McMaster University

Department of Psychology, Neuroscience & Behaviour

Daphne Maurer, Ph.D., F.R.S.C.
Distinguished University Professor

1 August 2021

Misconceptions of Memory and Trauma

How memory works. Memories are not what most people believe them to be. They are not stored in neuronal pigeonholes, waiting stably to be retrieved. They are neurochemical connections that are too unstable not to be changed by the neurochemical reactions forming recollections. For this reason, memories constantly change. On page 3 you will find an explanation of this for general readers, eight pages from one of my books.

Accurate memories. Memories constantly change. A lot of experimental evidence shows this. Events accompanied by a modest amount of emotion do tend to be remembered more accurately, but a large set of studies show that *traumatic* events tend to be remembered *less* accurately. These studies tested soldiers traumatized under standardized conditions in a military survival school. Page 12 is a three-page review of this work by two professors at the (US) National Center for PTSD, based at Yale.

Consistent memories. Consistency and accuracy are different dimensions. It is easy to be consistently wrong.

A review of 37 studies (page 15) found that in some of them, traumatic memories tended to become a little more consistent with time, but the difference was slight and not of practical significance. The authors conclude:

> Results of this review have implications for legal practice. Memory is a reconstructive process, which is prone to errors. Therefore, we cannot fully rely on its accuracy, completeness, and consistency. Eyewitness testimonies completed in situations where a victim is showing emotional reactions may not be entirely reliable. Testimonies completed at an early stage following a crime may be incomplete and inaccurate, in particular with regard to important details of an event. However, reports of criminal events can be highly consistent over time, without necessarily being accurate. Therefore, one has to be cautious using consistency as an indicator of accuracy and drawing conclusions from a single testimony.

Recovered memories. There is a great deal of evidence that interviewers can and do induce the "recovery" of erroneous memories. An overview of this literature starts at page 34. One of its authors is Elizabeth Loftus, Distinguished Professor at the University of California, Irvine. Loftus's research on memory has earned her 59 awards from around the world, plus seven honorary degrees.

Trauma-informed investigation. When this term means sensitive listening immediately after a trauma, the technique is sensible, but it commonly means asking leading questions long after the fact, sometimes to "recover" memories. This kind of "trauma-informed" investigation is inappropriate. Elizabeth Loftus (see above) co-authored a recent paper explaining the problems. It starts at page 65.

Trigger warnings. Trigger warnings tell people that they may encounter something they find upsetting. They are supposed to reduce anxiety but they do the opposite, because they work like a laboratory bell warning a mouse that it may be shocked. In this situation, the bell comes to frighten the mouse whether it is shocked or not. This is basic Pavlovian conditioning. Similarly, trigger warnings usually become more upsetting than the triggers themselves.

The top of page 73 shows this. It is a table by researchers at Harvard reporting results from the six studies of trigger warnings that the authors had been able to find as of 2020. One of the six did show a reduction in anxiety afterwards, but another one found an *increase,* and both of those effects were negligible. On the other hand, the three studies that measured it found the warning itself to trigger anxiety.

Moreover, trigger warnings are almost never heeded. An example of this is on the bottom of page 73.

In short, trigger warnings do harm rather than good.

False charges. Supposedly, no woman would lie about rape because no woman would subject herself to the inquisition that follows. However, a meta-analysis (page 74) based on police records found convincing evidence that across four anglophone countries, roughly 5% of rape reports were false. Apparently some women, like some men, want to take revenge upon somebody, and/or enjoy being the centre of attention, and/or develop psychotic distortions of reality.

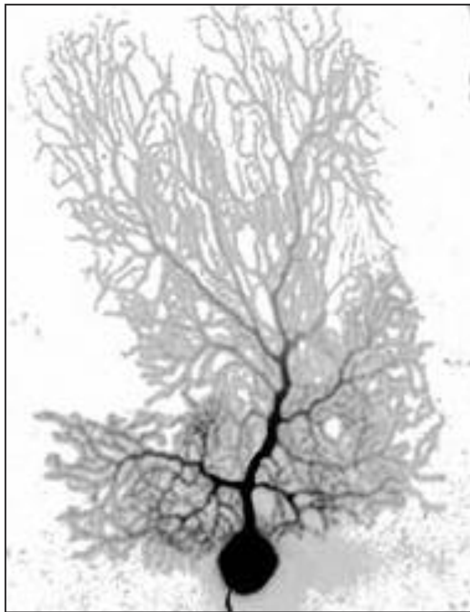
Five percent may sound like a low error rate, but in this context it is high. Imagine that after convicting 19 criminals, a judge orders a cop to pull a man off the street at random and throw him into jail. That is the equivalent of what happens when 5% of rape reports are false.

Moreover, 5% is a conservative estimate. Many of the less clear cases cases were undoubtedly false as well, so the real rate of false reports will be higher. Note, too, that less serious sexual assaults are easier to fabricate and less onerous to report, so falsifications of lesser claims are likely to be more common.

From Maurer, C., & Maurer, D. (2019). Pretty Ugly: Why we like some songs, faces, foods, plays, pictures, poems, etc., and dislike others. Newcastle upon Tyne: Cambridge Scholars Publishing. <prettyugly.info>

The Nervous System

In chapter two we showed that single-celled organisms behave tropically, like bouncing balloons. So do individual cells behave within the body of a multicellular organism like man—all individual cells, including neurons. Nerve cells come in an infinity of shapes but their primary tropic response is always the same. One end of the cell is extremely delicate. When something bumps this end, either physically or chemically, the impact shakes up subatomic particles. If the impact is sufficient, some of them push through the cell's membrane. The particles hold an electrical charge—they are ions—so their passage represents a minute discharge of electricity through the membrane. This discharge stimulates neighbouring ions, so that those ions push through the membrane as well. A chain reaction begins that leads to and across the main body of the cell, then off again along a fibrous projection. The reaction runs along that projection then jumps to a neighbouring cell and "fires" it. A short while later the cell's chemical machinery pulls the ions back in, leaving the cell ready to fire again.



All neurons work that way, from your head to your toe. Some neurons develop shapes that make them especially sensitive to energy in a specific form, but the sensitive end of most neurons is similar, a delicate tree of fine fibres called dendrites, after the Greek word for tree. In this photomicrograph you can see a dendritic tree extending from the body of a cell. In this image the second projection—the exit—is cut off at the bottom. It is called an axon, after the Greek for axle. Although this truncated axon may resemble an

axle here and in simplified drawings, in reality it looks more like a devil's tail with branching ends.[1]

Most of our neurons are surrounded by other neurons, either within strings of tissue called nerves or *en masse* within the brain. Neurons' dendrites and axons are thoroughly intermixed. They do not quite touch one another but by firing they form chemical connections that are more or less durable. This is another tropic process. Neurons sit in a chemical soup. When a neuron fires, the energy leaving its axon may pass through the soup to stimulate a nearby dendrite. At first the energy will cross to any dendrite that happens to be close enough, but as it passes through the soup, it alters the ingredients. Passing ions modify the local chemistry much as the energy of sparks in air converts oxygen into smelly ozone. Those chemical changes do two things: they facilitate the passage of more ions along the same route, and they inhibit the passage of ions nearby. The result is an *ad hoc* neurochemical connection that becomes stronger each time it is used. This connection is called a synapse, from the Greek for "connection."

Nerves to and from the brain end inside the eyes, ears, nose, mouth, skin, muscle, guts, and glands. We have some specialized nerve endings inside the eyes, ears, nose and mouth, but most nerve endings are generic and they are everywhere in the body except inside the brain. Most nerves respond to stimulation from our environment and body, then send it inward and upward to the spinal cord and brain. Others are oriented in the other direction. They receive stimulation from the brain and spinal cord, and then carry it to muscles, glands, and other tissues that may contract or release some chemical when stimulated.

All nerves lead to and from the brain, so the brain appears to control the nervous system, but the brain can no more control the entire nervous system than a general can control every part of an army. Small units will respond to firing automatically and independently. When you touch a hot stove, the nerves in your hand induce tropic responses that move your hand long before any neuronal firing reaches the brain. This is a neuronal company of infantry returning fire, using the same sort of dendritic connections as are in the brain.

Moreover, unlike a general, the brain does not effect control by issuing commands. The brain is a reactive structure that guides the

body's functioning not actively on its own volition but passively, deforming under the pressure of traffic. The brain acts like a blanket of snow covering streets. As a driver forces his car down a street, his tires form tracks. Since it is easier to drive in tracks than to push through fresh snow, drivers following him tend to follow in his tracks. Traffic deforms the snow and those deformations influence traffic in turn. The snow itself is passive. Within the brain, chemical responses to neuronal traffic form channels of synapses, and succeeding neuronal traffic tends to follow those synaptic channels. The brain ends up controlling neuronal traffic passively in that way.

This is not how psychologists usually think of the process. Our models usually contain feedback loops and forms of active control. However, those are engineering models looking at neuronal functioning from the outside. The brain's internal feedback and control mechanisms are not the sort that an engineer would envision. Firing neurons release chemicals, which dissipate in the surrounding space and, according to their concentration, increase or decrease the odds that energy will cross that space to fire nearby neurons. This is the kind of control that a tide exerts on fish in a brackish estuary, sending fresh-water fish swimming in one direction and salt-water fish in the other. To a fisherman the tide seems to control the fish, but the tide is not intending to do this and the fish are merely reacting to a chemical gradient, swimming toward or away from salt. Fundamentally, the system is reactive. Neurons do indeed fire on their own as well, and they do so continually, but spontaneous firing is disorganized energy: entropy. Spontaneous firing channels energy through existing pathways so that the consequences are organized and often look intentional, but entropic energy is random, not purposeful. If a tune runs through your mind's ear while you are frying eggs, this is because some random firing kicked off a channel formed previously by listening to that song.[2]

This presents a paradox. A reactive brain is a tropic mechanism yet we are hardly tropic creatures, and the brain is central to everything intelligent and adaptive that we do. It seems impossible that a reactive brain dealing with entropic activity could form an Aristotle.

To resolve this paradox we need to see the brain's intelligence as the sort of emergent phenomenon we described in chapter one. We shall begin by considering a theoretical device that we use every day: a

von Neumann machine. This is a theoretical machine that does three things: 1) reads numbers from some storage device, 2) adds or subtracts those numbers mechanically or electronically, then 3) returns the result to the storage device. It is the prototypical computer.

The computer on your desk is a von Neumann machine in the flesh (or rather, in the silicon). In principle it does nothing more than execute von Neumann's basic sequence, but it executes the sequence hundreds of millions of times per second. This simple mechanism permits a computer to do astonishingly complicated things, even to adapt its computations to circumstances and to learn. A computer may or may not have intelligence but the answer is sufficiently uncertain that academic careers are based on arguing the question. Much clearer is the cause of the uncertainty: the computations' inconceivably large scale. The apparent intelligence of a computer emerges from this scale.[3]

The brain's principle of operation is even simpler than a von Neumann machine's. The brain's functional principle is merely this: 1) local chemistry may permit energy from one neuron to transfer to an adjacent neuron, and 2) any such transfer of energy modifies the local chemistry to alter the efficiency of transfer in the future. To do a job of work, the brain requires scaling this principle to an unimaginable level of complexity: 10 billion neurons admitting 100 trillion possible connections.[4]

To get a sense of 100 trillion take a period from this book, abut it to another period, then repeat this 100 trillion times. Before you have finished, you will have circuited the earth 1000 times.[5]

But these do not define all of the possibilities, not by a long shot. A sequence of neurochemical reactions forms a wave across time. The possible connections of a neurochemical wave vary with its phase and frequency, so the possible connections are tenfold greater, on the order of 1000 trillion. Moreover, since a neurochemical reaction in one spot will facilitate and inhibit reactions nearby, a wave of reactions is likely to help shape a succeeding wave of reactions. This does more than make the system unimaginably complicated, it also makes the system unpredictably complex—deterministically chaotic in the general sense we developed in chapter two. Moreover, entropic neuronal firing continually alters the deterministic course.

Systems like this are difficult to comprehend in the abstract but if you own a home in the north, you understand one of them clearly. In the fall if you forget to drain unheated water pipes, the water will gradually chill and then will suddenly become ice. Emergent phenomena will include burst pipes, a flooded basement, and working overtime to cover the costs.

Scientists can follow little of the brain's complexity but with computers they have created virtual neural networks, computer programmes that model many complex and chaotic behaviours much as they are seen in the lab. These models involve computations rather than chemicals, so they say nothing about the neurochemical processing of the brain, but they do show that naturalistic, complex, *adaptive* behaviours can grow from the automatic responses of a simple, reactive machine. A computerized neural network has nowhere near enough power to write a philosophical treatise on the nature of beauty but it can categorize a set of faces by attractiveness much as humans do.[6]

The number of theoretical connections in the brain is so large as to seem infinite, yet at any given moment, the actual number of possible connections is limited, defined by the local chemistry. Any particular passage from axon to dendrite depends upon whether the chemicals around that axon are absorbing energy or passing it on. This depends in turn upon what energy passed nearby shortly before, and what the chemistry was like when the last energy passed by.

Connections begin with the first innervation within the embryo, they are shaped by the chemical and physical structure of the body, they are continually reshaped as the body develops, and as we shall see in chapter seven, they are influenced by the intrauterine environment: what a mother eats while she is pregnant can affect what her baby will enjoy.

