that it not only shields the memory from the impairing effects of stress, but is also enhanced beyond the unmodified original memory. Alternatively, if memory strengthening is not as effective (or the TSST is too damaging), then stress at retrieval could impair the memory to levels similar to the original memory (solid line in Fig. 1.2C), but memory strengthening would still occur, as shown by better performance (fewer errors) than the group that received no reactivations but stress at retrieval (dashed line in Fig. 1.2C). The former hypothesis (Fig. 1.2B) is consistent with Forcato and colleagues (2013), who showed that for the L1 memory, two reactivations before interference learning (R2 + L2) yielded performance comparable to a single reactivation with no interference learning (R1), suggesting that the R2+L2 group could have performed better than a no-reactivation (R0) group.
CHAPTER 2:  
METHOD  

2.1 Participants  

All subjects were college students recruited from the University of Notre Dame and were compensated with either course credit or cash payment. Age range was 18-27 years (M=19.47, SD=1.59). After providing consent for participation, subjects were screened for self-reported sleep, stress, and mental health disorders, irregular sleep habits, and medication/drug use. Participants were asked to refrain from alcohol and caffeine at least 2 hours before the experimental session time on each of the three days of the experiments.  

According to power analyses with G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) a sample size of N=24 per experiment (grey line in Figure 2.1) should have been appropriate to detect a medium-size effect (d=0.6), at least between the REACTIVATION conditions, with high power (~.85). That translates to n=6 per group, which is markedly low compared to other cognitive psychology experiments. Although this may have been possible, since Forcato and colleagues (2011,2013) had 5-8 participants per group in their studies, we believed that such low sample sizes were not rigorous enough and would not be appropriate to reach robust conclusions. We aimed for larger samples (at
least N=60 per experiment) to be collected. The reason for this change was two-fold: 1) Forcato and colleagues’ studies need to be replicated with larger samples and 2) providing higher confidence about the results of the current experiments. Experiment 1 had a final sample size of N=76, and Experiment 2, due to slow recruitment and experimental time constraints, had a final sample size of N=30. See sections 3.1.1 and 3.2.1 below for descriptive measures about each sample.

2.2 Materials

2.2.1 Demographic Information Questionnaire

This was a 24-item questionnaire asking for general information such as sex, age, drug and medication use, as well as occurrences of sleep or mental disorders, etc.
2.2.2 Three-day Sleep Log

This measure was a log of the dates, approximate bedtime, wake time, time to fall asleep, naps, and alcohol and caffeine consumption over the past three days. This sleep-pattern information is important because sleep and stress have an intricate relationship (Everly Jr & Lating, 2013); for example, sleep deprivation has been shown to increase cortisol levels both at baseline and in response to a stressor (Minkel et al., 2014). This questionnaire was used to assess the exclusion criteria mentioned above (section 2.1).

2.2.3 Positive and Negative Affect Schedule (PANAS)

The PANAS is a 20-item questionnaire asking participants to rate on a scale of one (not at all) to five (extremely) the extent to which they currently feel each of 20 positive or negative emotions (e.g., interested, defeated). This scale yields an overall score for positive affect (PA) and negative affect (NA) (Watson, Clark, & Tellegen, 1988).

2.2.4 Self-Assessment Manikin (SAM)

The Self-Assessment Manikin (SAM) is a speech-free method to capture the dimensions of valence and arousal affective reactions. It consists of two rows of pictograms, each of which has one dimension on a nine-step scale through stylized figures (M. M. Bradley & Lang, 1994).

2.2.5 State-Trait Anxiety Inventory (STAI)

The STAI is a 20-item questionnaire measuring the experience of trait and state anxiety. Participants are asked to rate the extent to which they feel a series of
statements (e.g., I feel pleasant), describe how they generally feel (trait) on a scale of one (not at all) to four (very much so). The state version asks how they feel at the moment (Spielberger, 2010). The PANAS, SAM, and STAI-state were used to assess the effects of the stressor (or control task) on mood and emotion.

2.2.6 Trier Social Stress Test

Stress was induced using the Trier Social Stress Test (TSST) (Kirschbaum et al., 1993), a procedure that reliably elicits moderate psychosocial stress in laboratory settings. The task consisted of an anticipation period of written speech preparation (10 min) and a delivery period that required participants to deliver a free speech from memory (5 min) and perform mental arithmetic (5 min) in front of two "judges," who were experimental confederates. Participants’ notes made during the preparation phase were abruptly taken from them immediately before the delivery period and they were asked to give the speech from memory. Furthermore, participants were told that they were going to be recorded for vocal and video analysis with the use of a microphone and a video camera that is connected to a 19-inch TV screen. To make the task additionally stressful, participants were able to watch themselves deliver the tasks in this TV monitor. The arithmetic task consisted of participants subtracting 13 from 1022 serially and without stopping for 5 minutes; if mistaken, they had to start again from 1022. If the subtraction seemed to be too easy, a harder subtraction was given.

In contrast, participants in the control group performed the 10-min delivery period sitting down in an empty room with no microphone, no video camera, and no
judges, but still delivering the speech quietly with their initial notes in hand. They were given the option of not delivering their speech out loud. Participants were told of their group assignment before the “speech” preparation, to help avoid anticipatory stress.

2.2.7 Experimental Room

The experiments were conducted in a small dark room (6’ x 6’) at the eMotion and eCognition Laboratory. The room had a personal computer, headphones, and colored lights reflecting to the back wall. The computer monitor was two feet away from the participant and a keyboard, illuminated by a small lamp, was provided. One foot behind the computer monitor there was a light box with nine colored light bulbs (3 red, 3 green, and 3 blue). See Figure 2.2 for overview of the room.
2.3 Procedure

2.3.1 General procedure of experiments 1 and 2

These experiments consisted of a three-day protocol: Day 1 was encoding, Day 2 was reactivation/reconsolidation, and Day 3 was testing (Figure 2.3 depicts experiment 2). Before participants started with the encoding phase on Day 1, they completed the demographic information questionnaire. Before the testing phase on Day 3, participants completed the 3-day sleep log, PANAS, STAI and the cortisol background questionnaire.
The task involved two types of trials: real trials and fake trials. Real trials consisted of a context presentation period (*context period*) followed by a syllable-pair presentation period (*syllable period*). Fake trials consisted of only a context period. A context period consisted of 6 seconds of the parallel presentation of three stimuli, depicted in Figure 2.4:

1. a colored light was projected into the back wall,
2. the presentation of an image (e.g., cascade #1),
3. the presentation of music (e.g., tango #1) through headphones.
A syllable period occurred immediately after the context period if the context was the one associated with the syllable period. A list of five nonsense-syllable pairs was presented visually during the syllable period. Every trial was preceded by a context period of five seconds, in which participants had to answer yes/no to whether the context presented in the context period is the one associated with the syllable-pair learning. This task was created to be a precise replication of the task created by Forcato and colleagues (2011, 2013). Each day the sessions occurred between 11:30AM-6PM for experiment 1 and between 12:30PM-5:15PM for experiment 2. The reason for the shorter time range of experiment 2 was to minimize cortisol extremes regarding the circadian peak and trough of cortisol.

During a real trial, the context was presented immediately before the syllable period, which started with the presentation of the cue syllable on the left side of the computer monitor and a white blank rectangular box on the right side (Figure 2.5). Participants had 5 seconds to use the blank box to type the response syllable, except in the first real trial, in which the response syllables were automatically filled out and

Figure 2.4 Graphical depiction of the context period
participants were instructed to just watch (i.e., not answer). After the 5 seconds, three situations were possible:

1. *If no syllable is typed*, the correct response syllable was presented in red font for 4 seconds,

2. *If an incorrect syllable is typed*, the incorrect response syllable was replaced by the correct response syllable in red font for 4 seconds, or

3. *If a correct syllable is typed*, this response syllable stayed on screen in black front for 4 additional seconds.
Immediately after this, the syllable pair disappeared and the next cue-syllable was presented a space below the disappeared syllable pair. The process was repeated until the whole list (5 syllable pairs in total) was presented. Every time a participant did not answer, or answered incorrectly, an error was computed. A real trial lasted 51 seconds: 5 seconds of context period and 45 seconds of syllable period.

2.3.2 Encoding (Day 1)

Each trial consisted of a combination of a red or blue light, a picture of a certain cascade, and a particular tango song. Only one combination of these stimuli, the specific context, was followed by the presentation of the syllable pairs. The real trial consisted of (Figure 2.6):

1. the red-colored light, projected in the back wall,
2. the picture of cascade #1 in the computer monitor,
3. the tango #1 music,

Encoding included 10 real trials pseudorandomly mixed with 22 fake trials, separated by 4-sec intervals. In the first real trial, the full list was presented (i.e., both cue and response syllables were presented), and in the following 9 trials, participants had the opportunity to type in the response syllables. Participants who did not reach a 60% (12 out of 20) correct response criterion for the last four real trials were excluded for analyses.
Participants arrived at the laboratory on Day 1 and completed the demographic information questionnaire. Before the encoding phase started, four “demo” trials were presented in which the instructions were thoroughly explained to the participants and questions, if any, were answered. These “demo” trials were similar in structure but had different contexts and syllable pairs. Day 1 lasted ~20 minutes in experiment #1 and ~40 minutes in experiment #2 because of saliva sample 1.

Because participants may have reached ceiling levels of learning on experiment 1 with 10 real trials, the encoding was halved for experiment 2 (5 real trials, 16 total trials).