Since human anatomy is the same everywhere, many infant behaviours are universal and appear to be wired into the brain, but most of what appears to be wired in is also influenced by the environment. Take the ability to learn language, for instance. After six months of gestation the ears mature enough to generate neurochemical responses to pressure waves. This means that during the last three months of gestation, pressure waves within the

mother's body begin to form neuronal pathways inside the fetus's brain where nerves from the ears enter it. Most of these pressure waves come from the pumps and plumbing within the mother's abdomen, but these are overlain forcefully by the pressure waves of the mother's speech. The latter are so frequent and pronounced that rhythmical constancies become the dominant force in shaping the pathways in those parts of the fetal brain. These neuronal structures become so attuned to the mother's idiosyncratic rhythms that after a baby is born, he will suck on a pacifier to choose her voice played through a loudspeaker instead of a stranger's. This attunement to her voice makes it easier for him to recognize specific sounds that she makes repeatedly. Once he recognizes those sounds, he comes to associate them with other sensations—with particular feelings and events—and eventually to treat them symbolically, as words. He also comes to recognize that these words usually appear in a certain order. Thus he begins to learn the syntax of his mother's language.[7]

Virtually all of these associations form within the cortex of the brain, the grey matter that forms the mammalian brain's outer layer. These cortical associations are the subject of this book. For ease of discussion, we shall commonly use phrases like "cortical processing" or "cortical control," but it is essential to keep in mind that this processing and control is fundamentally reactive, no less reactive than any other part of the brain. The cortex is merely the part of the brain that has the greatest density of neurons and thus permits the greatest variety of neuronal routes.

Memory and Dreams

Wilder Penfield was a neurosurgeon so audacious that he once operated on the brain of his own sister. As Penfield operated on a brain, he systematically explored it, stimulating it here and there and everywhere, to map what different parts of the brain do. (Brain surgery is done under local anaesthetic so that the patient's responses can guide the surgeon. This is possible because without nerve endings, the brain can sense nothing directly. Anaesthetic is required only to cut through the skin and skull.) Penfield found that stimulating any portion of the brain is likely to bring up a memory, so he concluded that the brain is a neural cabinet with chemical

pigeon holes for memories. Some of those memories had long been forgotten, so he hypothesized that everything we ever perceive is stored somewhere in the brain.[8]

Penfield's observations dovetailed nicely with notions of memory that are century-old furnishings of western intellectual life, notions of repression, suppression and recovery of memories, and "flashbulb" memories impressed indelibly by a traumatic experience. These notions of memory have become rooted in our beliefs and culture as deeply as geocentricity was rooted in medieval Europe. Unfortunately, a vast amount of rigorous, scientific evidence has shown these notions to be comparably wrong.[9]

You can see that this notion of memory is wrong if you just think about a memory from your childhood. Remember a family dinner. Visualize this in your mind's eye. Do you see yourself sitting at the table with everybody else around? If so, then it is obvious that you are not remembering what you saw. When you sat at that table you saw the people sitting around you but you did not see yourself. You have never seen yourself anywhere, save in a mirror. If ever you have an image of yourself anywhere doing anything, then you are not remembering that image, you are creating it.

A model of memory better fitting the scientific data comes from our model of the brain as neurochemical tire tracks through snow. These neuronal tracks are formed by experience, but they are not memories in and of themselves any more than strengthened fibres in biceps are memories of barbells. In various ways at various levels, experience helps to form every tissue in the body, but we would not say that the shape of any particular tissue *per se* is a memory. A memory is not neuronal tracks, it is a bolus of chemical energy rolling along neuronal tracks. This energy may come from sensory stimulation, or it may be internal neuronal energy forming thoughts, or it may be neuronal energy that is merely rattling around in a random way forming dreams, but a bolus of energy is required. You will never forget how to ride a bicycle but neither will you remember until you climb into the saddle. You cannot remember how to balance while you are sitting in a chair.

Since no single sensory stimulus is likely to be identical in every way to any other, and since no one state of the brain and body can ever

be identical to one that came before, each bolus of sensory energy slips and slides in the tire tracks a little differently. This means that each recreated memory is likely to be different from the last. Moreover, each passage of energy alters the tracks. The neuronal tracks deepen, so energy is more likely to find its way along them, but they also widen, shift, and distort. The brain has no way to know whether any particular chemical track is the same now as it used to be, or how much it changed, so we have no way to know how accurate our memories are. We may feel absolutely certain that we remember something accurately yet there is every likelihood we do not.

We can see this from research following the attack on New York's World Trade Center. Almost every American remembers what he was doing when he first heard about the attack. Most people believe that so fearful and dramatic an event sears consciousness, burning its image into the brain overtop the ordinary events of the day. To test this belief, on the day after the attack, Jennifer Talarico and David Rubin interviewed students at Duke University about what they were doing when it happened. They also asked each student about something else that had happened within the previous three days. They did this using carefully structured interviews. Either six weeks or 32 weeks later they interviewed each student again about both events. They used the same structure and gave all of the transcripts to third parties to check for inconsistencies and contradictions. The students remembered the 11th of September much more vividly than the other event, and much more confidently, but their memories of both were equally flawed.[10]

In short, a memory is not a fixture of the brain, it is a replayed perception, a perception that is divorced from sensory input (although a sensation may have triggered it) and is played back on a neurochemical machine that may start from any position and that frequently skips and distorts. The only difference between "real" waking memories and the bizarre memories of a dream lies with entropy. We dream while the brain has lost some of its power to organize energy because its chemical stores need replenishment. During sleep some neurochemical energy still enters the brain from sensory systems but it sloshes about largely at random, stimulating neuronal channels hither and yon. This tends to destroy tenuous neuronal connections, so that we forget much of what we did during the day (although it can also reinforce some recent connections,

perhaps to solidify the French vocabulary you reviewed just before switching off the light). When this stimulation passes through parts of the brain that form consciousness, we experience a dream.[11]

1. Adapted from a photomicrograph by Maryann Martone, Eric Bushong, and Ed Esquenazi. Released under the Creative Commons.

2. An excellent introduction to neuroscience is published on-line (in many languages) by the British Neuroscience Association and the European Dana Alliance for the Brain: *Neuroscience—Science of the Brain: An Introduction for Young Students* <https:/ /www.bna.org.uk/static/uploads/resources/BNA _English.pdf> (Accessed 11 February 2019)

3. von Neumann, John, "The First Draft Report on the EDVAC [30 June 1945]." *IEEE Annals of the History of Computing* 15, no. 4 (1993): 27-75.

4. These are orders of magnitude, not precise numbers. The first of these references has 10 billion neurons and 60 trillion synapses; the second has 20 billion neurons and 240 trillion synapses. G. M. Shepherd, "Introduction to Synaptic Circuits" in Shepherd, G. M. (ed.), *The Synaptic Organization of the Brain* (Oxford: Oxford University Press, 2004): 7. Christof Koch, *Biophysics of Computation: Information Processing in Single Neurons* (Oxford: Oxford University Press, 1999): 87.

5. Assuming the circumference of the earth to be 40,000 km and the diameter of a period to be 0.4 mm.

6. Wen-Chung Chiang, Hsiu-Hsia Lin, Chiung-Shing Huang, Lun-Jou Lo, and Shu-Yen Wan, "The Cluster Assessment of Facial Attractiveness Using Fuzzy Neural Network Classifier Based on 3D Moiré Features." *Pattern Recognition* 47, no. 3 (2014): 1249-1260. A. J. Kell and J. H. McDermott, "Deep Neural Network Models of Sensory Systems: Windows onto the Role of Task Constraints." *Current Opinion in Neurobiology* 55 (2019): 121-132.

7. For an introduction to neonatal hearing see Chapter 7 of Daphne Maurer and Charles Maurer, *The World of the Newborn* (New York: Basic Books, 1988.)
8. Wilder Penfield and Lamar Roberts, *Speech and Brain Mechanisms* (Princeton: Princeton University Press, 1959).

9. For a summary of modern views of memory see L. R. Squire, "Memory and Brain Systems: 1969-2009." *Journal of Neuroscience* 29, no. 41 (2009): 12711-12716.

10. J. M. Talarico and D. C. Rubin, "Confidence, Not Consistency, Characterizes Flashbulb Memories. *Psychological Science* 14, no. 5 (2003): 455-61.

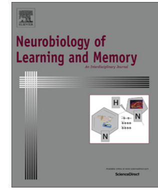
11. On the relationship between cortical firing and dreams: B. Baird, A. Castelnovo, O. Gosseries, and G. Tononi, "Frequent Lucid Dreaming Associated with Increased Functional Connectivity between Frontopolar Cortex and Temporoparietal Association Areas." *Scientific Reports* 8, no. 1 (2018): Article number 17798. On the incorporation of personal experiences: C. L. Horton and J. E. Malinowski, "Autobiographical Memory and Hyperassociativity in the Dreaming Brain: Implications for Memory Consolidation in Sleep." *Frontiers in Psychology* 6 (2015): 874.

Neurobiology of Learning and Memory 112 (2014) 101–103

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme



Review

Perspective: I believe what I remember, but it may not be true

Charles A. Morgan, Steven Southwick *

National Center for PTSD, Yale University, School of Medicine, United States

ARTICLE INFO

Article history:
Received 9 August 2013
Revised 18 December 2013
Accepted 20 December 2013
Available online 4 January 2014

Keywords:
Eyewitness memory
Stress
Forensics
Misinformation
False memory
Military

ABSTRACT

A growing number of research findings have challenged the conception that memory for traumatic events is highly accurate or even indelible in nature. Research involving soldiers indicates that realistic levels of high stress decrease the accuracy of eyewitness memory. In addition, recent findings from several studies show quite clearly that memories for stressful events – including those from combat trauma – are malleable and vulnerable to alteration by exposure to misinformation. Under high stress, our brains facilitate the formation of "gist" memories that allow us to avoid future dangers but which may not contain the detail and precision demanded by the judicial system. Although mental health professionals ought to play a role in educating the courts about mental illness and trauma, it is unwise for them to become advocates for the idea that traumatic memories are indelible, factual accounts of events.

© 2014 Published by Elsevier Inc.

1. I believe what I remember, but it may not be true

> *"I can't stop remembering. I see it all so clearly. Like it happened just yesterday. But it was 5 years ago. It was bad enough that I was raped in the first place but to have to relive it over and over is just too much."* – Trauma survivor with PTSD

In the summer of 1998, at the International War Crimes Tribunal in The Hague, a Bosnian-Croatian solider was tried for torture, inhumane treatment, and abetting the rape of a Muslim woman. The woman (Witness A), who was raped, was the primary witness in the trial; she had been diagnosed with post-traumatic stress disorder (PTSD) as a result of her experiences and was, in the court's eyes, *the person* who was best qualified to identify the person responsible for the crimes. However, there was a problem: Prior to appearing in court, the woman provided numerous and different sworn statements about the appearance and actions of the person responsible for her trauma. With respect to the appearance of the accused, shortly after the attack in 1995, Witness A described the defendant as being 5 ft 7 in. tall and as having short blond hair and blue eyes. However, at the trial in 1998, medical records showed the accused to be over 6 ft tall with dark hair and brown eyes. At the trial, Witness A performed an *in court* identification and, in the presence of the accused/defendant and while looking at him, described him as a, "rather thin young man [who had put on some weight], rather strong jaw or teeth" who had, "chestnut to black hair", which was "cut short and combed up".

What happened between 1995 and 1998 to witness A that made her give such a different account of events? How could the memory of her abuser's hair, height and build for which she described clear and distressing memories, change so dramatically over a period of three years?

The accuracy of memory for both neutral and emotionally arousing events has been a topic of great interest to animal and human researchers for years. An enormous body of scientific evidence has shown that emotionally arousing events are generally remembered better than emotionally neutral events. As early as 1890 William James observed "An impression may be so exciting emotionally as almost to leave a scar upon cerebral tissue." (James, 1890). From a neurobiological perspective, pre-clinical and clinical research has demonstrated that emotionally arousing experiences stimulate the release of endogenous stress hormones/neurotransmitters (e.g. corticosterone, glucose, epinephrine, norepinephrine), that these stress hormones can enhance consolidation of memory, and that they do it largely through effects on norepinephrine in the amygdala (McIntyre et al., 2012).

However, a growing number of research findings have challenged prior conceptions about the "indelible" nature of traumatic memories. For example, among combat veterans, a number of studies have shown that memories for traumatic events, experienced while in a war zone, can be inconsistent over time (Engelhard, van den Hout, & McNally, 2008; Krinsley, Gallagher, Weathers, Kutter, & Kaloupek, 2003; Roemer, Litz, Orsillo, Ehlich, & Friedman, 1998; Southwick, Morgan, Nicolaou, & Charney, 1997). In each of these research studies, the inconsistencies in memory were not limited to trivial events, but instead, included non-trivial

* Corresponding author.
E-mail address: Steven.Southwick@yale.edu (S. Southwick).

1074-7427/$ - see front matter © 2014 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.nlm.2013.12.011

events such as ‘‘being shot’’, ‘‘being shot at’’, ‘‘witnessing death of a friend’’, ‘‘viewing human remains’’, ‘‘being physically injured’’, and exposure to ‘‘firefights’’. In addition, inconsistencies in memories for traumatic events do not appear to be rare but are quite common and noted in 60–80% of the populations studied. It is important to note that inconsistency may not be the same as inaccuracy.

For many researchers and clinicians, (including us), the above noted data were difficult to reconcile with what traumatized patients report. We wondered whether the observed inconsistencies noted in military personnel might be caused by the imprecision inherent in the assessment methods used by researchers (i.e. the questionnaires), or that they might be caused by recall bias or by cognitive re-framing of events on the part of participants. Given that none of the research had ‘‘ground truth’’ evidence about the actual events, we reasoned that memories for the genuine events might be genuine but just not accurately assessed.

Over the past decade we have conducted a number of studies designed to address the issue of accuracy and indelibility of human memory for highly stressful events. To do this we have studied memory in active duty military personnel engaged in a highly stressful form of training where it was possible to objectively document the characteristics of a highly stressful event and compare these data to that of memory recall on the part of those exposed to these events.