Figure 2.6 Graphical depiction of the timing of real and fake trials
2.3.3 Reactivations/ No reactivations (Day 2)

Each reactivation trial was composed of the specific context that was used at list encoding in Day 1, followed by the presentation of a single cue syllable and the blank response box. After two seconds, participants were interrupted with a fake error message that stated: TRIAL INTERRUPTION: THE SESSION HAD TO BE SUSPENDED. OPEN THE DOOR AND THE NEW PROCEDURE WILL BE EXPLAINED (Figure 2.7). For experiment 1, a replication of Forcato and colleagues (2011)’s original experiment, there were four groups:

- R0: A group that received no reactivation on day 2 (skipped day 2).
- R1: A group that received a single reactivation on day 2.
- R2: A group that received 2 reactivations (separated by 4 minutes) on day 2.
- R4: A group that received 4 reactivations (separated by 4 minutes) on day 2.

Day 2 lasted about 10 to 30 minutes, depending on group assignment. For experiment 2, half of the participants were presented with two reactivation trials, separated by four minutes, whereas the other half skipped Day 2. By skipping Day 2, we warranted that participants had no incidental reactivation of the memory for the syllable-pairs. Day 2 lasted ~15 minutes. Saliva sample 2 was taken at the start of Day 2 and those that skipped Day 2 did not complete saliva sample 2.
2.3.4 Testing (Day 3)

The testing phase consisted of four real trials intermixed with 12 fake trials (context only); 16 trials in total. Participants had to type in the correct response syllable since the first trial, unlike the encoding phase in which the first trial was automatically completed.

Participants arrived at the lab and completed all questionnaires except the demographic information questionnaire (filled out in Day 1). In experiment 2, they had to complete saliva sample 3 during these questionnaires. After the participants were finished, they were immediately asked to fill out the SAM and were then randomly assigned to the stress (TSST) or control condition (in Haggar Hall 223-D). Participants filled out another SAM and completed saliva sample 4 immediately after the stress or control manipulation.

Figure 2.7 Graphical depiction of the reactivation trial presented on Day 2
For participants in experiment 1, the testing phase immediately followed the filling out of the questionnaires. In experiment 2, after completion of saliva 4 (post-manipulation), the experimenter and the participant walked back to Flanner Hall B09-A for the testing phase. Lastly, participants were debriefed and thanked for their participation. Day 3 lasted around 30 minutes for experiment 1 and around 1.5 hours for experiment 2, due to the TSST and saliva samples.

2.3.5 Memory measures and statistical analyses

Main results were reported and analyzed as mean number of errors for the testing phase (averaged between all 4 real trials). To perform a deeper (and exploratory) analysis of performance, we broke down the errors into three types (as in Forcato et al., 2011; 2013):

1. void-type errors: when nothing was typed in;
2. intralist-type errors: when the response-syllable was incorrect, but was included in the list;
3. confusion-type errors: when the response-syllable was not included in the list.

Using mean number of total errors, a 4 (Group: # of reminders) x 10 (Real Trials 1-10) mixed-measures ANOVA, with Real Trials as a within-subject factor, was performed to analyze the learning curves of each group during the encoding phase for experiment 1. A similar 2 (Stress manipulation: stress or control) x 2 (Reactivations) x 5 (Real Trials: 1-5) mixed-measures ANOVA, with Real Trials as a within-subject factor, was performed to analyze the learning curves of each group during the encoding phase for experiment 2. The Greenhouse-Geisser correction of degrees of freedom was used.
when the sphericity assumption was violated, as measured by the Mauchly’s Test of Sphericity. As another way of assessing learning (Forcato et al, 2011; 2013), a One-Way ANOVA (exp. 1) and a Two-Way ANOVA (exp. 2) were performed with the average of the mean number of total errors of the last four encoding (real) trials for experiment 1, and the last two real trials for experiment 2 (as encoding was halved for experiment 2 from 10 real trials to 5 real trials).

For experiment 1, a One-Way ANOVA was used to analyze overall error performance, followed by planned comparisons (LSD) between all four groups. For experiment 2, a Two-Way ANOVA (Stress/control and Reactivations) was used to analyze overall error performance, followed by planned comparisons (LSD) between all four groups (i.e., comparisons between the Stress manipulation groups and between the Reactivations groups; R0-control vs. R0-stress, R2-control vs. R2-stress, R0-control vs. R2-control, and R0-stress vs. R2-stress). As exploratory analyses, separate ANOVAs were performed for the three types of errors. The Bonferroni correction for multiple comparisons was used in any post-hoc (unplanned) comparison, and its use is specifically stated in the text. Effect sizes (partial eta squared and Cohen’s d) were presented for significant results.

To analyze the effectiveness of our stressor, we used the Self-Assessment Manikin (SAM)’s ratings of arousal and valence, as well as cortisol analysis derived from saliva samples, from before and after the manipulation. A 2 (Stress) x 2 (Reactivations) x 2 (Time: pre vs. post), mixed ANOVA, with Time as a within-subject factor was conducted for each rating (i.e., one for arousal and one for valence). Upon arrival in Day
1, participants were given ~15 min to acclimate to the laboratory setting prior to collection of the first saliva sample. Participants used the passive drool method using a straw to expectorate (i.e. no gum, cotton, or other saliva flow stimulants were used) and were instructed to fill the test tube to the 5mL line. They were allowed to drink sips of water throughout the experiment session, but only immediately after a sample. A total of 5 samples were collected: (Day 1) a sample collected ~15 min after participant arrival on Day 1, (Day 2) a sample collected ~5 min after participant arrival on Day 2 (for those who returned in Day 2), (Day 3-PRE) immediately before completion of the stress/control task, (Day 3-POST) immediately after completion of the stress/control task, and (DAY 3-END) after the final memory test. After each data collection session, vials were capped and frozen until later processing, which was done in-house using standard cortisol assay kits (MP Biomedical, Santa Ana, CA) and began with three freeze-thaw cycles and centrifugation. Cortisol levels were determined by solid-phase I radioimmunoassay using the protocol described by Wirth & Schultheiss (2006), using water-diluted standards/calibrators ranging from 0.5 to 50 ng/ml. In total, 1 assay was necessary to complete all samples. External controls were included and tested within expected ranges. A 2 (Stress) x 2 (Reactivations) x 3 (Day 3 Time points) mixed ANOVA, with Time points as a within-subject factor, was conducted to analyze how cortisol changed throughout experiment 2.
CHAPTER 3:

RESULTS

3.1 Experiment 1: Replication study

Experiment 1 aimed to replicate the general finding that multiple labilization/reconsolidations result in a strengthened memory the following day. For this purpose, we compared four groups: group R0 received no reactivations (skipped day 2), group R1 received 1 reactivation during day 2, group R2 received 2 reactivations during day 2, and group R4 received 4 reactivations during day 2. According to the early studies from Forcato and colleagues (2010), groups R2 and R4 should have better memory (as measured by less errors) than groups R0 and R1 at test, during day 3.

3.1.1 Descriptive and control variables

Eighty-four (84) participants were recruited for experiment 1, but 8 participants were excluded from analyses due to bad (<70% in last four trials) syllable-pair learning (5 participants), bad context learning (<50%, 1 participant) and missing Day 3 (2 participants), for a final sample size of N=76. The age range was 18-22 with M=19.41, SD=1.62, 60 females. Table 3.1 presents demographic information for each group.
### TABLE 3.1

**DEMOGRAPHIC INFORMATION, BY GROUP**

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Males</th>
<th>Females</th>
<th>Age M</th>
<th>Age SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>R0</td>
<td>24</td>
<td>9</td>
<td>15</td>
<td>19.29</td>
<td>1.04</td>
</tr>
<tr>
<td>R1</td>
<td>17</td>
<td>2</td>
<td>15</td>
<td>19.53</td>
<td>1.12</td>
</tr>
<tr>
<td>R2</td>
<td>22</td>
<td>6</td>
<td>16</td>
<td>18.82</td>
<td>0.96</td>
</tr>
<tr>
<td>R4</td>
<td>13</td>
<td>6</td>
<td>7</td>
<td>19.54</td>
<td>1.20</td>
</tr>
</tbody>
</table>

The STAI-trait, a measure of trait anxiety, was taken at the beginning of Day 1.

No differences in trait anxiety between the groups were found, F(3,72)=0.86, p=0.47.

The PANAS, a measure of state positive and negative affect, as well as the STAI-state, were taken at the beginning of Day 1 and at the beginning of Day 3. No differences were found in positive (PA) or negative affect (NA) on either day (F(3,72)=1.51, p=0.22, for PA on Day 1; (F(3,72)=0.39, p=0.76, for NA on Day 1; (F(3,72)=0.86, p=0.47, for PA on Day 3; (F(3,72)=1.07, p=0.38, for NA on Day 3). No differences were found for the STAI-state on Day 3, F(3,72)=1.74, p=0.17, but a significant difference was found for the STAI-state on Day 1, F(3,72)=3.29, p=0.025, partial $\eta^2=0.12$. Bonferroni-corrected post hoc comparisons show that group R0 (M=44, SD=4.94) was significantly higher (p=0.023) than group R1 (M=39.29, SD=5.89), but no other differences were found between the other groups. This difference was taken into consideration for memory analyses of Day 1 presented in the next section.
3.1.2 Day 1 – Encoding

Figure 3.1 shows the learning curve by groups. A 4 (Group) x 10 (Real Trials) mixed-ANOVA, with Real Trials as a within-subject factor, resulted in a main effect of Real Trial, \( F(5, 394) = 249.51, p < .001 \), partial \( \eta^2 = 0.78 \) showing effective learning, as measured by decreasing number of errors across real trials. We also observed a trend for the main effect of Group, \( F(3, 72) = 2.41, p = 0.74 \), but no significant Group X Real Trials interaction, \( F(16, 394) = 1.18, p = 0.28 \). These results were not changed when the STAI-state at Day 1, which was significantly different between the groups, was added as a covariate into the analysis: \( F(5, 390) = 4.54, p < .001 \), partial \( \eta^2 = 0.06 \) for main effect of Real Trials, \( F(3, 71) = 2.55, p = 0.06 \), for main effect of Group, \( F(5, 394) = 1.23, p = 0.24 \) for the Group X Real Trials interaction, and \( F(5, 394) = 0.66, p = 0.66 \) for the covariate X Real Trials interaction.

To explore the main effect of Group, which might suggest differences in learning for the groups, we averaged the last four trials of encoding and compared them between the groups (as did Forcato et al., 2007, 2009, 2011, 2013). As Figure 3.2 shows, although not statistically significant in a One-Way ANOVA (\( F(3, 72) = 2.11, p = 0.11 \)), it is clear that the groups had different levels of encoding by the end of Day 1, particularly groups R0 (\( M = 0.45, SD = 0.56 \)) and R4 (\( M = 0.92, SD = 0.79 \)) (\( t(35) = 2.14, p = 0.04 \), Cohen’s \( d = 0.65 \)). It is possible that the One-Way ANOVA did not reach significance due to the low sample sizes for groups R1 and R4, but we believe that this difference in encoding performance is real, and may be accompanied by a ceiling effect of learning (i.e., a floor
Figure 3.1 Encoding by trial. Error bars represent standard error of the mean (SEM).

Effect since performance is measured in errors). This was considered when analyzing the effects of the overall manipulation (# of reactivations) on Day 3.

Context learning was analyzed in a similar mixed-model 4 (Group) X 32 (Trials) ANOVA, with Trials as a within-subject factor. No main effect of Group, \(F(3,72)=1.16, p=0.33\) nor Group x Trials interaction was observed \(F(93,2232)=1.06, p=0.34\). As expected, a main effect of Trials was found, \(F(31,2232)=76.39, p<.001, \text{partial } \eta^2=0.52\), showing that all groups learned the context-syllable pairing at similar rates with increasing accuracy as the trials progressed. In a simplified manner, Figure 3.3 shows the percentage of correct context responses (out of all 32 trials) for all groups, which was not significantly different between the groups, \(F(3,72)=1.16, p=0.33\).
3.1.3 Day 2 – Manipulation

Participants were randomly assigned to a group after finishing Day 1. Participants either came back on Day 2 for 1, 2, or 4 reactivations or did not come back on Day 2 (skipped Day 2, 0 reactivations). The reactivation trial was composed of the context period, in which participants had to indicate whether the context presented was the context paired with the syllables the day before (Day 1) or not, followed by the ITE cue syllable for 2 seconds. It was designed so that participants could answer the context probe, but not the target syllable. Since group R0 skipped Day 2, they were not considered in this section’s analyses. Groups R1, R2, and R4 were not significantly different in their context response rate, \( F(2,51)=0.82, p=0.45 \), nor in their syllable
response rate $F(2,51)=2.15$, $p=0.13$. Note that these analyses are percentages of each group’s number of reactivation trials, i.e., group R1 is out of 1, group R2 is out of 2, and group R4 is out of 4 reactivation trials. In short, no differences in performance were found between the groups that received reactivation trials on Day 2.

3.1.4 Day 3 – Memory test

Figure 3.4 shows the test curve by groups. A 4 (Group) x 4 (Real Trials) mixed-ANOVA, with Real Trials as a within-subject factor, resulted in a main effect of Real Trial, $F(1, 135)=59.85$, $p<.001$, partial $\eta^2=0.45$ but no main effect of Group, $F(3,72)=1.55$,
Figure 3.4 Memory test at Day 3. Group R4 was significantly worse (i.e., more errors) at the first trial (T1), vs. Group R0 (p=0.07), vs. Group R1 (p=0.02), and vs. Group R2 (p=0.03). Error bars represent SEM.

$p=0.21$. Importantly, the interaction approached significance, $F(5, 135)=1.88$, $p=0.09$ ($p=0.057$ if sphericity is assumed, partial eta$^2=0.07$). Follow-up One-Way ANOVAs show that only the first test trial approached significance, $F(3, 76)=2.40$, $p=0.08$ ($F$'s<1.1, $p$'s>0.36 for test trials 2, 3, and 4). This difference was carried by group R4 (M=2.46, SD=1.66), who had more errors on test trial 1 than group R0 (M=1.63, SD=1.41, $p=0.07$), group R1 (M=1.24, SD=1.15, $p=0.02$), and group R2 (M=1.41, SD=1.14, $p=0.03$), although only group R1 survived multiple-comparisons correction (Bonferroni-corrected $p$-value for group R1 vs. group R4 $p=0.09$). We hypothesized multiple reactivations would result in better memory on Day 3. That is, group R4 (along with group R2) should have fewer
errors than groups R1 and R0; but the opposite pattern was found, R4 had more errors than R0. Nonetheless, the results reported here warrant caution because we observed that group R4 had more errors at the end of encoding in Day 1 (see section 3.1.2 and Figure 3.2 above). Therefore, a more appropriate way of assessing memory in this study was to use change scores, which would compare each individual’s performance on Day 3 with his or her own performance on Day 1, controlling for the group differences observed on Day 1. These analyses are presented in the next section.

Context learning at test (Day 3) was analyzed with a mixed-model 4 (Group) X 16 (Trials), with Trials as a within-subject factor. No main effect of Group, (F(3,72)=1.16, p=0.33) nor Group x Trials interaction was observed (F(18,442)=0.93, p<0.55). As expected, a main effect of Trials was found, F(6,442)=6.92, p<0.001, partial eta²=0.09), showing that all groups remembered the context-syllable pairing at similar rates, with increasing accuracy as the trials progressed. Similarly, percentage of correct context responses (out of all 16 test trials) was not significantly different between the groups, F(3,72)=0.98, p=0.41.

3.1.5 Change scores – Day 3 minus Day 1

A change score to compare performance from the end of Day 1 (encoding) to performance from the testing session at Day 3 was constructed by subtracting the average of the last four trials of encoding from the average of the four test trials (Figure 3.5). This comparison between the average of four trials of end of encoding and test was chosen because: 1) it is consistent with our previous analyses in section 3.1.2. (see
Figure 3.5 Change score, constructed by subtracting the average of the last four encoding trials from the average of the four test trials, by group. Group R2 was the only group who performed significantly better at test (Day 3). Error bars represent SEM.

Figure 3.2) and 2) using more trials should increase our variability, and therefore, help us overcome a possible ceiling learning effect (on Day 1). Although the One-Way ANOVA did not reach significance, F(3,72)=1.98, p=0.12, a priori tests were conducted. We had hypothesized that groups with multiple reactivations (R2 and R4) should have better performance on Day 3 than groups with a single (R1) or no reactivation (R0). As Figure 3.5 shows, group R2 (M=-0.17, SD=0.47) was the only group to improve from Day 1 to Day 3 (note that lower numbers are better since errors are being used to measure performance), reaching significance when compared to group R0 (M=0.23, SD=0.63), t(44)=2.43, p=0.02, Cohen’s=0.73, but not when compared to groups R1 (M=0.07, SD=0.54) and R4 (M=0.12, SD=0.62), t’s<1.53, p’s>0.13.
Sex has been reported to affect cognition (see D. I. Miller & Halpern, 2014), including memory performance, and because sex was not distributed equally between our groups (see Table 3.1) we added sex as a covariate to the change score ANOVA analysis. Similar to the results above, the One-Way ANOVA did not reach conventional significance, $F(3,71)=2.29$, $p=0.08$, partial $\eta^2=0.09$. Importantly, the sex covariate did not have a significance effect in this analysis, $F(1,71)=2.44$, $p=0.12$. In short, sex had no effect on the present results.

3.1.5.1 Change scores for specific types of errors – Day 3 minus Day 1

With the current task, errors can be divided into three specific types: 1) void errors, when nothing was typed in, 2) intralist errors, when the response-syllable was incorrect, but was included in the list, and 3) confusion errors: when the response-syllable was not included in the list. Analysis of performance with this breakdown, although exploratory, allowed for deeper understanding of the phenomenon studied in the current study.

Change scores, similar to those reported above, were used for these analyses; that is, we subtracted the average of the last four trials of encoding from the average of the four test trials for each error type, for each individual.

A One-Way ANOVA for void errors strongly trended towards significance, $F(3,72)=2.67$, $p=0.054$, partial $\eta^2=0.10$. Contrary to our predictions, post-hoc, uncorrected comparisons showed that group R4 ($M=2.4$, $SD=1.63$) had a higher change score (i.e., worse memory) than groups R0 ($M=1.51$, $SD=1.43$, $p=0.05$), R1 ($M=1.21$, $SD=1.43$), and R2 ($M=1.62$, $SD=1.54$).
SD=1.43, p=0.01), and R2 (M=1.26, SD=1.07, p=0.01) (Figure 3.6). Intralist and confusion errors were not significantly different between the groups, F(3,72)=0.77, p=0.51 for intralist, and F(3,72)=0.54, p=0.66 for confusion, which is unsurprising considering the small range of values for these two error types (for intralist M=-0.11-0.02, for confusion M=-0.02-0.13). The pattern of results for group R4, which is in the opposite direction as predicted was surprising and is considered in the next section.