Research data from our psychobiological investigations of military personnel who have volunteered for survival school training has provided robust evidence that it is a valid and reliable model for the study of realistic, threat-to-life stress in humans (Eid & Morgan, 2006; Morgan, Doran, Steffian, Hazlett, & Southwick, 2006; Morgan et al., 2000, Morgan, Wang, et al., 2001; Morgan et al., 2007). Unlike the stress experienced in mock crime or field studies, the stress experienced by survival school participants is applied in a uniform manner and, from a scientific perspective, controls for intensity and duration of stress across subjects. The extreme stress experienced by participants is pronounced and results in significant alterations of neurobiological processes (e.g. elevated cortisol) and psychological experiences (e.g., peritraumatic dissociation) that are on a par with those noted in individuals experiencing real-world, threat-to-life events such as landing on an aircraft carrier at night for the first time, skydiving for the first time, open-heart surgery, and actual combat (Morgan, Hazlett, Dial Ward., & Southwick, 2008; Morgan, Hazlett, et al., 2001). As such, the survival school provides a unique scientific opportunity for prospective, objective evaluations about the accuracy of human memory for events experienced during exposure to realistic stress. Since the stress of Survival School is a core feature of the school curriculum – and not specifically designed for research purposes – the opportunity to evaluate memories of soldiers in such training provides a way to study the impact of realistic stress on human memory that would not be otherwise ethically possible in traditional laboratories.

In order to evaluate the nature of human memory for realistic, high stress events (Morgan et al., 2004), we tested the eyewitness accuracy of soldiers who were exposed to interrogation stress. We assessed eyewitness memory in 500 soldiers to see if they could identify the individuals who conducted their high stress (during which the average heart rate was between 150 and 170 bpm) and low stress (during which the average heart rate was between 110 and 125 bpm) interrogations. Contrary to expectations, we found that regardless of the method used to assess memory (i.e. live line up, simultaneous or sequential presentation) eyewitness recall was less accurate for the high, compared to low stress interrogations.

At first glance these data may seem counterintuitive. If memory is generally better for emotional events compared to neutral events, why were soldiers less accurate in identifying interrogators who conducted their high stress interviews compared to interviewers who conducted the moderately stressful interviews?

While it is true that memory for emotional events is enhanced by the endogenous release of stress hormones/neurotransmitters, the relationship between memory and activation of stress hormones/neurotransmitters commonly fits a non-linear inverted-U shaped curve. Summarizing an extensive pre-clinical literature, Morley and Farr note, ‘‘The inverted U-shaped dose–response curve appears to be almost universal for memory mimetics.’’ (Morley & Farr, 2012). Studies testing memory enhancing or degrading effects of epinephrine, norepinephrine, neuropeptide Y, and steroid hormones, among others, as well as drugs that modulate a number of these stress-related hormones/neurotransmitters, have typically reported inverted U-shaped curves. Thus, retention for emotional/stressful events is generally best when stress and accompanying activation of stress-related hormones/neurotransmitters is moderate rather than minimal or extreme in nature. (Morley & Farr, 2012).

It is not known why high doses of stress tend to compromise retention. Based on preclinical findings, it is possible that high levels of stress impair hippocampal function by blocking induction of LTP, or that high levels of stress impair prefrontal cortical capacity to attend to the stressful event and/or to facilitate memory formation through interactions with the temporal lobe (Diamond, Campbell, Park, Halonen, & Zoladz, 2007). It is also possible that high levels of stress hormones/neurotransmitters engage a homeostatic response, such as the activation of opiate mechanisms, that down-regulate memory formation (Gold & Korol, 2012).

In addition to the findings noted above, we found that if soldiers were shown photographs of their interrogators that were taken *at the time of the interrogations*, eyewitness accuracy for high and low stress events was 66% and 86%, respectively. These data provided evidence that high stress impairs accuracy of eyewitness memory *and* that people differ in their vulnerability to this negative impact of stress.

In our follow up studies of eyewitness memory in military personnel at survival school (Morgan, Southwick, Steffian, Hazlett, & Loftus, 2013; Morgan et al., 2007) we replicated and extended our findings: With respect to eyewitness accuracy for body and facial features traditionally used by police and courts for identification purposes (i.e., race, gender, body size, hair length, hair color, shape of face, eye color, shape of ears, facial hair, teeth) the only features for which eyewitness accuracy was greater than 50% were for the interrogator's hair length, race, gender, teeth and approximate size and body shape. The majority of subjects were wrong with respect to a person's hair color, shape of face, facial hair, eye color, shape and appearance of ears.

In addition to high rates of inaccuracy in face identification, we also found that eyewitness memory for stressful events was malleable: In our study involving 861 soldiers we found that those who were exposed to misinformation were more likely to commit errors in eyewitness memory (Morgan et al., 2013). Compared to controls, soldiers who were exposed to misinformation in the form of a photograph (i.e. a picture of someone different from their interrogator) immediately after their interrogation were 40% more likely, when given a memory test 24 h later, to misidentify the man from the photograph as the person who conducted their interrogation.

Further, when exposed to misinformation contained in a questionnaire, 27% of soldiers falsely identified their interrogator as having brandished a weapon; 51% of soldiers who were exposed to misinformation in a video falsely recalled seeing caches of rocket propelled grenades. On average, soldiers who had formed false memories, tended to have higher confidence in the accuracy of their memories compared to soldiers who did not have false memories.

This finding – that memory for highly stressful events can be altered by misinformation – does not appear to be an artifact of the survival school setting. In a recent study, Lommen, Engelhard, and van den Hout (2013) found that exposure to misinformation significantly altered memory for combat related events in 26% of 249 soldiers who had just returned from their tour of duty in Afghanistan.

Taken together the research findings make it clear that we need to reformulate our previous assumptions about voluntary human memory (Brewin et al., 2012) for traumatic events. Even though we may strongly believe in the accuracy of our memories for emotionally salient experiences – what we believe may not, in fact, be true. For while we remain confident that emotional events are remembered better than neutral events, there is a caveat – memory for novel and highly stressful events is vulnerable to substantial error. Under high stress, it seems that our brains facilitate the formation of "gist" memories that allow us to avoid future dangers. However, for a number of reasons, memories formed under high stress may not contain the detail and precision demanded by the judicial system: First, at the time of encoding, high levels of stress, arousal and norepinephrine, may exceed the "sweet spot" that is optimal for the formation of accurate memories; Second, following the initial encoding, stress related memories may be susceptible to alteration through post stress exposure to information/misinformation conveyed during interactions with the police, legal professionals, the media, friends and family.

As clinicians, we work with people who suffer from distressing memories about stressful and traumatic events; in our clinical experience, it is not necessary to view these traumatic memories as indelible or as high fidelity recordings of what happened to the person. By working with a patient's memories and their appraisal of what they have experienced, we can still provide effective treatment and relief from suffering.

Although we believe that mental health professionals ought to play a role in educating the courts about mental illness and how trauma may affect people, we believe it is unwise for us to become advocates within the judicial system for the idea that traumatic memories are indelible, supra-normal and factual accounts of events. To do so may only undermine the credibility of our profession in the eyes of the court and contribute to miscarriages of justice.

References

Brewin, C. R., Huntley, Z., & Whalley, M. G. (2012). Source memory errors associated with reports of posttraumatic flashbacks: A proof of concept study. *Cognition, 124*, 234–238.

Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J., & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: A synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law. *Neural Plast: 60803. Review, 7*, 7.

Eid, J., & Morgan, C. A. III, (2006). Dissociation, hardiness and performance in military cadets participating in survival training. *Military Medicine, 171*(5), 436.

Engelhard, I. M., van den Hout, M. A., & McNally, R. J. (2008). Memory consistency for traumatic events in Dutch soldiers deployed to Iraq. *Memory, 16*(1), 3–9. Jan.

Gold, P. E., & Korol, D. L. (2012). Making memories matter. *Frontiers in Integrative Neuroscience, 6*, 116. doi: 10.3389/fnint.2012.00116. Epub 2012 Dec 18.

James W., (1890), 1983. The Principles of Psychology, The Harvard University Press, Cambridge, MA.

Krinsley, K. E., Gallagher, J. G., Weathers, F. W., Kutter, C. J., & Kaloupek, D. G. (2003). Consistency of retrospective reporting about exposure to traumatic events. *Journal of Traumatic Stress, 16*(4), 399–409. Aug.

Lommen, M., Engelhard, I., & van den Hout, M. (2013). Susceptibility to long-term misinformation effect outside of the laboratory. *European Journal of Psychotraumatology*, 4.

McIntyre, O., McGaugh, J. L., & Williams, C. L. (2012). Interacting brain systems modulate memory. *Neuroscience and Behavioral Reviews, 36*, 1750–1762.

Morgan, C. A., III, Doran, A., Steffian, G., Hazlett, G., & Southwick, S. M. (2006). Stress induced deficits in working memory and visuo-constructive abilities in special operations soldiers. *Biological Psychiatry, 60*(7), 722–729.

Morgan, C. A., III, Hazlett, G., Baranoski, M., Doran, A., Southwick, S. M., & Loftus, E. (2007). Accuracy of eyewitness identification is significantly associated with performance on a standardized test of recognition. *International Journal of Law & Psychiatry, 30*(3), 213–223.

Morgan, C. A., III, Hazlett, G. A., Dial Ward., M., & Southwick, S. M. (2008). Baseline dissociation and prospective success in special forces assessment and selection. *Psychiatry, 5*(7), 52–57.

Morgan, C. A., III, Hazlett, G. A., Doran, T., Garrett, S., Hoyt, G., Baranoski, M., et al. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Psychiatry and the Law, 27*(3), 265–279.

Morgan, C. A., III, Hazlett, G., Wang, S., Richardson, G., Schnurr, P., & Southwick, S. M. (2001). Symptoms of dissociation in humans experiencing acute uncontrollable stress: a prospective investigation. *American Journal of Psychiatry, 158*(8), 1239–1247.

Morgan, C. A., III, Southwick, S. M., Steffian, G., Hazlett, G. A., & Loftus, E. F. (2013). Misinformation can influence memory for recently experienced, highly stressful events. *International Journal of Law & Psychiatry, 36*(1), 11–17. Jan-Feb.

Morgan, C. A., III, Wang, S., Hazlett, G., Rassmusson, A., Anderson, G., & Charney, D. S. (2001). Relationships among cortisol, catecholamines, neuropeptide Y and human performance during uncontrollable stress. *Psychosomatic Medicine, 63*, 412–442.

Morgan, C. A., III, Wang, S., Mason, J., Hazlett, G., Fox, P., Southwick, S. M., et al. (2000). Hormone profiles in humans experiencing military survival training. *Biological Psychiatry, 47*, 891–901.

Morley, J. E., & Farr, S. A. (2012). Hormesis and amyloid B protein: physiology of pathology? *Journal of Alzheimers Disease, 29*, 488.

Roemer, L., Litz, B. T., Orsillo, S. M., Ehlich, P. J., & Friedman, M. J. (1998). Increases in retrospective accounts of war-zone exposure over time: The role of PTSD symptom severity. *Journal of Traumatic Stress, 11*(3), 97–605.

Southwick, S. M., Morgan, C. A., III, Nicolaou, A. L., & Charney, D. S. (1997). Consistency of memory for combat-related traumatic events in veterans of operation desert storm. *American Journal of Psychiatry, 154*, 173–177.

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Clinical Psychology Review 25 (2005) 935–953

CLINICAL PSYCHOLOGY REVIEW

Consistency of memory for emotionally arousing events: A review of prospective and experimental studies

Anne E. van Giezen [a,*], Ella Arensman [a,b], Philip Spinhoven [a,c], Gezinus Wolters [d]

[a] Department of Clinical and Health Psychology, Leiden University, P.O. Box 9555 2300 RB Leiden, The Netherlands
[b] National Suicide Research Foundation, Cork, Ireland
[c] Department of Psychiatry, Leiden University Medical Center, The Netherlands
[d] Department of Cognitive Psychology, Leiden University, Leiden, The Netherlands

Received 29 November 2004; received in revised form 7 March 2005; accepted 22 April 2005

Abstract

Although emotionally arousing events are more memorable than ordinary daily life events, the nature of memories for emotionally arousing events is widely debated. On the one hand, researchers consider memories for highly emotional events as malleable and subject to distortion, while on the other hand these memories are perceived as both indelible and consistent over the lifetime. Up till now, a systematic comparison of research findings on consistency of memory for emotional events is lacking. This paper is the first effort to summarize available studies on consistency of memory for emotionally arousing events and to address methodological limitations and suggestions for future research as well. In general, findings show that quality of the selected studies is sufficient to good, with studies with victims of assault and studies on war-exposure reaching higher quality scores than studies on flashbulb memories and experimental memory studies. Victims of assault or war-exposure tend to amplify their memories for the event, while results from flashbulb memory research and experimental research suggest that memory for emotional events is either stable or diminishes over time.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Consistency; Accuracy; Memory; Emotional events; Trauma; Review

* Corresponding author. Tel.: +31 715273744; fax: +31 715274678.
E-mail address: giezen@fsw.leidenuniv.nl (A.E. van Giezen).

0272-7358/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cpr.2005.04.011

1. Introduction

Although it is generally acknowledged that an event is more accurately and more vividly remembered when it is emotionally arousing (Bremner, Krystal, Charney, & Southwick, 1996; Rolls, 1990), differences in opinion about the nature and recall of memories for traumatic events remain. The nature of memories for traumatic events is widely debated in the popular and scientific literature.

On the one hand, researchers consider memories for traumatic events as both indelible and accurate over the entire lifetime (Bohannon, 1988; Brown & Kulik, 1977; Conway et al., 1994; Pillemer, 1984), while on the other hand such events are also perceived as malleable and subject to substantial distortion and alteration (Southwick, Morgan, Nicolaou, & Charney, 1997). According to Mechanic, Resick, and Griffin (1998), for example, emotional memories can be subject to failed, confabulated, or incomplete recall.