3.1.6 End-of-encoding performance as a covariate

Another way of addressing the fact that learning was already significantly different by the end of Day 1 is by using performance at Day 1 as a covariate in the analyses of the memory test at Day 3. A 4 (Group) x 4 (Real Trials) mixed-ANCOVA, with Real Trials as a within-subject factor and the average of the last four encoding trials as a covariate resulted in a main effect of Real Trial, F(2, 136)=17.39, p<.001, partial \( \eta^2 \) =0.18, but no main effect of Group, F(3,71)=1.66, p=0.18. The Covariate x Real Trials interaction was also significant, F(2, 136)=4.50, p=0.01, partial \( \eta^2 \)=0.06, showing that differences in the Real Trials of Day 3 (test) are in part explained by differences already present at the end of Day 1 (the averaged last four Real Trials of encoding). Importantly, the Group x Real Trials interaction, which was nearing significance in the previous analysis in section 3.1.4, was not significant, F(8, 136)=1.58, p=0.16, showing that this ANCOVA was not able to detect the differences detected by the change score in section 3.1.4.
3.1.7 Interim discussion for Experiment 1

Experiment 1 aimed to replicate the finding that multiple reactivations (groups R2 and R4) strengthen a memory whereas a single or no reactivation does not (groups R0 and R1). We observed that one of the multiple-reactivations group, R2, did result in better memory on Day 3 compared to a single or no reactivation, albeit only significantly so when compared to the no reactivation group (R0). Also, this was only observed when using our change score, i.e., when memory was assessed in reference to how much it changed from Day 1 to Day 3. Next, provided is some explanation as to why this pattern of results may have occurred in experiment 1.
First, in experiment 1, it is clear that a majority of our participants reached ceiling levels of learning (Ms=85%-91%, SDs=11%-16% compared to 85% in Forcato et al., 2011). This is a particularly troublesome issue for the current experiment since we were expecting to replicate a memory strengthening effect, and if a memory has hit ceiling levels of performance, there is little room for improvement that can be caused by our manipulation (i.e., multiple reactivations). Although speculative, a possible reason for this can be explained by the overachieving nature of undergraduate students at the University of Notre Dame. This has been a problem in other studies in our laboratory. Also, ten real trials of learning might have been too many, which is the main reason why experiment 2 included only five real trials (see below).

Second, although randomization was used for group assignment, the groups achieved different levels of learning (on average) by the end of encoding (Day 1), which is why a change score was used. For this reason, participants were pseudorandomized into groups for experiment 2, depending on their performance at the end of Day 1 (i.e., the average level of performance at the end of Day 1 was kept as similar as possible throughout all four groups).

Finally, it is possible that the strengthening effect was not observed in group R4 for three reasons. First, low power as sample size was much lower for group R4 (n=13) than for group R2 (n=22) due to time constraints. Second, it is possible that some of the participants in group R4 were able to actually input (type in) the answer for the cue syllable since they had four chances to do so, an action that according to Forcato and colleagues (2011;2013) does not labilize the memory, and therefore, should not
strengthen the memory. This is unlikely though, since no differences were found in our analysis of correct responses on Day 2 (see section 3.1.3 above). Third, although complete speculation, it is possible that suspicion of the verity of the reactivation “error” may be a boundary condition for labilization>reconsolidation>strengthening to occur in the current paradigm. That is, participants in group R4 had to sit through four “errors” and may have figured out that it was not a real interruption, but an experimental farce. These factors should be considered in future experiments. For example, with large enough sample sizes, post-hoc analyses of experimental suspicion regarding Day 2 may shed some light on this matter.

3.2 Experiment 2: Testing the strengthening effect against a stressor before memory retrieval

Experiment 2 aimed to test the efficacy of the ‘strengthening via multiple reactivations’ effect by using an ecologically valid psychosocial stressor (TSST) before the final memory test during Day 3. A 2 (Stress condition: stress vs. control) x 2 (Reactivations: 0 vs. 2) design was used for this study which resulted in four groups: group R0C received zero reactivations (skipped Day 2) and went through the control version of the TSST on Day 3, group R0S received zero reactivations and went through the TSST on Day 3, group R2C received 2 reactivations on Day 2 and went through the control version of the TSST on Day 3, and group R2S received 2 reactivations and went through the TSST. Two and zero reactivations were chosen because these conditions have been used more consistently by the Forcato group (C. Forcato, personal
communication, March 18 2017; see Fernández, Bavassi, Forcato, & Pedreira, 2016; Forcato et al., 2013, 2011). Additionally, this comparison was the only one that reached significance in our replication experiment (see Figure 3.5 above). We expected stress to overall impair memory (main effect of stress), but that 2 reactivations would, at least, protect the memory at Day 3.

3.2.1 Descriptive and control variables

Forty-seven (47) participants were recruited for experiment 2, but 17 participants were excluded from analyses due to bad (<60% in last two trials) syllable-pair learning (9 participants) and missing Day 3 (8 participants), for a final sample size of N=30. The age range was 18-23 with M=19.78, SD=1.41, 16 females. Due to the low sample size, low context learning was not used as an exclusion criterion for experiment 2. Table 3.2 presents sample size and demographic information for each group.

Mood and anxiety measures were analyzed with a 2 (Stress) x 2 (Reactivations) Two-Way ANOVA for each measure. The STAI-trait, a measure of trait anxiety, was taken at the beginning of Day 1. No significant effects in trait anxiety were found, for main effect of Stress F(1,24)=0.19, p=0.67, for main effect of Reactivations F(1,24)=0.99, p=0.33, nor the interaction F(1,24)=0.90, p=0.35. The PANAS, a measure of state positive and negative affect, as well as the STAI-state, were taken at the beginning of Day 1 and at the beginning of Day 3. No differences (main effects or interactions) were found in positive (PA), negative affect (NA), or state anxiety on Day 1, F’s<1.91, p’s>0.18. For Day 3, only PA resulted in a main effect of Reactivations, F(1,23)=5.28, p=0.03, partial
eta^2=0.18, showing that participants that received 2 reactivations (on Day 2) had higher positive affect (M=27.73, SD=7.17) on Day 3 than those who received no reactivations (M=21.40, SD=7.82). This variable was taken into consideration in the memory analyses for Day 3, presented below in section 3.2.4.2. No other main effects or interactions were significant for Day 3, F’s<2.38, p’s>0.14.

### TABLE 3.2

DEMOGRAPHIC INFORMATION, BY GROUP

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Males</th>
<th>Females</th>
<th>Age M</th>
<th>Age SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC</td>
<td>8</td>
<td>2</td>
<td>6</td>
<td>20.13</td>
<td>1.73</td>
</tr>
<tr>
<td>ROS</td>
<td>7</td>
<td>3</td>
<td>4</td>
<td>19.29</td>
<td>1.25</td>
</tr>
<tr>
<td>R2C</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>19.14</td>
<td>0.69</td>
</tr>
<tr>
<td>R2S</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>20.38</td>
<td>1.51</td>
</tr>
</tbody>
</table>

#### 3.2.2 Day 1 – Encoding

Figure 3.7 shows the learning curve by groups. A 2 (Stress) x 2 (Reactivations) x 5 (Real Trials) mixed-ANOVA, with Real Trials as a within-subject factor, resulted in a main effect of Real Trial, F(3,76)=135.59, p<.001, partial eta^2=0.97, showing effective learning, as measured by decreasing number of errors across real trials. No other main effect was significant, for main effect of Stress, F(1,26)=0.01, p=0.92 and for main effect of Reactivations, F(1,26)=0.31, p=0.59. No interaction was significant, for 3-way interaction, F(3, 76)=1.43, p=0.24, for Real Trials x Stress, F(3, 76)=0.86, p=0.46, for Real
Trials x Reactivations, $F(3, 76)=0.76$, $p=0.52$, for Stress x Reactivations, $F(1, 26)=0.54$, $p=0.47$. These results show that syllable-pair learning was comparable in all four groups.

Figure 3.8 shows the average of the mean number of total errors of the last two real trials of encoding. We used the last two trials, as opposed to the last four as was the case for experiment 1, because encoding was halved for experiment 2 (from 10 to 5 real trials). Supporting the above analysis, a 2 (Stress) x 2 (Reactivations) Two-Way ANOVA revealed no interaction, $F(1,26)=0.09$, $p=0.76$, main effect of Stress, $F(1,26)=0.09$, $p=0.76$, or main effect of Reactivations, $F(1,26)=0.01$, $p=0.91$, once again, showing that learning was comparable between groups. Context learning was analyzed in a similar mixed-model 2 (Stress) x 2 (Reactivations) x 16 (Trials) ANOVA, with Trials as a within-
subject factor. As expected, a main effect of Trials was found, $F(15,390)=22.85$, $p<.001$, partial $\eta^2=0.98$, showing that all groups learned the context-syllable pairing at similar rates with increasing accuracy as the trials progressed. No other main effects were found, for Stress, $F(1,26)=0.01$, $p=0.98$, for Reactivations, $F(1,26)=0.21$, $p=0.65$. The 3-way interaction trended towards significance, $F(15,390)=1.61$, $p=0.07$, partial $\eta^2=0.06$, reflected in a similar trend for the Trials x Reactivation interaction, $F(15,390)=1.64$, $p=0.06$, partial $\eta^2=0.06$. The Trials x Stress interaction was not significant, $F(15,390)=1.14$, $p=0.31$, as well as the Stress x Reactivation interaction, $F(1,26)=0.45$, $p=0.51$. Follow-up One-Way ANOVAs were performed, comparing reactivation conditions (0 vs. 2), for each trial (16 total real trials). With a family-wise alpha of
0.05/16 = 0.003, none of the ANOVAs survived the Bonferroni correction for multiple comparisons (the closest was trial #11 at F(1,28)=6.15, p=0.02). In a simplified manner, Figure 3.9 shows the percentage of correct context responses (out of all 16 trials) for all groups. A 2 (Stress) x 2 (Reactivations) Two-Way ANOVA showed no interaction, F(1,26)=0.45, p=0.51, no main effect of Stress, F(1,26)=0.01, p=0.98, and no main effect of Reactivations, F(1,26)=0.21, p=0.65. In short, all groups learned the syllable-context pairing at comparable rates.