These different viewpoints reflect two distinctive concepts of how the brain stores memories. The first view is the so-called 'static view of memory', stating that the brain stores memories of an event like a time capsule or videotape. Despite the time that elapses, all memories will be accurate, stable, and unchanged from their original occurrence. The second concept is referred to as the 'dynamic view of memory'. According to this view, memories can change over time and they are influenced by new events. Recollections of the original event can change over time. They can be incomplete, distorted, or even more complete (Zola, 1998). Support for the assumption that memories for trauma are indelible and consistent over time mainly seems to come from research into "flashbulb memory". "Flashbulb memories" are remarkably vivid, detailed, and accurate recollections of the circumstances in which one first learned of an unexpected and shocking event, like the death of Princess Diana or the assassination of President Kennedy. Those flashbulb memories resemble a "photographic print", complete with extreme detail (Brown & Kulik, 1977). Memory for these types of events is usually persistent and accurate, although some distortion can occur (Christianson, 1989; Neisser & Harsch, 1992; Schacter, 1996). Experimental research into emotion and memory also shows that memory for emotionally arousing events is generally accurate, with distortions occurring at the level of peripheral details instead of central aspects of an event (Cahill & McGaugh, 1995; Christianson, 1992; Heuer & Reisberg, 1990).

The notion of a number of authors that memory for trauma is malleable and subject to distortion is mainly based on experimental research into inaccuracies in the recall of neutral information, the misinformation effect and other memory errors, e.g., source monitoring errors, time-slice errors, and false childhood memories (Hyman & Loftus, 1998; Loftus, 1993, 2003; Loftus & Ketcham, 1994). The underlying assumption of their theoretical approach is that deficits in recalling traumatic events are attributed to the same processes of decay and interference as are deficits in memories for neutral information and ordinary events. Moreover, proponents of this viewpoint state that memory distortions can be induced by giving false post-event information. According to Hyman and Pentland (1996), even entire scenes of stressful events can be fabricated and inserted into autobiographical memory. Studies of real-life traumatic events also indicate that memories of these events are malleable and subject to substantial distortion (Fergusson, Horwood, & Woodward, 2000). However, there are also indications that these memories are generally persistent and often impressively accurate (Goodman et al., 1999).

A growing number of studies has investigated the extent to which people are consistent in their recall of a specific event over time. Among these studies, different terms and definitions are used to indicate consistency of memory. Sometimes, the term "accuracy' is used interchangeably with consistency. However, they refer to different concepts. A report can be highly consistent over time without necessarily being accurate. Consistency of memory can best be described as the same information being

reported at different points in time. Responses are also consistent even though no information at all can be remembered at different points in time. Inconsistency can be defined as a change in responses between reports over time. Here, a distinction can be made between omission errors, referring to a reduction in information in reports compared to previous reports, and commission errors, referring to an increase in information over time. In order to assess (in)consistency, assessments using exactly the same instrument should be performed on at least two different occasions.

As opposed to consistency, accuracy can be defined as the agreement between recall and an objective record or facts of what has occurred. In order to assess accuracy, an 'objective' assessment of facts has to be available, allowing verification of reported facts, which is in line with the definition by Fivush (1993): "Accuracy is the agreement between the individuals recall and either an objective record of the event or social consensus from other participants of the event as to what occurred" (p. 22).

In sum, although a high degree of inconsistency in recall does not support the notion that memory for traumatic events is indelible, both clinical and experimental research findings suggest that memories of emotionally arousing events are often well retained. Moreover, there is little basis in supposing that they are less susceptible to distortion than are memories of neutral information or ordinary events (Zola, 1998). Individuals seem to gradually recall traumatic events in the course of time because information about the event is available but often not directly accessible (Brewin, Dalgleish, & Joseph, 1996; Melchert & Parker, 1997). In the absence of empirical evidence for real-life traumatic events, however, it remains undecided whether inconsistencies always imply inaccuracy in recall of memory for real-life traumatic events.

Apart from conceptual problems, research has pointed out that other factors may influence consistency. One factor concerns the severity of the traumatic event. Those persons who suffer from more exposure to severe trauma may encode or retrieve information during the trauma differently (Foa, Molnar, & Cashman, 1995; Van der Kolk & Fisler, 1995) and are at higher risk for the development of PTSD at a later stage than those exposed to less severe trauma (Bremner & Brett, 1997; Koopman, Classen, & Spiegel, 1994; Marmar et al., 1994). Recollections of severe trauma are expected to be more consistent than less severe trauma (Krinsley, Gallagher, Weathers, Kutter, & Kaloupek, 2003). Another factor is the degree of involvement in the traumatic event. Events that are experienced directly by victims have found to be more consistently remembered than events that are experienced without personal involvement, for example, witnessing a crime. Furthermore, there appears to be a relation between psychological and psychiatric symptoms experienced at the time of, or shortly after an emotionally arousing event and inconsistency. Posttraumatic stress symptoms and peritraumatic dissociation in reaction to a traumatic event appear to be associated with amplification of memory following the event (Mechanic et al., 1998; Southwick et al., 1997; Wyshak, 1994). With regard to the nature of the information recalled, there are indications for a higher degree of consistency for the central core of information compared to specific details (Schacter, 1996). Another factor is the developmental phase in which the event to be remembered occurred. There are indications for more inconsistent memory reports when the events are experienced at a younger age (Ghetti, Goodman, Eisen, Qin, & Davis, 2002). Ultimately, the length of time between occurrence of an emotionally arousing event and first assessment of recall, as well as the time between first and subsequent assessments may also have an influence on the degree of consistency of reports of the event (Winningham, Hyman, & Dinnel, 2000).

The aim of this paper is to systematically screen and review available studies focusing on consistency of memories with regard to emotionally arousing negative events. Examining the available evidence, each of the studies will be reviewed according to aspects that have shown to be related to consistency of memory

for emotionally arousing events. Although the nature and recall of traumatic memories is widely debated, a systematic screening of the literature and comparison of research findings on consistency of memory is lacking. Also, relatively few studies have been conducted including samples of victims of a stressful or traumatic event. Prior to comparison of the study findings, the methodological quality of the studies will be systematically screened. Specific methodological problems will be discussed, followed by recommendations for future studies in this area. Implications for clinical and legal practice will be discussed as well.

2. Methods

2.1. Literature search

A literature search was carried out using the following electronic bibliographic databases: PsycINFO: 1887 to 2004, MEDLINE: 1966 to 2004, and Current Contents: 1995 to 2004. A computer search was carried out, using a wide range of key words[1] to indicate consistency of memory for emotionally arousing events. Reference lists of available reviews and studies were screened as well.

2.2. Inclusion criteria

Study reports were included if the following inclusion criteria were met: (a) The features of the event studied had to be either in line with criterion A of the DSM-IV definition: "a person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others" (American Psychiatric Association, 1994, p. 427), or presence of an emotionally arousing negative event without specific reference to the DSM criterion; (b) differences in reports of a traumatic event had to be assessed on at least two different times. Inconsistency of recall was operationalized as the total number of responses that changed from time 1 to time 2.

2.3. Grouping of studies

Specific groups of victims or samples or differences between types of events, and situations in which the event occurred were used as a basis to categorize the different studies. The first cluster (A) comprised studies of victims of sexual and physical assault; the second cluster (B) comprised studies of war-zone exposure; the third cluster (C) comprised studies of flashbulb memories and the fourth group (D) comprised experimental studies including both victims and non-victims.

2.4. Quality assessment

In order to compare the results of the different studies, their methodological quality should be taken into account. Quality of the studies included in this review was rated by two independent reviewers (AvG, EA). Cohen's κ was used to assess the agreement between the two reviewers and was found to be

[1] Keywords used in the literature search: consistency, stability, accuracy, reliability, narrative, discrepancies, change, autobiographical memory, retrospective, recall, memory, account, report, recollections, and flashbulb memory.

0.80. We formulated four criteria in order to rate the quality of the method used to examine consistency of memory. These were:

(1) Operationalization of consistency of memory: If consistency of memory was clearly operationalized, the study received a score of 1. If no clear operationalization was provided, a score of 0 was given.
(2) Selection of the sample: Representative samples received a score of 1. If no representative sample was used, a 0 score was given.
(3) Type of assessment of memory of the event: (a) If data were primarily collected through interviews in a face to face setting, the publication received a positive score. If data were collected through self-report questionnaires, a 0 score was given. (b) If a standardized procedure was used, including the use of similar instruments at first and later assessments, the publication received a positive score, if no standardized procedure was used, a 0 score was given. (c) If data were collected through validated instruments, the study received a positive score. If no validated instrument was used, a 0 score was given. If criteria a, b, and c were all rated positive, a score of 1 was given. If only one criterion was rated positively, the publication received a score of 0.33, if two were rated positive, a score of 0.66 was given.
(4) Appropriate statistical analyses: If statistical analyses were used appropriately, without violation of the assumptions of statistical tests, and if important parameters were reported, a score of 1 was given. If statistical analyses were used inappropriately, i.e., if assumptions were violated, the publication received a 0 score.

For each study report, a total quality score was calculated based on the number of criteria that were fulfilled. All four criteria could receive a total score of 1, so quality scores ranged from 0 (lowest quality) to 4 (highest quality).

3. Results

3.1. Studies included in the review

Thirty-seven studies were identified through the combined search strategies and found to be eligible for inclusion in the study. The identified studies were grouped according to the criteria described in the Method section. Table 1 summarizes the clusters of studies included in the review, details of the design, methods, and sample. For each of the studies, the research quality was assessed according to the criteria described above. The total quality score of the studies varied from 1.33 to 4, with an average of 3. The earliest study was published in 1984 and the most recent study was published in 2004.

3.2. Studies including victims of assault

Studies included here comprised victims of childhood sexual abuse, recent sexual abuse, or physical maltreatment. Seven papers were included in this cluster, comprising eight separate studies. The overall quality of the studies in this cluster was high ($M = 3.5$), with three studies scoring positive on all four criteria. Four studies were given a score of 3.66 and one study obtained a score of 1.33. In four out of

Table 1
Methodological aspects and quality of the studies by cluster

Author(s), year	*N*	Population	Response, first assessment (%)	Response, second assessment (%)	Age	% men	Method of assessment, interview(I)/ self-report(S)	Interval event, first assessment	Interval, time 1, time 2	Quality
Cluster A victims of assault										
Foa et al., 1995	14	Sex. assault	n.a.	n.a.	30.1	0	S+I	4.85 years	1 month	1.3
Mechanic et al., 1998	92	Rape	n.a.	n.a.	29.2	0	S+I	2 weeks	2.5 months	4
Goodman et al., 1999	50	Abuse	80	92.6	39.9	42	I+S	–	2 weeks	4
Fergusson et al., 2000	1265	Cohort	78	n.a.	21	50.2	I	Min. 2 years	3 years	4
Zoellner et al., 2001	30	Assault	n.a.	n.a.	31.4	0	I+S	10.13 days	10 weeks	3.6
Zoellner et al., 2001	60	Assault	n.a.	n.a.	31.4	0	I+S	31.67 days	86.13 days	3.6
Ghetti et al., 2002	222	Abuse	n.a.	n.a.	7.3	45	I	–	3 days	3.6
Aalsma et al., 2002	217	Abuse	n.a.	48.9	17	17	S	–	7 months	3.6
Cluster B war-zone exposure										
Wyshak, 1994	30	Refugees	n.a.	n.a.	44.3	40	S	–	1 week	3
Southwick et al., 1997	59	Gulf War	95	n.a.	29.9	78	S	1 month	2 years	3.6
Roemer et al., 1998	460	Somalia	n.a.	n.a.	26.7	92	S+I	1 year	21.3 months	3.6
Niles et al., 1999	38	Vietnam	n.a.	34.3	39.7	100	S+I	30 years	4–8 years	4
King et al., 2000	2942	Gulf War	n.a.	22	30.2	92	S	5 days	18–24 months	3.6
Bramsen et al., 2001	137	Cambodia	55	30	29.1	97	S	2.5 years	9 months	3.3
Herlihy et al., 2002	39	Refugees	n.a.	90.7	39.5	53.5	I+S	–	3–32 weeks	3
Wessely et al., 2003	2370	War veterans	n.a.	78	30–44	78.7	S	5.5 years	3 years	3.6
Krinsley et al., 2003	76	Vietnam	n.a.	88.4	49	100	S+I	30–39 years	2–7 days	4
Cluster C flashbulb memories										
Pillemer, 1984	121	Assassination attempt Reagan	45	47	39	49	S+I	1 month	6 months	2.6
McCloskey et al., 1988	45/27	Challenger	90	7	n.a.	n.a.	S	3 days	9 months	2.3
Christianson, 1989	40	Olof Palme	n.a.	10	37.8	50	I	42 days	12 months	3.6

Neisser & Harsch, 1992	44	Challenger	n.a.	41.5	Students	47	S	1 day	30 months	2
Norris & Kaniasty, 1992	65	Hurricane Hugo	70	18	44	n.a.	I	6 weeks	9 months	2.6
Schwarz et al., 1993	12	School shooting	80	50	n.a.	n.a.	S	5 months	12 months	1.6
Weaver, 1993	22	Bombing Irak	n.a.	18	19–23	n.a.	S	2 days	3,12 months	1.6
Conway et al., 1994	396	Thatcher	n.a.	n.a.	Students	n.a.	S	24 days	12 months	2.6
Neisser et al., 1996	76/41/44	Earthquake	n.a.	26.2/13.2/ 31.9	Students	n.a.	S	2/3/15–21 days	18 months	3.3
Christianson & Engelberg, 1999	203	Estonia	n.a.	32	n.a.	59.6	I	1 day	12 months	3.6
Schmolck et al., 2000	63	O.J. Simpson verdict	n.a.	60	Students	52	S	3 days	15,32 months	2.6
Winningham et al., 2000	65	O.J. Simpson verdict	n.a.	n.a.	Students	12	S	5 h/1 week	8 weeks	2.6
Curci et al., 2001	124/105	Death of Mitterand	n.a.	34.5	28.4/34.5	29.2/48.3	S	1–2 months	1 year	2.3
Hornstein et al., 2003	66/49	Diana	100	76/53	Students	29/30	S	1 week	3–18 months	2.3
Talarico & Rubin, 2003	54	September 11	n.a.	n.a.	18.7/17.8/ 19.1	22/33/22	S	1 day	1,6,32 weeks	2.3
Smith et al., 2003	93	September 11	n.a.	n.a.	19.3	25.8	S	1 week	6 months	2.6
Tekcan et al., 2003	483?	September 11	n.a.	n.a.	28.89	35?	S	3 days	6,12 months	2.6
Lee & Brown, 2003	142	September 11	n.a.	n.a.	Students	?	S	4–24 h/ 10 days	7 months	2.6
Cluster D experimental studies										
Fisher & Cutler, 1995	33	Students	n.a.	n.a.	Students	n.a.	S	40 min	5 days	2.3
Fisher & Cutler, 1995	27	Students	n.a.	n.a.	Students	n.a.	S	20 min	2 days	2.3
Fisher & Cutler, 1995	135	Students	n.a.	n.a.	Students	n.a.	I	1–3 days	2 weeks	2.3
Fisher & Cutler, 1995	85	Students	n.a.	n.a.	Students	n.a.	I	1–2 days	1 week	2.3
Brewer et al., 1999	62	Students	n.a.	n.a.	24.8	30.6	S	–	2 weeks	2.6
Candel et al., 2004	52	Students	n.a.	79	20	11.5	S	–	3–4 weeks	2.3

21

eight studies, the victims were all female. The majority of studies used standardized interview measures for the assessment of consistency of memory. The time period between the index event and time 1 ranged from 10 days to 4.85 years. The interval between time 1 and time 2 varied from 3 days to 36 months, with an average interval of 6.6 months. For detailed information on the methodological aspects of the studies, see also Table 1.