3.2.3 Day 2 – Reactivation manipulation

Participants were randomly assigned to a reactivation condition after finishing Day 1. Participants either came back on Day 2 for 2 reactivations/reminders or did not come back on Day 2 (skipped Day 2, 0 reactivations). The reactivation trial was exactly the same as in experiment 1. Because participants did not come back for Day 2 (i.e., groups ROC and ROS), the Reactivations variable was taken out of this section’s analyses. A One-Way ANOVA using only the Stress variable showed that control participants, compared to the stress participants, were not significantly different in their context response rate, F(1,28)=0.14, p=0.71, nor in their syllable response rate F(1,28)=2.15, p=0.15. That is, all participants who received the reactivation manipulation performed similarly during the reactivation trial on Day 2.
3.2.4 Day 3 – Stress manipulation and memory test

Participants were randomly assigned to the stress or control condition at the beginning of Day 3. Following this procedure, they were walked to the experimental room to complete a final memory test. First, we analyzed the Self-Assessment Manikin (SAM) scale to check whether the stressor (i.e., TSST) was effective. Following, results of the memory test are presented.

3.2.4.1 Stress manipulation check

Subjective (SAM scale) and objective (salivary cortisol) measures were used to analyze the effectiveness of our stressor. The SAM scale, which asked for arousal ratings (calm [1] to excited [9]) and valence ratings (negative [1] to positive [9]) before and after
the TSST/control, was used to assess the effectiveness of our stressor, the TSST. Note that 2 participants were removed from these analyses because they did not complete the SAM adequately (i.e., final sample n=28).

Figure 3.10 shows the arousal ratings from the SAM, by group. A 2 (Stress) x 2 (Reactivations) x 2 (Time: pre vs. post) mixed ANOVA revealed no 3-way interaction, F(1,24)=0.34, p=0.57, no Time x Reactivations interaction, F(1,24)=0.39, p=0.54, nor a main effect of Reactivations, F(1,24)=1.09, p=0.31. Significant main effects of Stress, F(1,24)=9.30, p=0.006, partial eta²=0.28, and Time, F(1,24)=22.75, p<0.001, partial eta²=0.49, were found, but importantly, the Time x Stress was also significant, F(1,24)=29.01, p<0.001, partial eta²=0.55. These patterns are confirmed with pairwise comparisons, in which the control and stress participants rated their arousal similarly before the manipulation (pre), t(27)=0.22, p=0.83, but the stress participants (M=7.20, SD=1.08) were more aroused/excited than controls (M=4.54, SD=1.71) after the manipulation (post), t(26)=4.98, p<.001, Cohen’s d=1.95. Finally, the Stress x Reactivations interaction was also significant, F(1,24)=6.59, p=0.02, partial eta²= 0.22, which suggests that the TSST did not affect both stress groups similarly or that the control-TSST did not affect both control groups similarly.

To explore this effect One-Way ANOVAs, with all four groups were conducted separately for pre-arousal ratings and post-arousal ratings. For pre-arousal, the ANOVA tended towards significance, F(3,25)=2.47, p=0.09, but none of the individual comparisons survived the Bonferroni correction. Importantly, for post-arousal, the ANOVA was strongly significant, F(3,24)=10.94, p<0.001, partial eta²=0.06. Bonferroni-
Figure 3.10 Arousal ratings from the SAM. Control groups rated their arousal levels similarly before (pre) and after (post) the control-TSST (paired t’s n.s.). Arousals increased for the stress groups (paired t’s>2.69, p’s<0.03). Error bars represent SEM.

corrected comparisons showed that the R2S group was significantly different from both control groups (p<0.001 vs. R2C and p=0.04 vs. R0C), but R0S was only significantly different from R2C (p=0.001). Importantly, the R0S and R2S groups were not significantly different from each other, p=0.99). As Figure 3.10 shows, whereas no change between pre and post (paired t-tests) was found for the control groups (t(7)=0.23, p=0.83 for R0C, t(6)=0.21, p=0.84 for R2C), an increase in arousal was found for the stress groups (t(7)=2.69, p=0.03 for R0S, Cohen’s d=1.63, t(8)=5.93, p<.001, Cohen’s
d=3.59 for R2S). In short, our stressor was effective in increasing arousal, whereas our control condition did not.

Figure 3.11 shows the valence ratings from the SAM, per group. A 2 (Stress) x 2 (Reactivations) x 2 (Time: pre vs. post) mixed ANOVA revealed no 3-way interaction, $F(1,24)=0.12, p=0.73$, no Time x Reactivations interaction, $F(1,24)=1.54, p=0.23$, nor a main effect of Reactivations, $F(1,24)=1.44, p=0.24$. Significant main effects of Stress, $F(1,24)=4.61, p=0.04$, partial eta$^2=0.16$, and Time, $F(1,24)=13.17, p<0.001$, partial eta$^2=0.35$, were found, but importantly, the Time x Stress was also significant, $F(1,24)=4.15, p=0.05$, partial eta$^2=0.15$. These patterns are confirmed with pairwise comparisons, in which the control and stress participants rated their valence similarly before the manipulation (pre), $t(26)=1.21, p=0.24$, but the stress participants ($M=4.33, SD=1.76$) rated their experience as more negative (i.e., 1=most negative, 9=most positive) than controls ($M=6.00, SD=1.30$) after the manipulation (post), $t(26)=2.88, p=0.008$, Cohen’s $d=1.13$. No Stress x Reactivation interaction was found, $F(1,24)=0.05, p=0.83$. Finally, as Figure 3.11 shows, whereas no change between pre and post (paired t-tests) was found for the control groups ($t(7)=0.68, p=0.52$ for R0C, $t(6)=1.92, p=0.10$ for R2C), a decrease in valence (i.e., rated more negative) was found for the stress groups ($t(7)=3.21, p=0.02$, Cohen’s $d=1.94$ for R0S, $t(8)=2.83, p=0.02$, Cohen’s $d=1.71$ for R2S). In short, our stressor was effective in making participants feel more negatively (i.e., decrease in our valence scale), whereas our control condition did not.
Figure 3.11 Valence ratings from the SAM. Control groups rated their arousal levels similarly before (pre) and after (post) the control-TSST (paired t’s<1.92, p’s>0.10). Valence decreased (more negative) for the stress groups (paired t’s>2.83, p’s=0.02). Error bars represent SEM.

As mentioned in the Procedure (section 2.3), five saliva samples were collected throughout this experiment. Sample 1 was taken at the beginning of Day 1, sample 2 was taken at the beginning of Day 2 (for those in the R2 groups), sample 3 was taken on Day 3 before the stressor, sample 4 was taken immediately after the stressor, and sample 5 was taken after the memory test, at the end of the experiment (about 45 minutes after the beginning of the stressor). Note that because the samples for two participants were misplaced, the final sample size was N=28. Figure 3.12 (top) shows the timeline for all groups, whereas Figure 3.12 (bottom) shows the timeline collapsing both
Figure 3.12 Salivary cortisol throughout the experiment for all groups (top) and collapsed by stress/control condition (bottom). Stress (ROS and R2S groups collapsed) increased cortisol immediately after the stressor ($p = 0.01$). Errors represent SEM.
control groups and both stress groups. Day 1 and Day 2 cortisol was not different between the groups, F(3,24)=1.03, p=0.39 for Day 1, F(3,24)=1.03, p=0.33 for Day 2 between groups R2C and R2S. Therefore, because Day 1 and Day 2 cortisol levels were similar, and to increase the power to detect a meaningful interaction, a 2 (Stress) x 2 (Reactivations) x 3 (Day 3 Time points) mixed ANOVA, with Time points as a within-subject factor, was conducted. A main effect of Time points was found, F(2,35)=7.31, p=0.005, partial eta²=0.23, as well as a trend for a main effect of Stress, F(1,24)=3.85, p=0.06, partial eta²= 0.14, but no main effect of Reactivations, F(1,24)=2.59, p=0.12. No Time point x Reactivation interaction was found, F(2,34)=2.44, p=0.59. Importantly, a Time point x Stress interaction was found, F(2,34)=6.10, p=0.01, partial eta²=0.20, showing that before (t(26)=0.28, p=0.78) the stressor and by the end of the experiment (t(26)=1.86, p=0.08) stress and controls did not have significantly different levels of salivary cortisol (see Figure 3.12 bottom). Only immediately after the TSST did stress participants (M=6.33, SD=5.04) had significantly higher levels of salivary cortisol than control participants (M=2.25, SD=2.60, t(26)=2.69, p=0.01, Cohen’s d=1.06). No 3-way interaction was found, F(2,35)=1.45, p=0.25. Thus, our stressor was effective at increasing cortisol levels in the stress groups participants.

3.2.4.2 Memory test on Day 3

Figure 3.13 shows the test curve by groups. A 2 (Stress) x 2 (Reactivations) x 4 (Real Trials) mixed ANOVA, with Real Trials as a within-subject factor, resulted in a significant main effect of Real Trials, F(2,51)=56.59, p<0.001, partial eta²=0.78 showing
that participants committed fewer errors as the test went on. No other main effects
were found, for Stress, $F(1,26)=0.04$, $p=0.85$, or for Reactivations, $F(1,26)=0.04$, $p=0.85$.
Importantly (and surprisingly), no interactions were significant, for Stress x
Reactivations, $F(1,26)=0.27$, $p=0.61$, for Real Trials x Reactivations, $F(2,51)=1.67$, $p=0.20$,
for Real Trials x Stress, $F(2,51)=1.01$, $p=0.37$, and for the 3-way interaction, $F(2,51)=0.54$,
$p=0.58$.

Because positive affect (PA), as measured by the PANAS, was significantly
different at Day 3 (see section 3.2.1), we added this variable as a covariate to the above
analyses. Although virtually all main effects and interactions stayed the same (i.e., not
significant), and the covariate main effect ($F(1,25)=0.17$, $p=0.69$) and the Real Trials x
covariate interaction ($F(2,49)=0.65$, $p=0.52$) were not significant, the main effect of Real
Trials lost its statistical significance, $F(2,49)=2.34$, $p=0.11$. This may suggest that, although not related to the lack of group differences per se, this difference in PA during Day 3 may explain some of the variance regarding the memory performance curve during test at Day 3.