Although studies used different terms to indicate consistency, the primary focus of all eight studies was to examine consistency of memory for the index event over time. All studies found discrepancies in reports of the index event over time.

In their study, including female rape victims, Mechanic et al. (1998) found that, while memory deficits were common 2 weeks following the assault (32%), memory for details of the index event improved significantly over a 3-month period ($p<0.05$). Only 16% reported significant amnesia 3 months after the rape experience.

Fergusson et al. (2000) found that reports of childhood physical and sexual abuse were unstable and inconsistent over time ($p<0.01$). Half of those who reported abuse at one assessment failed to report it at another one. Inconsistency was unrelated to the respondent's psychiatric status. The results from Aalsma, Zimet, Fortenberry, Blythe, and Or (2002) showed that stability of reports of childhood sexual abuse (CSA) by adolescents was poor. Fifty-eight percent reported no abuse at both points in time, 20% reported CSA on both occasions, while 22% were inconsistent reporters of CSA. According to Ghetti et al. (2002), children's reports of sexual abuse were more consistent than were reports of physical abuse. The memories of older children were more consistent compared to those of younger children ($p<0.01$) and reports of sexual abuse from girls were more consistent than reports from boys ($p<0.01$).

Goodman et al. (1999) interviewed men and women with a serious mental illness about adult physical and sexual abuse, childhood sexual abuse, and PTSD symptoms. They found that reports of abuse were fairly consistent, with women being even more consistent than men. Men showed a tendency to report significantly less experiences of sexual abuse since age 16 and sexual abuse in the past year at the second assessment than at the first one ($p<0.05$). Both men and women were consistent in their reports of the severity of PTSD symptoms.

3.2.1. Studies including a psychotherapeutic intervention

Intervention studies differ from the studies above, because the effect of the intervention can be of influence on consistency of memory. Zoellner, Sacks, and Foa (2001) conducted two studies among female victims of recent assault. The results of study 1 indicate that memory of the fear associated with the assault remained stable, while memory of general emotional intensity and memory of dissociative intensity for the traumatic event changed over time ($p<0.06$), reflecting a trend that did not reach statistical significance. The second study compared treatment seeking assault victims, with either acute or chronic PTSD symptoms. Victims diagnosed as having acute PTSD reported a decrease in memory for emotional and dissociative intensity for the index event ($p<0.05$), while victims with chronic PTSD reported an increase in memory for general emotional and dissociative intensity of the index event 12 weeks after the initial report ($p<0.05$). According to Foa et al. (1995), narratives of rape tended to become longer from pre- to post-treatment assessment ($p<0.08$), reflecting a trend. Furthermore, the percentage of aspects representing actions and dialogue decreased ($p<0.06$) and the aspects representing thoughts and feelings increased 18 weeks after the initial report ($p<0.05$). This increase in organization of the narratives was related with an improvement in depression after treatment ($p<0.02$). However,

reliability and validity of the findings of the latter study is questionable since most of the quality criteria were not fulfilled.

3.3. Studies of war-zone exposure

Studies included here comprised war veterans and mainly male subjects. Nine studies fulfilled the inclusion criteria. Two studies included soldiers that served in peace-keeping missions, two studies included Gulf War veterans, one study included both Gulf War and Bosnia veterans, two studies included refugees, and two studies included Vietnam veterans. The quality scores of the papers in this cluster were generally high ($M=3.5$); all papers were given a score of 3 or higher. One study scored positive on all criteria, five studies received a score of 3.6, and one study received a score of 3.3. Detailed information on methodological aspects of the studies is presented in Table 1. Most studies used self-report questionnaires to assess consistency of memory for the index event. Men were over represented in most studies, ranging from 40% to 100%. The length of time between the index event and time 1 ranged from 5 days to 39 years, which is an extremely wide range. The interval between time 1 and time 2 varied from 1 week to 79 months, with an average interval of 23.3 months.

All nine studies described changes in reports of war-zone exposure over time. Three studies also found an interaction between PTSD symptom severity and amplification of memory. According to Niles et al. (1999), most reports of trauma exposure are relatively stable over time. However, in their study with treatment seeking Vietnam veterans, some participants showed dramatic changes in report over time. Southwick et al. (1997) found that recollections of combat-related events were not stable over time ($p<0.02$). Eighty-eight percent of Gulf War veterans changed at least one of their responses over time. Seventy percent of the respondents recalled an event at time 2 they had not reported at time 1, while 46% reported an event at time 1 and failed to mention it at time 2. According to Southwick et al. (1997), inconsistency of memory for combat exposure appeared to be significantly associated with PTSD symptoms ($p<0.002$). Respondents with higher PTSD scores 2 years after their return from the Gulf War tended to amplify their memory for combat exposure. King et al. (2000) also found that reports of war-zone stressor exposure appeared to be vulnerable to some shifts over time. In accordance with the findings of Southwick et al. (1997), most changes in answers in their study were of the no to yes type. The total number of changes in report of stressor exposure from time 1 to time 2 was modestly associated with PTSD symptom severity at time 1 and time 2 (0.14, 0.24, $p<0.001$). Results from Roemer, Litz, Orsillo, Ehlich, and Friedman (1998) also support the dynamic view of memory. They found that memories of war-zone exposure were not fixed or indelible, but that they were variable over time. Soldiers who had served in the peace-keeping mission in Somalia reported a significant increase in the frequency of war-zone exposure ($p<0.001$) over a 1–3-year period. PTSD symptom severity at follow-up, in particular intrusive symptoms, appeared to be significantly associated with change in reports of war-zone exposure ($p<0.01$). Wessely et al. (2003) also found inconsistency in reported military hazards over time. They found that Gulf War veterans showed an increase in reports over time ($p<0.01$), while soldiers who served in Bosnia showed no significant increase of memory over time.

Bramsen, Dirkzwager, van Esch, and van der Ploeg (2001) found that 88% of the peace-keepers changed at least one answer from time 1 to time 2 on a 16-item trauma checklist, although none of the participants showed an increase in the number of reported events over time. Also, inconsistency in

reports of exposure to stressful events was not significantly associated with PTSD symptom severity, which is in contrast with findings from previous studies by Southwick et al. (1997), Roemer et al. (1998), and King et al. (2000).

Wyshak (1994) examined refugees from Southeast Asia in a psychiatric clinic about traumatic events and psychiatric symptoms. Results showed that number of traumatic events reported increased slightly but not significantly from time 1 to time 2. Psychiatric symptom scores increased significantly from time 1 to time 2 ($p<0.05$). Number of traumatic events reported was positively correlated with both PTSD-related symptoms and non-PTSD-related symptoms at time 1 ($p<0.003$, $p<0.007$). A remarkable finding is that those who were inconsistent in their reports of the traumatic events showed less severe symptoms over time ($p<0.009$). According to Herlihy, Scragg, and Turner (2002), discrepancies in the reports of asylum seekers were common. They interviewed Kosovan Albanians and Bosnians about a traumatic event and a non-traumatic event twice over a period of 3 to 32 weeks. Results show that discrepancies were found for all participants. More discrepancies were found in the peripheral details than in central details ($p<0.05$). Also, the length of time between the interviews and posttraumatic stress symptoms was significantly related to the number of discrepancies. Participants with high levels of posttraumatic stress at time 1 showed more discrepancies when the time between interviews became longer ($p<0.05$). Krinsley et al. (2003) interviewed male Vietnam veterans twice on exposure to traumatic events. Results showed that average number of reported times being exposed to traumatic events increased significantly from interview 1 to interview 2 ($p<0.05$). Only 11% reported the same number of events on both occasions, whereas 38% reported fewer events over time and 51% reported more events. In addition, events that were directly experienced were more consistently remembered than events that were experienced from a witness perspective. Also, more severe traumatic events were remembered more consistently. With regard to the period in life in which the event occurred, results showed that the number of reported events from adulthood and childhood both increased over time, but only childhood events showed a significant amplification from T1 to T2 ($p<0.05$).

3.4. Studies of flashbulb memories

Studies included in this cluster comprised studies of flashbulb memories. All studies included unexpected events that were experienced by or known to a large group of people. However, there was considerable diversity with regard to the types of events included, which limited a strict comparison of the research findings. Eighteen studies were included in this cluster. The studies were grouped into three categories according to similarities of the nature of the events: (1) assassination attempts, (2) disasters, and (3) public persons. Eventually, two studies that could not be included in one of the three categories were classified as (4) ‘other‘ studies, of which the results are described separately.

The overall quality of the research papers in these clusters was relatively low ($M=2.5$), due to the frequent use of self-report questionnaires and student populations. Quality scores ranged from 1.6 to 3.6, with only three studies obtaining a score higher than 3 (see also Table 1). The studies in this cluster differ from the studies in clusters A and B, in that the subjects in the studies in cluster C were often not directly exposed to traumatic events, but they were indirectly confronted with an emotionally arousing negative event, e.g., obtaining information on an assassination of a celebrity or disaster, through the media. The majority of studies used self-report questionnaires to assess consistency of memory. Most subjects

included in these studies comprised either students or community samples. The time period between the initial event and time 1 varied from 1 day to 5 months. The interval between time 1 and time 2 varied from 1 week to 30 months.

3.4.1. Assassinations and assassination attempts

Pillemer (1984) found in his study of the 1981 assassination attempt on Reagan that memories for this flashbulb event were highly consistent over a 6-month interval. The intensity of the emotional reaction and the extent to which the event came as a surprise, as reported by the subjects, were significantly associated with consistency of memory ($p<0.04$).

In contrast, Christianson (1989) reported that memories for assassination of the Swedish Prime Minister Olof Palme were significantly ($p<0.001$) reduced over time. Eighty percent of the subjects showed consistency over time with respect to the central aspects of the event. Regarding specific details, only 53% of the subjects' recollections were consistent.

3.4.2. Disasters

McCloskey, Wible, and Cohen (1988) found that flashbulb memories for the explosion of the space shuttle challenger were quite consistent over time, although they were also subject to normal forgetting and inaccuracies. Over a 9-month retention interval, 60.7% of the responses were consistent, 5.6% of the responses were inconsistent, 8.4% were not remembered at time 2, 6.5% of the responses were more specific, and 18.7% were more general. Neisser and Harsch (1992) also questioned participants on their memories for the explosion of the challenger. Most subjects told the same stories at time 1 (1 day) and time 2, with an average interval of 2.5 years. However, most of these memories were not very consistent, since they showed major discrepancies between the recall and the original report. Over 25% of the subjects were inconsistent on all questions. In their study, Norris and Kaniasty (1992) found that self-reports regarding Hurricane Hugo were remarkably stable over a 9-month period. Eighty-eight percent of the responses were the same at time 1 and time 2. According to Christianson and Engelberg (1999), recall of the Estonia ferry disaster was fairly inconsistent. After 6 months, subjects were able to recall the central aspects of the event to some degree (69%), but they showed considerable loss of memory for specific details (47%).

Neisser et al. (1996) studied recollections of three groups of informants–two in California and one in Atlanta–regarding the 1989 Loma Prieta earthquake in California. Results showed that personal involvement appeared to be predictive of greater consistency. The Californians' recall (99%, 96%) was more consistent than the recollections of the inhabitants of Atlanta (55%).

Talarico and Rubin (2003) studied the consistency of memories for the terrorist attacks of September 11th and of an everyday event. Results indicated that both the flashbulb and everyday memories show a decline over time. However, subjects had more confidence in their flashbulb memories. In another study on consistency of memory of September 11th, Smith, Bibi, and Sheard (2003) found that Canadian Psychology students' autobiographical memory for details when learning about the event was significantly ($p<0.001$) more consistent (65.3%) than their memory for the emotionally arousing event itself (30.9%). These findings were confirmed by Tekcan, Ece, Gulgoz, and Er (2003) who investigated consistency of memory following the event of 11th. Consistency of autobiographical memory was significantly higher than consistency of memory for event details ($p<0.01$). Results from Lee and Brown (2003) showed that 66.5% of the participants in their study on September 11 were categorized as having flashbulb memories. Although subjects gave significantly less information at a 7-month follow-

up ($p<0.01$), most of the reports were partially (12.7%) or completely (75.5 %) consistent, only few reports were completely inconsistent (11.8%).

3.4.3. Public persons

In a study on the acquittal of O.J. Simpson, Winningham et al. (2000) found that memory reports after an 8-week interval were less consistent when the first report followed within 5 h of the acquittal than when the report was delayed (1 week after the event) ($p<0.02$). According to Schmolck, Buffalo, and Squire (2000), recollections of the O.J. Simpson trial were more accurate at a follow-up after 15 months (50%) than after 32 months (29%). However, this difference was not significant. Conway et al. (1994) found that memories of U.K. citizens for the resignation of Margaret Thatcher were fairly consistent over time, while memories of a non-U.K. group were not. Over 86% of the U.K. citizens had flashbulb memories, compared to 29% of the none-U.K. group 1 year following the event. The level of importance attached to the event and the level of affective responses were significantly associated with the formation of flashbulb memories ($p<0.01$). Also, Curci, Luminet, Finkenauer, and Gisle (2001) found that memory for the death of former French President Mitterrand remained more stable over time for French people compared to those with Belgian nationality ($p<0.001$). Although both groups showed impaired memories over time, the French were also more confident about their recollections of flashbulb memory attributes.

Hornstein, Brown, and Mulligan (2003) examined US student's flashbulb memories concerning the death of Princess Diana. The results showed that the reports were both accurate and consistent over an 18-month interval. Emotional intensity and rehearsal (talking about the incident) were both related to accuracy.

3.4.4. Other studies

Weaver (1993) found that psychology students' memories of both the bombing of Iraq in 1991 and a personal event were subject to normal forgetting. These memories were not very accurate following a 3-month delay, but showed little change from 3 months to a year. Confidence levels for memories of the circumstances when learning about the bombing of Iraq were significantly higher ($p<0.05$).

Schwarz, Kowalski, and McNally (1993) found in their study on a manmade disaster that all self-reports of school personnel after a school shooting were subject to change over time. Subjects showed either a decrease or an increase in reported information over a 12-month period. Decreases were more common. An increase in reported information appeared to be associated with more PTSD symptoms at initial assessment ($p<0.01$), while a decrease appeared to be associated with a reduction in symptoms of anxiety and depression and an improvement of self-confidence ($p<0.04$).

3.5. Experimental studies

Studies included in this cluster comprised experimental studies. Three studies were included in this cluster. The quality scores were rather low (2.3) and were the same for all three studies. All studies included a student population.

In a study by Brewer, Potter, Fisher, Bond, and Luszcz (1999), 62 undergraduate students were asked to watch a video, showing a bank robbery. They completed a questionnaire 1 h after watching the video. Two weeks later, respondents were interviewed again using the same questionnaire. Except two, all respondents reported contradictory information.

In their experiment, Candel, Smeets, and Merckelbach (2004) showed female students an emotional fragment of the movie "American History X". Students were asked to provide a detailed written account of the movie fragment on two different occasions, immediately afterwards and 3 to 4 weeks later. The mean number of discrepancies between the two accounts was 6.13 (S.D. 2.41, range 2–13). Results show that inconsistencies were not related to accuracy of the testimonies.

Fisher and Cutler (1995) performed four experiments with undergraduate students to explore the relationship between consistency and accuracy of eyewitness testimony. Overall, the average proportion of consistent statements was 76%. In accordance with Candel et al. (2004), consistency of reports was not found to be predictive of accuracy.

4. Discussion

The current review first of all shows that worldwide relatively few studies have been conducted on consistency of emotionally arousing events. Moreover, with exception of the studies of flashbulb memories, the studies in all other categories have been published in recent years, which is surprising considering the ongoing discussion about the nature and recall of memories for traumatic events over the last decennia. This might be partly explained by progress of scientific research in this area, where initially, research into memory processes underlying the recall of neutral events was expected to provide answers that could be generalized to the recall of emotionally arousing events. Research evidence, however, has revealed that both emotional reactions at the time of occurrence of an event and present mood may distort memories of the event (McNally, 2003). The current review is the first effort to investigate this issue in detail by screening and comparing available studies.

Results of this review show that overall quality of the studies varied from sufficient to good, ensuring the reliability and validity of the outcomes of most of the studies included. The quality scores of both studies with assault victims and victims of war-exposure were high, in contrast with the overall quality of the flashbulb studies and experimental studies. In the flashbulb memory cluster, as well as in the experimental cluster, criteria for the selection of the sample, use of interview measures, and use of a standardized instrument were often not fulfilled.

Furthermore, all studies in cluster C (flashbulb memories) showed either fairly consistent memory reports over time or a tendency towards a decline in memory reports (normal forgetting), while in clusters A (victims of assault) and B (war-zone exposure) amplification of memory was more common with nine studies finding an increase in memory reports for the initial event. In cluster D (experimental studies), all studies found inconsistency of memory reports. The differences between the clusters in the degree and direction of inconsistency can be partly explained by the nature of the traumatic event, the extent of involvement, and the amount of psychological or psychiatric symptoms present.

With regard to these factors of influence to consistency, results show that, although studies included in this review were all focussing on consistency of memory, the concept of consistency was not always clearly distinguished from the concept of accuracy, due to the absence of operational definitions for both concepts. It is generally known that "accuracy" of memory for emotionally arousing events is difficult to establish, because detailed objective reports of the traumatic event are often not available. Instead of assessing accuracy, research has indicated that it has been more feasible to compare memory reports obtained soon after the event with reports obtained at a second point in time after the event. "Consistency is better measurable and is a necessary (though not sufficient) condition for accuracy" (Talarico & Rubin,

2003, p. 455). According to McNally, "consistency of reports over time has served as a proxy measure for memory accuracy" (p. 49).

From this review, it appears that the degree of involvement in and the severity of emotionally arousing events tends to be associated with greater consistency over time or amplification of memory (Neisser et al., 1996). In line with these findings, results from the study of Krinsley et al. (2003) show that full-fledged traumatic events were more consistently remembered. Studies including victims who were personally involved in or directly exposed to the negative event revealed a different pattern than flashbulb studies or experimental studies, in which subjects were not personally involved in an emotional event. Instead of consistent responses or a decline in memory over time, their memory reports showed an amplification of memory.

Van der Kolk and Fisler (1995) already provided evidence for the fact that the emotional intensity of an emotionally arousing event interferes with the construction of a coherent narrative of what happened. Gradually, individuals seem able to recall more information about the event, which apparently is available but not directly accessible. Irrespective of the level of arousal during encoding, presence of PTSD and dissociative symptoms or current mood may also influence the consolidation and retrieval of memories for the traumatic event (McNally, 2003). Of those studies that looked at the relation of symptoms of PTSD and/or dissociation and inconsistency, the majority showed that an increase in PTSD symptoms was significantly associated with inconsistency of reports. Respondents with more PTSD symptoms were more likely to amplify their memory for the emotionally arousing event (King et al., 2000; Roemer et al., 1998; Schwarz et al., 1993; Southwick et al., 1997; Wyshak, 1994; Zoellner et al., 2001). Non-PTSD-related symptoms were also found to be related to the amount of reported information (Herlihy et al., 2002). Thus, trauma-related symptoms are negatively influencing the consistency of memory of a traumatic event over time. With regard to the nature of reported information, memory of stressful events appears to be more extensive than memory of non-stressful events; especially the central and critical details are better remembered, but not the peripheral details (Bremner et al., 1996). Of those studies who investigated differences between consistency of reports of central and peripheral information, all found that central aspects of the event were more consistently remembered than the specific details (Christianson, 1989, 1992; Herlihy et al., 2002).

The period in life in which a traumatic event occurs tends to be associated with memory inconsistency. Krinsley et al. (2003) and Ghetti, Goodman, Eisen, Qin, & Davis (2002) found that events occurring at a younger age are associated with less consistent memory reports. This effect may be partly explained by the long time period between the event and follow-up assessments.

The studies in this review used different time-periods between occurrence of the event and baseline assessment, and between baseline and second assessment to assess consistency over time. Overall, it appeared that the degree of consistency increased with a longer time period between the event and first assessment, and decreased as the time period between first and subsequent assessments became longer.

According to Winningham et al. (2000), the consistency of memories of flashbulb events depends on the point in time when the initial memory report of the event was obtained. A longer delay between a flashbulb event and the initial report appears to be associated with greater consistency of memory. Herlihy et al. (2002) and Schmolck et al. (2000) mention that not only the interval between the initial event and the first assessment matters, but that the length of time between baseline and second assessment also influences consistency of reports. They found that, when more time elapsed between assessments, reports became less consistent.

4.1. Methodological considerations

Although overall quality of the studies appeared to be sufficient or good, we identified a number of shortcomings among the studies, limiting a strict comparison of the findings. A common problem concerns the fact that consistency of memory was often not defined and measured in a standard way across all studies. Many studies used self-report questionnaires to assess consistency and some studies failed to use the same questionnaire at follow-up, which makes it almost impossible to assess consistency of memory over time. Often, no details were provided on the psychometric quality of the used instruments. Also, different coding systems were used to compare memory reports over time.

Overall, the majority of studies included small samples, which often did not appear to be representative. In the category of experimental studies, only three studies were identified, limiting generalization of the outcomes. In some of the clusters, especially in cluster A and cluster B, either women or men were overrepresented.

A specific characteristic of flashbulb memory studies, which is absent in the other clusters, is the important role of the media. Therefore, it is likely that flashbulb memories are partly determined and probably strengthened over time by contacts with newspapers, internet, and television (Winningham et al., 2000).

4.2. Implications for clinical and legal practice

In cases where an individual has been personally affected by a stressful or traumatic event, the absence of a complete reconstruction of the event in the first phase after the event is likely to be related to the emotional reactions at the time of the event and/or at later points in time when the event is discussed again. On the basis of the review, it can be concluded that it is likely that additional details of an emotionally arousing event may be remembered at a later stage. It is, however, unlikely that an individual will not remember the event at all or that the information concerning the event will be entirely distorted at a later stage. If an individual is not able to remember certain details at one point in time, it is recommended to address this again, at a later stage. When this is done, it is very important, both in clinical and in legal practice, to avoid the use of suggestive questions. Using standard and non-suggestive questions will enhance the consistency of reports (Krackow & Lynn, 2003).

Results of this review have implications for legal practice. Memory is a reconstructive process, which is prone to errors. Therefore, we cannot fully rely on its accuracy, completeness, and consistency. Eyewitness testimonies completed in situations where a victim is showing emotional reactions may not be entirely reliable. Testimonies completed at an early stage following a crime may be incomplete and inaccurate, in particular with regard to important details of an event. However, reports of criminal events can be highly consistent over time, without necessarily being accurate. Therefore, one has to be cautious using consistency as an indicator of accuracy and drawing conclusions from a single testimony (Morgan et al., 2004; Porter, Spencer, & Birt, 2003).

With regard to clinical practice, results indicate that therapists are not able to draw conclusions about the occurrence or accuracy of traumatic events in a patient's history. Fortunately, in most professional guidelines of memories of traumatic events, this limitation is acknowledged and therapists are advised to refrain from conclusions about the veracity of memories of trauma in clinical and legal settings (Health Council of the Netherlands, 2004).

4.3. Recommendations for future research

Assessment of accuracy of memory is complex and often consistency is used as a proxy measure for accuracy. Therefore, future studies on consistency of memory should include operational definitions of consistency. The methods used to assess consistency vary considerably across the different studies, thus limiting the possibilities to conduct more quantitative meta-analyses. Standardized measures for the assessment of consistency of memory are needed. Furthermore, use of interview measures to assess reports of an emotionally arousing event instead of self-report questionnaires is recommended. It is generally believed that interview measures are more reliable and more valid than self-report questionnaires (Dohrenwend, 1973; Southwick et al., 1997). More research is needed to investigate inconsistency of memory over time in more detail. For example, a distinction may be made between differences in consistency of reports of central and peripheral information and objective versus subjective aspects of an event. Inconsistencies may also be specified in omission and commission errors. In order to determine the association between consistency and accuracy, prospective studies are required addressing both concepts.

According to McNally (2003), not only mood at the time of occurrence of the event, but also at the time of reporting the event or discussing the event may influence the consistency of the reported information. Therefore, assessment of stress- or trauma-related psychological reactions, psychiatric symptoms, as well as general mood assessments should be carried out when investigating consistency of memory. Future studies should also take into account critical periods in which the diagnosis of trauma-related disorders are determined, such as acute stress disorder, which is being assessed within 1 month following an event, and PTSD, which is assessed within 3 months. The influence of specific psychological and psychiatric symptoms, such as dissociative and PTSD symptoms on the recall and consistency of memories of emotionally arousing events, so far has not been fully explained by experimental studies among non-victims. Even though in a number of these studies arousal of stress was generated in an artificial way, it is impossible, due to justifiable ethical limitations, to fully replicate traumatic stress in a laboratory (Terr, 1991). Therefore, studies should consider including both victims and non-victims in experimental studies assessing general information processing aspects as well as specific recall processes for traumatic events.

References

Aalsma, M. C., Zimet, G. D., Fortenberry, J. D., Blythe, M., & Orr, D. P. (2002). Reports of childhood sexual abuse by adolescents and young adults: Stability over time. *Journal of Sex Research*, *39*, 259–263.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bohannon, J. N. (1988). Flashbulb memories for the space shuttle disaster: A tale of two theories. *Cognition*, *29*, 179–196.

Bramsen, I., Dirkzwager, A. J. E., van Esch, S. C. M., & van der Ploeg, H. M. (2001). Consistency of self-reports of traumatic events in a population of Dutch peacekeepers: Reason for optimism. *Journal of Traumatic Stress*, *14*, 733–740.

Bremner, J. D., & Brett, E. (1997). Trauma-related dissociative states and long-term psychopathology in posttraumatic stress disorder. *Journal of Traumatic Stress*, *10*(37), 49.

Bremner, J. D., Krystal, J. H., Charney, D. S., & Southwick, S. M. (1996). Neural mechanisms on dissociative amnesia for childhood abuse: Relevance to the current controversy surrounding the "false memory syndrome". *American Journal of Psychiatry*, *153*, 71–82.