3.2.5 Change scores – Day 3 minus Day 1

For experiment 2, the change score was created by subtracting the average of the last two trials of encoding from the average of the first two test trials (Figure 3.8). Note the difference from the change score used for experiment 1, which used the average of the last four encoding trials compared to the average of all (four) test trials. As mentioned above, the last two trials were used because encoding was halved for experiment 2 (from 10 to 5 real trials). Additionally, although we do not assume a ceiling effect on learning by the end of Day 1 for experiment 2, this is another useful way of looking at the data, and may account for subtle differences in Day 1 learning that statistical analyses cannot detect.

A 2 (Stress) x 2 (Reactivations) Two-Way ANOVA revealed no main effect of Stress, $F(1,26)=0.02$, $p=0.89$, no main effect of Reactivations, $F(1,26)=0.29$, $p=0.56$, and more importantly, no Stress x Reactivations interaction, $F(1,26)=0.09$, $p=0.77$. Adding PA on Day 3 as a covariate made no difference: main effect of covariate, $F(1,25)=0.01$, $p=0.99$, main effect of Stress, $F(1,25)=0.02$, $p=0.89$, main effect of Reactivations, $F(1,25)=0.23$, $p=0.63$, Stress x Reactivations interaction, $F(1,25)=0.08$, $p=0.78$. In short,
we observed no effect of our stressor in memory retrieval, nor was the ‘strengthening via multiple reactivations’ effect replicated.

3.2.5.1 Change scores for specific types of errors – Day 3 minus Day 1

As in experiment 1, with the current task, errors can be divided into three specific types: void errors, intralist errors, and confusion errors. Analysis of performance with this breakdown, although exploratory, allowed for deeper understanding of the phenomenon studied in the current study. Change scores, similar to those reported above, were used for these analyses; that is, we subtracted the average of the last two trials of encoding from the average of the first two test trials for each error type, for each individual.

2 (Stress) x 2 (Reactivations) Two-Way ANOVAs were conducted for each error type. For void errors, no significant effects were found, for main effect of Stress, F(1,26)=0.12, p=0.73, for main effect of Reactivations, F(1,26)=0.40, p=0.53, for the interaction, F(1,26)=0.15, p=0.70. For intralist errors, no significant effects were found, for main effect of Stress, F(1,26)=1.27, p=0.27, for main effect of Reactivations, F(1,26)=0.01, p=0.94, for the interaction, F(1,26)=0.14, p=0.71. For confusion errors, no significant effects were found, for main effect of Stress, F(1,26)=0.04, p=0.84, for main effect of Reactivations, F(1,26)=0.04, p=0.84, for the interaction, F(1,26)=0.12, p=0.73. In short, there were no condition or group differences in any of the specific error types in experiment 2. The following section presents some possibilities as to why experiment 2 did not yield the hypothesized results.
3.2.6 Interim discussion for Experiment 2

Experiment 2 aimed to test the magnitude of the strengthening effect via multiple reactivations by disrupting retrieval via psychosocial stress. In this experiment, we expected to replicate the strengthening finding that two reactivations would lead to a stronger memory (i.e., less errors on Day 3 for the R2 groups), compared to no reactivations, as well as memory to be impaired by our stressor on Day 3 (i.e., more errors for the stress groups). The current data did not support either of those hypotheses.

Due to time constraints, but mostly due to high levels of attrition (8/47, 17%) and a high number of poor performers during encoding at Day 1 (9/47, 19%), the final sample size for this experiment was N=30 for four groups (7-8 participants per group, see Table 3.2). Although, in theory, power analysis with G*Power resulted in n=6 per group (at least for experiment 1, see section 2.1), it would have been irresponsible to adhere to those numbers. In practice, experiment 2 suffered from lack of power with such a small sample size; conventional cognitive psychology/neuroscience experiments have group sample sizes of n=20 at least. To fully answer whether the strengthening effect can be observed under the TSST/control conditions, and whether stress at retrieval can negatively affect performance in this task, a larger sample size is needed.

Another possibility, at least for the lack of a detrimental main effect of stress on memory, is that our stressor was not effective enough. Nonetheless, our manipulation check analyses do not support this idea (see Figures 3.10, 3.11, 3.12). Additionally, the rich context may have served as a strong retrieval cue, which could have counteracted
against the detrimental effect of stress at retrieval. Further discussion of these ideas is provided in the general Discussion (next section).

3.3 Collapsed data for experiments 1 and 2

In order to answer whether multiple reactivations strengthened a memory compared to no reactivations and to increase overall statistical power, the samples from both experiments were compiled and analyzed. Specifically, groups R2, R4, and group R2C were compiled into a Multiple Reactivations (MR, n=42) group and groups R0 and R0C were compiled into a Zero Reactivations (0R, n=32) group. For these analyses, the same proportion of the encoding session was used for all groups. That is, for participants in experiment 1, the last four encoding trials (E7-E10) were used, whereas for experiment 2, the last two encoding trials (E4-E5) were used. The logic behind this is that while using the last two trials in both datasets would be numerically consistent, we believe these two comparisons are not comparable in practice. Using trials 7-10 for experiment 1 and using trials 4-5 for experiment 2 entails the last 40% of the encoding session at Day 1. In contrast, using trials 9-10 for experiment 1 would entail using only the last 20% of the encoding session and it would be inappropriate to compare it to the 40% (trials 4-5) of experiment 2. In short, using the average of E7-E10 for experiment 1 and E4-E5 for experiment 2 should, in practice, reflect the most similar levels of learning between the different experiments.

The groups were similar by the end of encoding, t(72)=1.50, p=0.14, and at test, t(72)=0.09, p=0.93. Nonetheless, when a change score similar to those in the sections
above was used, the OR group (M=0.39, SD=0.86) had a decrease in performance (i.e., more errors), whereas the MR group (M=0.04, SD=0.65) had virtually no change from the end of Day 1 to Day 3, t(72)=1.99, p=0.05, Cohen’s d=0.47 (Figure 3.14). In short, although a small-to-medium effect (0.35 difference, Cohen’s d=0.47), multiple (2 and 4) reactivations led to a stronger memory compared to no reactivations.

![Figure 3.14](image.jpg)

Figure 3.14 Change score from the end of encoding to test, by group. The Multiple Reactivations group performed significantly better than the 0 Reactivations group. Error bars represent SEM.
4.1 Summary of results and overview

Memory reconsolidation is the process by which stable memories can be updated/modified (Dudai & Eisenberg, 2004). The current project aimed to investigate one of the hypothesized functions of reconsolidation: memory strengthening (Forcato et al., 2014; Lee, 2008; Sara, 2000a).

In the first experiment, we aimed to show (replicate) a memory strengthening effect via reconsolidation mechanisms in an episodic memory paradigm (Forcato et al., 2011, 2013). We reliably showed that multiple reactivations, or at least two reactivations, compared to a single or no reactivation, result in memory improvement over a three-day period. Nonetheless, four reactivations did not show this effect.

The second experiment was designed to further explore the ‘strengthening-via-multiple-reactivations’ phenomenon, by testing the efficacy of the memory strengthening by applying a psychosocial stressor before retrieval, an intervention that is known to reliably impair memory retrieval (see Vogel & Schwabe, 2016). In our study, stress before retrieval in Day 3 did not have a detrimental effect on memory performance. In addition, multiple (2) reactivations on Day 2 did not result in better memory performance in Day 3 compared to no (0) reactivations in Day 2. Section 4.3
presents further discussion into why the expected results were not observed in experiment 2. Strengthening via reconsolidation mechanisms has implications for real-life learning, as well as pinning this process against a naturalistic stressor such as the Trier Social Stress Test (TSST), which could emulate stress levels comparable to those elicited by real-life experiences, such as exams.

4.2 Discussion of experiment 1

Reconsolidation is thought to provide two functions to adaptive learning: memory modification or updating (see Lee, Nader, & Schiller, 2017) and memory strengthening (see Forcato et al., 2014). Animal work has used contextual fear conditioning to show that memory strengthening is dependent on reconsolidation, not consolidation, mechanisms when two trials are separated by a day (i.e., one trial each day). When both trials were presented in a single session (one day), memory strengthening recruited consolidation (brain-derived neurotrophic factor), but not reconsolidation mechanisms (zif268) (Lee, 2008). In another study using inhibitory avoidance (IA) learning, Inda and colleagues (2011) showed that, whereas three reactivations by full test trials resulted in extinction of the IA memory (shown by reduced IA memory), three reactivations trials by 10 sec of context-only presentation resulted in increased IA memory. Of importance, note that each session, including the reactivation sessions, were separated by two days. Furthermore, this effect was dependent on protein synthesis, as injection of cloheximide, a protein-synthesis inhibitor, after each reactivation disrupted memory performance to pre-training levels.
Evidence from animal models supports the general idea that memory strengthening is a possible function of reconsolidation, but it is important to point out differences in the animal (Inda et al., 2011; Lee, 2008) and human work by Forcato and colleagues. Most obvious is the nature and complexity of the memory systems tested. As most reconsolidation research to date, the animal work presented above investigated associative fear memories, for which both molecular- and systems-levels processes have been well documented. Forcato and colleagues (2011, 2013), as well as the current studies, looked at declarative-episodic memories, which are more complex in nature, and the processes, particularly molecular processes, are not as well understood. Nonetheless, the Forcato task does have strong associative learning components: the pairs of nonsense syllables, as well as the context learning specific to the presentation of the syllable pairs. Most importantly, in both animal studies reported above, reactivation trials were presented in separate days, whereas in the Forcato experiments, they were presented temporally close to each other. In fact, a boundary factor for this strengthening effect is the condition that the reactivation trials are presented within the ‘reconsolidation window’ (Duvarci & Nader, 2004); that is, reminders presented within 5 minutes or 2hr result in strengthening but not when they are presented 24hr apart (experiments 4 and 5 in Forcato et al., 2011). In agreement, Wichert and colleagues found that, although reconsolidation of whole scenes can be disrupted by new learning after reactivation (Wichert et al., 2013b), memory strengthening does not occur if multiple reactivations are separated by 24hr (Wichert et
al., 2013a). In short, the time-dependent boundaries for memory strengthening seem to be different for distinct memory systems and/or species.

4.2.1 Mechanisms for reconsolidation, and its strengthening function

By now, it is well accepted that reconsolidation and consolidation are, although similar, quite distinctive processes (see Tronson & Taylor, 2007). Reconsolidation is a two-phase process that starts with memory destabilization/labilization and finishes (if uninterrupted) with memory re-stabilization. In this section, possible mechanisms for episodic memory reconsolidation will be discussed, first in the viewpoint of the general labilization/restabilization process, and then in the viewpoint of the strengthening function specifically.

Schwabe and colleagues (Schwabe, Nader, Wolf, Beaudry, & Pruessner, 2012) found that propranolol (a beta-adrenergic receptor antagonist) after reactivation, but not propranolol only or reactivation only, reduced emotional memory. In this study, participants were in an fMRI scanner during reactivation and during the final memory test. Reactivation + propranolol resulted in increased hippocampus and amygdala (crucial regions for emotional memory processing) activation during test, showing that these brain regions needed a stronger activation for successful memory retrieval. More importantly, these same structures were activated during reactivation, independent of whether participants received propranolol or placebo, which shows that this activity is not related to drug administration. In a recent experiment, Forcato and colleagues (2016) found similar effects with another episodic memory task. In this task, participants
had to associate full words with a background picture. Memory was reactivated on Day 2 and a second word list was learned to disrupt the original memory, inside the scanner. Memory was tested on Day 3, outside the scanner. As in their previous studies (Forcato et al., 2009), different types of reminders were presented, and only the one that included a mismatch component (incomplete word + sudden interruption) triggered labilization/reconsolidation. More importantly, that kind of reminder recruited the left hippocampus more than the other conditions. This supports Schwabe et al.’s findings and is in agreement with animal work that suggests a crucial role for the hippocampus during reconsolidation (Debiec et al., 2002; Inda et al., 2011; Milekic & Alberini, 2002; Morris et al., 2006), although note that most of this work involves contextual fear conditioning, which relies heavily on the hippocampal system. It is possible that the role of the hippocampus (particularly the left hippocampus) in the labilization/reconsolidation of episodic memories is that of identifying a prediction error (Kumaran & Maguire, 2006, 2007). Kumaran and colleagues found that the left hippocampus was more strongly activated when participants encountered objects out of order (2006) or out of position (2007). These findings reiterate the importance of the mismatch component in the labilization/reconsolidation of memories (Pedreira et al., 2004).

Non-invasive brain stimulation during reactivation, coupled with fMRI during test, have also been used to study reconsolidation mechanisms at the systems level in humans. Using a procedural (finger-tapping) memory task, Censor and colleagues (Censor, Dayan, & Cohen, 2014) found that after disrupting reconsolidation-dependent
memory modification with stimulation to the primary motor cortex (M1), participants had decreased activation of the supplementary motor area. More importantly, reconsolidation disruption via stimulation also decreased functional connectivity between early-learning areas (such as the cerebellum) and late-learning areas (such as the striatum), suggesting that reconsolidation recruits brain regions involved in early and late stages of memory processing. If this is also the case for episodic memories, particularly for the task used in our experiments, we would expect that memory reactivation/reconsolidation would recruit early-learning regions, like the hippocampus and sensory cortices (due to the enriched context in this particular task), as well as late-learning regions, like parahippocampal regions and a more distributed neocortical network that would include the prefrontal cortex (PFC). This is supported by studies that found increased hippocampal activity during reactivation of episodic memories (Forcato et al., 2016; Schwabe, Nader, et al., 2012) and studies that show casual involvement of the PFC, by stimulating the PFC after reactivation, resulting in memory enhancement (Sandrini et al., 2013).

Regarding the precise mechanisms for a strengthening function specifically, direct research is lacking. Forcato and colleagues (2011) suggested that, at least for episodic memories, since no strengthening was found when the subsequent reactivation was presented outside of the reconsolidation window (somewhere between 2-24hr in this task), for memory strengthening to occur, reactivation-dependent processes are required to be ‘mounted on’ previous reactivation-dependent processes. That is, this nested mechanism would result in a repeated cascade of molecular pathways that result
in either a higher expression or increased number of macromolecules (e.g., AMPA receptors) available for the retrieval of information. Animal models that parallel this phenomenon are needed to test this mechanism.

It has also been suggested that reconsolidation occurs at a synaptic level (Hu & Schacher, 2014), and therefore, that it regulates plasticity in a more general way (i.e., not only regarding learning and memory) in the central nervous system (Bonin & De Koninck, 2015). Hu and Schacher (2014) found that in cell cultures of the slug, *Aplysia*, after inducing persistent synapses (i.e., LTP), reconsolidation was triggered only by serotonin application, which induces transient heterosynaptic potentiation (i.e., activation of many synapses), not by direct stimulation, which induces transient homosynaptic potentiation (i.e., a single synapse at a time) (Figure 4.1). Not every synapse reactivation induces synaptic reconsolidation (Figure 4.1B). Similarly, in behavior, reconsolidation has boundary conditions, i.e., not every reactivation/retrieval triggers reconsolidation. It may be that in situations where reactivation does not trigger reconsolidation the behavioral stimuli failed to cause sufficient synaptic activation (Figure 4.1D). This might explain why repeated labilization/reconsolidation cycles are needed to strengthen a memory via reconsolidation mechanisms. Assessing the veracity of these ideas poises a future challenge, even for nonhuman research, as animal models that are able to measure complicated behavior in a complicated neuronal system (such as conditioning or IA in rodents) cannot stimulate single synapses (the closest method being optogenetics, which still stimulates hundreds of neurons simultaneously).
Figure 4.1 Taken from Bonin and Konick, 2015 explaining Hu and Schacher, 2014. In Aplysia, reconsolidation was triggered only by stimuli that generated heterosynaptic but not homosynaptic plasticity. A) Initial stimulation causes LTP. B) The reactivation of the potentiated synapse alone (homosynaptic activation) does not necessarily trigger synaptic reconsolidation. C) Reactivation of several synapses, including the potentiated synapses (heterosynaptic activation), can trigger the depotentiation and repotentiation processes of synaptic reconsolidation. D) Behaviorally, only stimulus that cause heterosynaptic activation will induce reconsolidation.
4.2.2 Limitations of experiment 1

As mentioned in section 3.1.6, experiment 1 resulted in participants who received 2 reactivations (but not 4) having better memory on Day 3 compared to Day 1. This experiment likely suffered from ceiling levels of learning at Day 1, as can be observed in Figure 3.2, at least for groups R0, R1, and R2 (note the low mean number of errors and smaller error bars). For this reason, experiment 2 had half of the training trials that experiment 1 had. Additionally, group R4 (n=13) suffered from a low sample size, which undoubtedly affected the power to see significant differences in this group, i.e., it prevented us from detecting the expected memory strengthening effect.

Besides these parametrical limitations, other explanations for the lack of a strengthening effect in group R4 could be related to the nature of the task itself. The role of context in human episodic reconsolidation is still an open question. In the work by Hupbach and colleagues (Hupbach et al., 2007, 2011, 2008) that describes memory updating through reconsolidation by intrusions of a second set of learned items into recall of a first set of items (see Figure 1.1), context has been established as the most important factor for reconsolidation-mediated updating to occur (Hupbach et al., 2008). Only when original and new learning occurred in the same spatial context was reconsolidation-mediated updating observed. Further, just presenting the experimenter and/or the reminder question was not enough for reconsolidation to be triggered (Hupbach et al., 2008). In contrast, when Forcato and colleagues (2009) tested the different components of their reminder/reactivation trial, just presenting the context was not enough to trigger reconsolidation, whether it was to show disruption via
interference learning (Forcato et al., 2007), memory updating (Forcato, Rodríguez, Pedreira, & Maldonado, 2010), or memory strengthening (Forcato et al., 2011, 2013). Notably, the context in the Forcato et al. experiments is more complex and novel than the context in the Hupbach et al. experiments. Animal work suggests that context is particularly important to trigger reconsolidation-mediated updating. Winters and colleagues (Winters et al., 2009) found that early or weakly encoded object memory in rats could be easily destabilized by reactivation, but a context change was needed when the memory was old or strongly encoded. Nonetheless, in this case, this change in context could be considered a mismatch/prediction error, a boundary condition for reconsolidation that some argue is crucial in this process (Exton-McGuinness et al., 2015; Pedreira et al., 2004).

In the task used in the presented studies here, although the (very rich) context did not change in the reactivation trial, the mismatch component was clearly present in the type of reminder/reactivation trial used (see Forcato et al., 2009). Nonetheless, it could be argued that the assumption of prediction error in the current task is not what is necessary for reconsolidation to be triggered. In the current task, participants had to go through the same ‘error’ message in every reactivation trial. It is possible that participants in group R4 did not believe that the error is real by the end of day 2, i.e., after they had received all four reactivation trials. In fact, whereas 75% of group R2 reported not being suspicious of the ‘error’ on Day 2, only 54% of group R4 reported not being suspicious in a questionnaire regarding their actions during the experimental days at the end of Day 3, and only in group R2 did we observe the predicted strengthening.
effect in experiment 1. The awareness/belief of the prediction error is not something that has been systematically investigated, most likely because it is impossible to study in animal models and because only the Forcato group has incorporated it in their research. It is possible that the coupling/functional connectivity between the hippocampus, as a detector of prediction errors, and the PFC as a regulator of awareness of prediction error, might predict the efficiency of the reactivation trials, i.e., the magnitude of the strengthening effect. Nonetheless, neither fMRI study discussed above reported correlations between any sort of brain activation and performance on Day 3 (but note that neither study investigated memory strengthening, but memory disruption, via propranolol or via new learning). Although complete speculation, this idea might be worth pursuing for future endeavors.