Brewer, N., Potter, R., Fisher, R. P., Bond, N., & Luszcz, M. A. (1999). Beliefs and data on the relationship between consistency and accuracy of eyewitness testimony. *Applied Cognitive Psychology*, *13*, 297–313.

Brewin, C. R., Dalgleish, T., & Joseph, S. (1996). A dual representation theory of posttraumatic stress disorder. *Psychological Review*, *103*, 670–686.
Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*, 73–99.
Cahill, L., & McGaugh, J. L. (1995). A novel demonstration of enhanced memory associated with emotional arousal. *Consciousness and Cognition*, *4*, 410–421.
Candel, I., Smeets, T., & Merckelbach, H. (2004). Accuracy, completeness and consistency of emotional memories. *American Journal of Psychology*, *117*, 595–609.
Christianson, S. A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin*, *112*, 284–309.
Christianson, S. A., & Engelberg, E. (1999). Memory and emotional consistency: The MS Estonia Ferry disaster. *Memory*, *7*, 471–482.
Conway, M. A., Anderson, S. J., Larsen, S. F., Donnely, C. M., McDaniel, M. A., McClelland, A. G. R., et al. (1994). The formation of flashbulb memories. *Memory and Cognition*, *22*, 326–343.
Curci, A., Luminet, O., Finkenauer, C., & Gisle, L. (2001). Flashbulb memories in social groups: A comparative test–retest study of the memory of French President Mitterand's death in a French and a Belgian group. *Memory*, *9*, 81–101.
Dohrenwend, B. S. (1973). Life events as stressors: A methodological inquiry. *Journal of Health and Social Behavior*, *14*, 167–175.
Fergusson, D. M., Horwood, L. J., & Woodward, L. J. (2000). The stability of child abuse reports: A longitudinal study of the reporting behaviour of young adults. *Psychological Medicine*, *30*, 529–544.
Fisher, R. P., & Cutler, B. L. (1995). The relation between consistency and accuracy of eyewitness testimony. In G. M. Davies, S. M. A. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (pp. 21–28). Berlin: De Gruyter.
Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G. S. Goodman, & B. L. Bottoms (Eds.), *Child victims, child witnesses: Understanding and improving testimony* (pp. 1–24). New York, NY: Guilford Press.
Foa, E. B., Molnar, C., & Cashman, L. (1995). Change in rape narratives during exposure therapy for posttraumatic-stress-disorder. *Journal of Traumatic Stress*, *8*, 675–690.
Ghetti, S., Goodman, G. S., Eisen, M. L., Qin, J., & Davis, S. L. (2002). Consistency in reports of sexual and physical abuse. *Child Abuse and Neglect*, *26*, 977–995.
Goodman, L. A., Thompson, K. M., Weinfurt, K., Corl, S., Acker, P., Mueser, K. T., et al. (1999). Reliability of reports of violent victimization and posttraumatic among men and women with serious mental illness. *Journal of Traumatic Stress*, *12*, 587–599.
Health Council of the Netherlands. (2004). *Disputed Memories*. Health Council of the Netherlands, The Hague, publication no. 2004/02.
Herlihy, J., Scragg, P., & Turner, S. (2002). Discrepancies in autobiographical memories—implications for the assessment of asylum seekers: Repeated interviews study. *British Medical Journal*, *324*, 324–327.
Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory and Cognition*, *18*, 496–506.
Hornstein, S. L., Brown, A. S., & Mulligan, N. W. (2003). Long-term flashbulb memory for learning of Princess Diana's death. *Memory*, *11*, 293–306.
Hyman, I. E., & Loftus, E. F. (1998). Errors in autobiographical memory. *Clinical Psychology Review*, *18*, 933–947.
Hyman, I. E., & Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language*, *35*, 101–117.
King, D. W., King, L. A., Erickson, D. J., Huang, M. T., Sharkansky, E. J., & Wolfe, J. (2000). Posttraumatic stress disorder and retrospectively reported stressor exposure: A longitudinal prediction model. *Journal of Abnormal Psychology*, *109*, 624–633.
Koopman, C., Classen, C., & Spiegel, D. (1994). Predictors of posttraumatic stress symptoms among survivors of the Oakland/Berkeley, Calif, firestorm. *American Journal of Psychiatry*, *151*, 888–894.
Krackow, E., & Lynn, S. J. (2003). Is there a touch in the game of twister? The effects of innocuous touch and suggestive questions on children's eyewitness memory. *Law and Human Behavior*, *27*, 589–604.
Krinsley, K. E., Gallagher, J. G., Weathers, F. W., Kutter, C. J., & Kaloupek, D. G. (2003). Consistency of retrospective reporting about exposure to traumatic events. *Journal of Traumatic Stress*, *16*(4), 399–409.
Lee, P. J., & Brown, N. R. (2003). Delay related changes in personal memories for September 11, 2001. *Applied Cognitive Psychology*, *17*, 1007–1015.
Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist*, *48*, 518–537.

Loftus, E. F. (2003). Make belief memories. *American Psychologist*, *58*, 867–873.

Loftus, E. F., & Ketcham, K. (1994). *The myth of repressed memory: False memories and allegations of sexual abuse*. New York: St. Martin's Press.

Marmar, C. R., Weiss, D. S., Schlenger, W. E., Fairbank, J. A., Jordan, B. K., Kulka, R. A., et al. (1994). Peritraumatic dissociation and posttraumatic stress in male Vietnam theatre veterans. *American Journal of Psychiatry*, *151*, 902–907.

McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism? *Journal of Experimental Psychology. General*, *117*, 171–181.

McNally, R. J. (2003). *Remembering trauma*. Cambridge, MA: The Belknap Press of Harvard University Press.

Mechanic, M. B., Resick, P. A., & Griffin, M. G. (1998). A comparison of normal forgetting, psychopathology, and information-processing models of reported amnesia for recent sexual trauma. *Journal of Consulting and Clinical Psychology*, *66*, 948–957.

Melchert, T. P., & Parker, R. L. (1997). Different forms of childhood abuse and memory. *Child Abuse and Neglect*, *25*, 1053–1068.

Morgan III, C. A., Hazlett, G., Doran, A., Garrett, S., Hoyt, G., Thomas, P., et al. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry*, *27*, 265–279.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about challenger. In E. Winograd, & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of* '*flashbulb*' *memories* (pp. 9–31). New York: Cambridge University Press.

Neisser, U., Winograd, E., Bergman, E. T., Schreiber, C. A., Palmer, S. E., & Weldon, M. S. (1996). Remembering the earthquake: Direct experience vs. hearing the news. *Memory*, 337–357.

Niles, B. L., Newman, E., Erwin, B. A., Fisher, L. M., Kaloupek, D. G., & Keane, T. M. (1999). Stability and fluctuation of veterans' reports of combat exposure. In L. M. Williams, & V. L. Banyard (Eds.), *Trauma and memory* (pp. 311–317). Thousand Oaks: Sage Publications.

Norris, F. H., & Kaniasty, K. (1992). Reliability of delayed self-reports in disaster research. *Journal of Traumatic Stress*, *5*, 575–588.

Pillemer, D. B. (1984). Flashbulb memories of the assassination attempt on President Reagan. *Cognition*, *16*, 63–80.

Porter, S., Spencer, L., & Birt, A. R. (2003). Blinded by emotion? Effect of the emotionality of a scene on susceptibility to false memories. *Canadian Journal of Behavioural Science*, *35*, 165–175.

Roemer, L., Litz, B. T., Orsillo, S. M., Ehlich, P. J., & Friedman, M. J. (1998). Increases in retrospective accounts of war-zone exposure over time: The role of PTSD symptom severity. *Journal of Traumatic Stress*, *11*, 597–605.

Rolls, E. T. (1990). A theory of emotion and its application to understanding the neural basis of emotion. *Cognition and Emotion*, *4*, 161–190.

Schacter, D. L. (1996). Searching for memory. The brain, the mind, and the past. New York: Basic Books.

Schmolck, H., Bufffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: Recollections of the O.J. Simpson trail verdict after 15 and 32 months. *Psychological Science*, *11*, 39–45.

Schwarz, E. D., Kowalski, J. M., & McNally, R. J. (1993). Malignant memories: Post-traumatic changes in memory in adults after a school shooting. *Journal of Traumatic Stress*, *6*, 545–553.

Smith, M. C., Bibi, U., & Sheard, D. E. (2003). Evidence for the differential impact of time emotion on personal and event memories for September 11, 2001. *Applied Psychology*, *17*, 1047–1055.

Southwick, S. M., Morgan, C. A., Nicolaou, A. L., & Charney, D. S. (1997). Consistency of memory for combat-related traumatic events in veterans of Operation Desert Storm. *American Journal of Psychiatry*, *154*, 173–177.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14*, 455–461.

Tekcan, A. I., Ece, B., Gulgoz, S., & Er, N. (2003). Autobiographical and event memory 9/11: Changes across one year. *Applied Cognitive Psychology*, *17*, 1057–1066.

Terr, L. C. (1991). Childhood traumas: An outline and overview. *American Journal of Psychiatry*, *148*, 10–19.

Van der Kolk, B. A., & Fisler, R. (1995). Dissociation and the fragmentary nature of traumatic memories: Overview and exploratory study. *Journal of Traumatic Stress*, *8*, 505–525.

Weaver, C. A. (1993). Do you need a "flash" to form a flashbulb memory? *Journal of Experimental Psychology. General*, *122*, 39–46.

Wessely, S., Unwin, C., Hotopf, M., Huli, L., Ismail, K., Nicolaou, V., et al. (2003). Stability of recall of military hazards over time. Evidence from the Persian Gulf War of 1991. *British Journal of Psychiatry*, *183*, 314–322.

Winningham, R. G., Hyman, I. E., & Dinnel, D. L. (2000). Flashbulb memories? The effects of when the initial memory report was obtained. *Memory*, *8*, 209–216.

Wyshak, G. (1994). The relation between change in reports of traumatic events and symptoms of psychiatric distress. *General Hospital Psychiatry*, *16*, 290–297.

Zoellner, L. A., Sacks, M. B., & Foa, E. B. (2001). Stability of emotions for traumatic memories in acute and chronic PTSD. *Behaviour Research and Therapy*, *39*, 697–711.

Zola, S. M. (1998). Memory, amnesia, and the issue of recovered memory: Neurobiological aspects. *Clinical Psychology Review*, *18*, 915–932.

Annu. Rev. Clin. Psychol. 2006. 2:469–98
doi: 10.1146/annurev.clinpsy.2.022305.095315
Copyright © 2006 by Annual Reviews. All rights reserved
First published online as a Review in Advance on January 16, 2006

RECOVERED MEMORIES

Elizabeth F. Loftus
Department of Psychology and Social Behavior, University of California, Irvine, California 92697-7085; email: eloftus@uci.edu

Deborah Davis
Department of Psychology, University of Nevada, Reno, Nevada 89557; email: debdavis@unr.nevada.edu

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Key Words repression, influence, false memory, therapy, childhood sexual abuse

■ **Abstract** The issues surrounding repressed, recovered, or false memories have sparked one of the greatest controversies in the mental health profession in the twentieth century. We review evidence concerning the existence of the repression and recovery of autobiographical memories of traumatic events and research on the development of false autobiographical memories, how specific therapeutic procedures can lead to false memories, and individual vulnerability to resisting false memories. These findings have implications for therapeutic practice, for forensic practice, for research and training in psychology, and for public policy.

CONTENTS

INTRODUCTION . . . 470
EVIDENCE FOR REPRESSION AND RECOVERY OF MEMORIES OF TRAUMA . . . 471
- Retrospective Studies . . . 471
- Prospective Studies . . . 473
- Case Histories . . . 473

EVIDENCE FOR THE EXISTENCE OF FALSE MEMORIES OF ABUSE . . . 476
- False Memories of Real-Life Trauma . . . 476
- Experiences of the “Retractors” . . . 477
- Laboratory Research on the Malleability of Autobiographical Memory . . . 478
- Summary . . . 480

THERAPEUTIC PROCESS AND FALSE MEMORIES OF ABUSE . . . 480
- A Priori Assumptions Regarding Abuse . . . 480
- Confirmation Biases and the Dangers of Specific Hypothesis Testing . . . 481
- Plausibility-Enhancing “Evidence” . . . 482
- Adopting and Confirming the Belief in Abuse . . . 483

THERAPY AND SOCIAL INFLUENCE . . . 489
- Individual Vulnerability to False Memories . . . 489

Ability to Resist 490
Motivation to Resist 492
RECOVERED MEMORIES AND PUBLIC POLICY 492

INTRODUCTION

On November 28, 1989, George Franklin was arrested for the murder of his daughter's childhood playmate—a murder that had allegedly occurred almost 20 years earlier (see MacLean 1993). The evidence against him? Nothing but the recently "recovered" memory of his now 29-year-old daughter, who claimed to have repressed her memory of having witnessed the murder two decades ago. The Franklin case was far from unique. Parents were being accused and convicted of other terrible crimes, primarily childhood sexual abuse, sometimes involving years of horrific abuse that was allegedly repressed in memory. Typically, these accusations arose on the basis of memories that adult children had "recovered" during psychotherapy. Some mental health professionals were even promoting the notion that numerous victims had experienced horrific satanic ritual abuse about which they were harboring repressed memories (Rogers 1992, Wright 1994).

But scholarly analyses (e.g., Holmes 1990) were revealing that there was little in the way of support for widespread assumptions among therapists and in popular folklore that traumatic memories are particularly likely, relative to nontraumatic memories, to be "repressed" and later recovered intact through techniques such as hypnosis, guided imagery, and other suggestive therapeutic procedures. Analyses of how false memories could develop in the therapeutic setting soon followed (e.g., Lindsay & Read 1994, Loftus 1993, Loftus & Ketcham 1994, Ofshe & Watters 1994, Tavris 1993) and sparked a heated response from the therapeutic community (e.g., Alpert 1995, Terr 1994, Whitfield 1995) and alleged survivors of sexual abuse and their supporters that marked the beginning of a controversy that has been among the most vitriolic and emotionally charged in the history of psychology. This debate, known as the memory wars, has been referred to as "psychology's most fiercely contested ground" (Crews 2004). But underlying a very practical side of the debate—centered on real-life concerns for victims of either true abuse or of false allegations—is another debate surrounding the very nature of memory and how it works: whether memory might work differently for traumatic versus more ordinary events, and whether it might be distorted or confabulated as a result of therapeutic procedures commonly employed by some therapists.