4.3 Discussion of experiment 2

The effects of stress on memory have been extensively documented (Gagnon & Wagner, 2016; Het, Ramlow, & Wolf, 2005; Schwabe, Joëls, Roozendaal, Wolf, & Oitzl, 2012). Glucocorticoid receptor concentration is high in the rat (Mcewen et al., 1968; McEwen, Weiss, & Schwartz, 1969; Reul & Kloet, 1985) and human hippocampus (Seckl, Dickson, Yates, & Fink, 1991). In short, the combination of the fast-acting catecholamines (adrenaline, noradrenaline, NA) with the slower-acting glucocorticoids (cortisol in humans) in the amygdala as a result of a stressor shifts the brain into a “memory formation mode” (Schwabe et al., 2012) by mediating memory-relevant structures such as the hippocampus and the PFC. Moderate (but not low or high) levels
of cortisol seem to enhance encoding of new information (Diamond et al., 2007; Domes, Rothfischer, Reichwald, & Hautzinger, 2005; McCullough, Ritchey, Ranganath, & Yonelinas, 2015; Shields, Sazma, McCullough, & Yonelinas, 2017). Consequently, retrieval of old information is suppressed, as cognitive and synaptic resources are focused on encoding processes.

Recent evidence suggests that the timing of retrieval in reference to the stressor is also crucial. Schwabe and Wolf (2014) used a modified version of the cold pressor test that includes a socio-evaluative component (SECPT) before recognition testing and observed decreased performance only when the stressor was administered 25 or 90 min, but not immediately before, the retrieval test. Interestingly, salivary cortisol levels were only elevated at 25 min (not in the immediate condition) and back to baseline levels at 90 min. The later effect (90 min) was suggested to be an effect of late genomic cortisol actions, which are thought to be responsible for the beneficial effects of stress on consolidation; in this case, taking resources for consolidating the stressful experience and impairing retrieval of the learned material (Schwabe & Wolf, 2014). In another experiment, participants learned neutral and negative material (words and pictures) and 24 hr later were asked to recall half of the items in an oral-examination-like manner (which included a social evaluation component like the TSST and SECPT) or in a normal free recall test (Schönfeld, Ackermann, & Schwabe, 2014). The other half of the stimuli was recalled 25 min after with a normal free recall test. Strikingly, items that were recalled in a stressful manner were better remembered than those recalled in a standard free recall test. It is important to note that, at this time point, only blood
pressure was increased, but not salivary cortisol, which peaks 20-40 min following the stressor (Sapolsky, Romero, & Munck, 2000). Performance was positively correlated with the increase in blood pressure, showing that this effect is related to fast-acting autonomic processes (i.e., mediated by catecholamine release). Items (although this effect was only significant for negative stimuli) that were recalled 25min after the stressor were less remembered than those in the control condition. At this time point, blood pressure was back to baseline levels, whereas cortisol was significantly increased. Performance was negatively correlated with this cortisol increase. In short, retrieval during stress, in which NA levels are high but cortisol is not, seems to improve memory, whereas memory tested in the 20-30 min range post-stressor, when NA levels are low again but cortisol is high, seems to be impaired.

In the current study (experiment 2), memory was tested slightly outside of this ‘sweet spot’ (~25 min post-stressor) of high cortisol (minutes from end of stressor/control to beginning of testing session: M=18.75, SD=5.44, range 11-38). Although a significant increase in salivary cortisol was observed in our stressed participants (see Figure 3.12), it is possible that the lack of a (detrimental) main effect of stress in the current study was because cortisol levels were not increased enough by our stressor. In fact, the most consistent evidence showing a detrimental effect of stress/cortisol on memory retrieval has elicited extremely high salivary cortisol levels (see Gagnon & Wagner, 2016). Many of the most robust findings have used exogenous administration of cortisol (in tablet form, ~18 ng/mL) whereas results with natural stressors like the TSST are more inconsistent, eliciting cortisol levels 3-7.25 ng/mL. In
average, the highest levels in our stress groups were ~6 ng/mL, which is in range of previous literature. In short, the TSST may have not elicited a strong enough cortisol response to impair memory retrieval – at least in this task, which may have robust retrieval cues that undermine the effects of stress (see below).

It has been suggested that the detrimental effects of stress on retrieval of episodic memories also depends on the testing method used. Reports in which cued recall (De Quervain et al., 2003; Lupien et al., 1997) was used seem to show a stronger impairment from stress than those that used a recognition test, but not as strong as when free recall (Kuhlmann, Kirschbaum, & Wolf, 2005; Kuhlmann, Piel, et al., 2005; Kuhlmann & Wolf, 2005) was used. Retrieval tasks that provide few retrieval cues and/or require the generation of the learned material are more affected by stress than those that provide more retrieval cues (i.e., free recall>cued recall>recognition). In the current study, cued recall was used. More importantly, it could be argued that the rich context (specific unfamiliar tango song, specific colored light, and specific background picture) of the task provided additional strong retrieval cues. Specifically, context cues increase medial temporal lobe and PFC activation (Mather et al., 2006; Rasch, Büchel, Gais, & Born, 2007), so the rich context of the current task may have overruled the increased retrieval threshold caused by the stressor. This idea is supported by an elegant study by Schwabe and Wolf (Schwabe & Wolf, 2009b), in which stress-related performance decreases in an object-location task were only observed when the memory was tested in another context (room B without odor), different from where the material was initially learned the day before (room A with odor). If memory was tested in the
same context, no stress-at-retrieval effect was observed. Again, considering the richness of the context used in the current task, and the fact that participants were in the same room and were presented with the same context all three days, it is possible that retrieval cues were stronger than the stressor used. In short, consideration of the context, both at the moment of reactivation (see discussion above) and at the moment of retrieval, should be warranted in future research.

4.3.1 Limitations of experiment 2

The most obvious limitation of experiment 2 is the small sample size (7 or 8 per group). Decent sample sizes are particularly important for stress studies because it has been shown that the TSST affects males and females differently (Kirschbaum, Wüst, & Hellhammer, 1992; Kudielka & Kirschbaum, 2005) and that the effects of stress in memory also vary with sex (e.g., Buchanan & Tranel, 2008; Wolf, Schommer, Hellhammer, McEwen, & Kirschbaum, 2001). Menstrual cycle also affects reactivity to a stressor (Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999; Stephens, Mahon, McCaul, & Wand, 2016) and its relationship with memory (e.g., Andreano, Arjomandi, & Cahill, 2008; Andreano & Cahill, 2006). Finally, large sample sizes are also needed because inter-individual variability in cortisol samples tends to be substantial, as 20-30% of individuals do not mount a cortisol response to psychosocial stressors (Kajantie & Phillips, 2006; Kudielka & Kirschbaum, 2005; R. Miller, Plessow, Kirschbaum, & Stalder, 2013; Otte et al., 2005).
According to our salivary cortisol analyses and the ratings from the Self-Assessment Manikin (SAM) scale our stressor was effective in increasing arousal, negative feelings, and cortisol levels in the stress groups. Nonetheless, it is possible that our stressor was not strong enough, in that it did not produce a robust stress response (increased activity in sympathetic nervous and HPA systems) and that the self-report ratings were not real. That is, since participants were told in what condition they were immediately before the start of the stress/control manipulation, it is possible that participants completed the SAM to adhere to the expectation of each condition. Additionally, the SAM was administered on a single sheet of paper, in which participants filled out one side pre-manipulation and the other side post-manipulation. This method allowed them to go back to their pre-manipulation ratings and answer according to what they thought were the experimenter’s expectations. Future studies should 1) reiterate to the participants that they should answer truthfully on how they feel at that exact moment and 2) make sure that participants cannot access their pre-manipulation response. Finally, saliva samples are necessary to objectively measure stress reactivity (at least HPA activity). The addition of heart rate monitoring would allow for an objective measure of sympathetic nervous system activity.

4.4 Conclusion

The current project aimed to investigate one of the proposed functions of memory reconsolidation in humans: memory strengthening. The field of memory reconsolidation is still very young; some may argue that it (re)started only 17 years ago
with Nader at el. (2000). Work in declarative memory reconsolidation is even younger and scarcer, with the first evidence reported ten years ago with Forcato et al. (2007) and Hupbach et al. (2007). To the best of our knowledge, research investigating the strengthening function of reconsolidation in humans is even more limited, and virtually specific to one research group (with the exception of Wichert, Wolf, and Schwabe (2013a), who showed null results) in Argentina (Forcato et al., 2011, 2013, 2014).

Therefore, the current research expanded significantly on the declarative memory field by attempting to replicate the strengthening effect in humans and testing its efficacy against stress at retrieval. The current project (partially) replicated the strengthening effect by showing that two reactivations in Day 2 result in a stronger memory in Day 3, but, most likely due to lack of power, was not able to fully investigate the effect of stress-induced retrieval impairment on the strengthened memory.

It is clear that further research is warranted in order to 1) fully understand the role and mechanisms of reconsolidation in humans and 2) fully understand its strengthening function so that we can create evidence-based interventions that are rooted in reconsolidation research. Memory strengthening is an infinitely useful mechanism for individuals with developmental and learning impairments as well as individuals with neurodegenerative disorders that destroy our memory systems, such as Alzheimer’s disease.


Fernández, R. S., Bavassi, L., Forcato, C., & Pedreira, M. E. (2016). The dynamic nature of the reconsolidation process and its boundary conditions: Evidence based on
human tests. *Neurobiology of Learning and Memory*, 130, 202–212. https://doi.org/10.1016/j.nlm.2016.03.001


McClelland, J. L., McNaughton, B. L., & O’Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory.


Aversive Stimuli. *PLOS ONE*, 7(6), e38849. https://doi.org/10.1371/journal.pone.0038849


124