On the one side were primarily practicing therapists who argued that there was and still is overwhelming support for the psychoanalytic notion of repressed memories (sometimes referred to by other terms, such as dissociated memory or traumatic amnesia; e.g., Brown et al. 1998). Traumatic memories such as those of sexual abuse were viewed as fundamentally different from more ordinary memories because they tend to be encoded in ways that render them inaccessible in everyday life. Moreover, suggestive memory recovery procedures and therapeutic

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

interactions were viewed as necessary to break through the barrier of repression and bring memories into conscious awareness, which was in turn viewed as necessary for the patient to improve. Therapists who supported such a position tended to view the incidence of "repressed" memories of abuse as relatively high, and therefore the frequency with which such memories were "recovered" in therapy as unsurprising. Many such therapists also viewed even extremely bizarre claims such as satanic ritual abuse among children and adults as credible. And finally, they argued that false memories for such events were particularly unlikely.

On the other side were the clinical, social, and cognitive researchers who had long studied the fallibility and suggestibility of memory. Beginning with the assumption that, if anything, memories for trauma are stronger than are those for ordinary events, these researchers viewed traumatic experiences as unlikely to be repressed and as subject to the same sources of distortion and confabulation as memories of other kinds of experiences. These scholars and scientists found no compelling evidence that people massively repress sexual abuse and then reliably recover the memories later (see Piper at al. 2000, Pope et al. 1998). In one survey, 79% said that there was no support for the statement, "Traumatic experiences can be repressed for many years and then recovered," or that the data were inconclusive (Kassin et al. 2001). Moreover, the repression skeptics worried that suggestive procedures used by some psychotherapists to try to extract allegedly buried trauma memories (such as direct suggestion that the patient was probably abused, guided imagery, hypnosis, age regression, or dream analysis) could lead to false memories—even such seemingly improbable false memories as those of satanic abuse.

In the following sections, we review evidence concerning the existence of recovered memories. We focus upon the controversial sense of this term, which involves memories of abuse that are "recovered" during suggestive psychotherapy. We also review evidence for the existence and mechanisms of creation of false memories and discuss how these processes apply in therapy.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

EVIDENCE FOR REPRESSION AND RECOVERY OF MEMORIES OF TRAUMA

Most fundamentally, to demonstrate that memories can be repressed and later recovered, at least three things must be verified: (*a*) that the abuse did take place, (*b*) that it was forgotten and inaccessible for some period of time, and (*c*) that it was later remembered (see, e.g., Pope & Hudson 1995). Studies used to support repression generally do not meet these criteria.

Retrospective Studies

In a retrospective memory study, individuals are interviewed today and asked whether they were abused in the past as well as other questions assessing the

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

continuity of memory over time, such as whether they ever forgot the abuse. In scores of such studies, some individuals will claim that they were abused and that there was a time when they forgot the abuse (see, e.g., Briere & Conte 1993, Melchert 1996), but the inherent common flaws of these studies render them virtually uninterpretable (see, e.g., Brenneis 2000; Kihlstrom 1998, 2005; McNally 2003b, 2004).

Perhaps the most fundamental flaws are lack of validation of the abuse and lack of assessment of the conditions under which the "memory" was retrieved. Though some studies have attempted to validate reported abuse, criteria for validation are often suspect, relying, for example, on participant reports that they had verified the abuse (see, e.g., Herman & Schatzow 1987) or the outcome of legal proceedings (see, e.g., Burgess et al. 1995). Often, researchers fail to report rates of verification separately for those who always remembered the abuse versus those who report periods of amnesia or repression, leaving open the question of the rate of verified "recovered" memories (see, e.g., van der Kolk & Fisler 1995). Still others mix self-reports of verification from patients with apparently more objective verification by police or therapists, but without clear delineation of the frequency of each (see, e.g., Kluft 1997). Nonetheless, although few instances of abuse have been conclusively verified, a number of studies have reported instances of apparent verification of once unrecalled abuse (see McNally 2003b).

Other studies have not attempted to validate the abuse at all, and have used persons whose memories were recovered in suspect circumstances without comparison to those whose memories were recovered more naturally. Incredibly, one of the most influential studies of this type recruited subjects through a national network of therapists specializing in treatment of abuse survivors (Briere & Conte 1993). Of those claiming past abuse, 59% reported experiencing a time when they could not remember the abuse. Given a variety of methodological issues (see McNally 2003a), the claim of past nonmemory is uninterpretable. In other studies, patients claiming recovered memory of abuse appear to have undergone questionable procedures such as hypnosis or guided imagination. Indeed, fully two thirds of those reporting periods of amnesia in Roe & Schwartz's (1996) study reported first recovery of the memory during hypnosis. In another study, participants who never remembered abuse but who had joined incest survivor groups to help them remember were classified as having repressed abuse (Herman & Schatzow 1987).

A second general set of concerns surrounds the interpretation of forgetting and recovery. Episodes of abuse may not be experienced as traumatic or even labeled as abuse at the time, and hence forgetting cannot be regarded as traumatic repression. Indeed, one study found that women who reported having forgotten their abuse rated it as having been less upsetting when it occurred than those who had never forgotten (Loftus et al. 1994). In other studies, participants' own interpretations of failures to remember abuse have included failure to understand the experience as abusive until later and deliberate attempts not to think about it. Many reported that they could have remembered if they tried or if they had been reminded or

asked about it. However, since all such reports are either guesses about whether the person could remember or subjective assessments of reasons for failure to remember rather than responses to actual attempts to remember, they are difficult to interpret.

Further, it is well documented that one can fail to remember that one previously remembered the abuse. For some apparently verified instances of recovered memories, it has been shown that the person actually did remember during the alleged amnestic period, but later forgot previously remembering and talking about the abuse to others (see Brenneis 2000). In one study, women who claimed that they had undergone periods of forgetting their abuse said later in the same interview that they had never forgotten (Fivush & Edwards 2004). Finally, studies of memory for real-life traumata of all sorts suffer from additional problems such as misinterpreting general difficulty with everyday memory as reflecting repression of a specific traumatic event or failure to rule out injury and organic causes of amnesia.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Prospective Studies

In a prospective memory study, individuals with a record of abuse or other trauma in the past are later interviewed to see what they remember. One well-known study (Williams 1994) involved women who had reported sexual abuse that had occurred when they were aged 10 months to 12 years old. In interviews some 17 years later, 38% did not mention the abuse incident. These results are frequently used to support the notion that a significant percentage of women repress their memories of sexual abuse. But numerous critics have questioned this interpretation (Kihlstrom 2005, Loftus et al. 1994, McNally 2003b), noting the myriad reasons other than repression that could cause participants to fail to mention the abuse. Some had experiences as children when they were so young (under age 2) that childhood amnesia would lead us to expect no memory for the abuse. Even if they did remember it, some may have simply not wanted to report the abuse to an interviewer for reasons such as embarrassment or lack of rapport (Della Femina et al. 1990). Moreover, by design, participants were not asked directly about the abuse, and had they been asked, may well have been able to report it. Finally, normal forgetting occurs for all sorts of events, even ones that would have been rather upsetting or traumatic.

Nearly a decade later, Goodman et al. (2003) published a conceptual "replication" of the Williams study. Participants had been involved in a study of the effects of criminal prosecutions on sex abuse victims when they were ages 3–17 and were interviewed by the authors three times 10–16 years later. By the final interview, only 8% did not report the abuse. Although their study has been criticized for using a "prosecution" sample, its results do cast doubt on the claim that large percentages of women have repressed their memories and have no awareness of real past abuse.

Case Histories

A third source of evidence offered to support the claim of massive repression is anecdotal cases (sometimes called "anecdata"), where a therapist writes an account

of a case history along with an interpretation that the patient has repressed and later reliably recovered the memory. Other reports involve a set of case histories analyzed and possibly "verified" by the researcher. But one problem with case histories is that the therapist/author is typically the only person who has access to the "data," which are often subjective and not convincingly subjected to objective external verification. Few instances exist where the selectivity has been scrutinized, but one clear example can be found in the 1997 case history of "Jane Doe" (Corwin & Olafson 1997), who was videotaped in 1984 recounting specifics of sexual abuse allegedly committed by her mother. Eleven years later, when Jane was 17, she was videotaped again. This time, she at first did not remember the abuse, and then she did. The therapists published an account of Jane Doe's life, her allegedly repressed/recovered memory, and the case was cited as verified (e.g., Gleaves et al. 2004).

Loftus & Guyer (2002a,b) used public records and newspaper clippings and eventually located Doe's family. From court documents and other information, they learned that the case was not even remotely a proven case of repressed memory. In fact, much newly discovered evidence cast doubt on whether abuse had occurred at all and pointed to the very real possibility that the abuse narrative had been planted in Jane's mind by individuals who wished to remove her from her mother. This scrutinized case is a cautionary tale that raises questions about the role of case histories in medicine, science, and mental health. Case histories can be compelling, but they are bounded by the motivations and interpretations of the storyteller.

Problems with motivational biases characterize many case histories involving litigated events. Claims of repression are sometimes necessary in order to file suit after delays that would normally exceed statutes of limitations. Nor can out-of-court settlements be taken as proof of claims of abuse. Innocents sometimes settle to avoid legal and emotional costs and risks of litigation. Such issues limit the weight of evidence provided by allegedly "corroborated" cases of recovered memories of trauma involved in legal proceedings.

Notwithstanding these problems, there are a large number of case reports, some with better verification than others. Schooler et al. (1997) give the impression that they discovered several case histories for which the necessary three-pronged evidence specified by Pope & Hudson (1995) was obtained: The abuse did occur, it was forgotten for some period of time, and it was later remembered. Unfortunately, even in some of these the "documentation" was merely the subject's word. Moreover, these cases provided no evidence for repression per se. In fact, most memory recoveries reportedly occurred outside of therapy, and some subjects reported that they had forgotten the event even though it was later discovered that they had discussed it with others during the period of "amnesia."

Brenneis (2000) analyzed "verified" case histories provided by Schooler et al. (1997) and others, noting that the most adequately verified accounts tend to share several features: (*a*) the memories were typically not (with one exception) recovered in the context of therapy, and in all cases the moment of recovery was

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

unrelated to any therapeutic activity, (*b*) the memories were triggered by external events that "reminded" the person of the original abuse, (*c*) once prompted, the memories "completely unwound instantly" (Schooler et al. 1997, p. 271) and required no interpretation or deciphering, (*d*) the memories were mostly for single, not repeated, events, (*e*) assailants were primarily nonfamily members, and (*f*) the events dated from age 9 and older. Finally, information from friends and relatives often revealed that despite the person's claim of amnesia, the event had been discussed with others during the period of claimed amnesia. These features stand in contrast to the typical pattern for memory recovery in therapy, which involves very effortful and gradual recovery through extremely suggestive therapeutic processes and memory recovery procedures. When memories emerge in this context, they tend to begin as vague and lacking detail but to unfold and become more vivid and elaborate over time. Further, Brenneis (2000) noted that "this constellation of features—multiple events over long periods of time, beginning in early childhood, involving bodily penetration, and enacted by male family members—is seldom if ever found among verified recovered memory cases" (p. 75).

Wagenaar & Crombag (2005) provide a devastating analysis of the case of JR—a man Schooler interviewed some nine years after he allegedly recovered memories of abuse by a priest—in which they note that the JR narrative contains innumerable unproven and often suspicious assumptions. Schooler says that JR recovered his memory without a therapist, but JR was in therapy at the time. Schooler implies that JR had no motive, but JR actually did start a lawsuit. Schooler was impressed by a second individual who apparently also implied wrongdoing by the priest, but that second individual was not independent, as he came forth after learning about JR. As Wagenaar & Crombag (2005) note, case histories like that of JR do not meet two of the most important criteria for empirical research: public control and replicability. It is virtually impossible for readers to check most of the details of JR's story because of the anonymity. We cannot critically question those who provided information nor can we have access to case information. The story is simply hearsay. (See Wagenaar & Crombag 2005, Chapter 5, for an excellent discussion of when case histories can be used successfully to illustrate something important.)

In sum, there is little support for the notion that trauma is commonly banished out of awareness and later reliably recovered by processes beyond ordinary forgetting and remembering. Although there have been some apparent instances of verified lost and recovered memories (see, e.g., Brenneis 2000, Gleaves et al. 2004), it is not clear how much scrutiny has been applied, and crucial questions of the base rate at which such verified instances occur and how it depends upon the circumstances of retrieval (through specific procedures during therapy versus as the result of a retrieval cue that occurred in everyday life) remain essentially unanswered. Yet over the past couple of decades, many persons have reported having experienced massive abuse that was repressed and recovered, which raises the question of whether some or all such "memories" might be false.

EVIDENCE FOR THE EXISTENCE OF FALSE MEMORIES OF ABUSE

Recovered memory therapy advocates note that no controlled experimental evidence confirms that false memories of traumatic events can be implanted. Indeed, ethics restrict experimentation on the impact of memory recovery procedures on recovery of true traumatic memories or implantation of false ones, making laboratory evidence specific to trauma and abuse for both camps rare. Given such limitations, critics of "recovered memories" offer several kinds of evidence demonstrating that false memories for mundane and relatively traumatic autobiographical events can be implanted.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

False Memories of Real-Life Trauma

For most claims of massive repression and recovery, there is little confirming or disconfirming evidence. But some "memories" can be shown to be factually, psychologically, geographically, or biologically impossible. As with case histories of alleged lost and recovered memories, those of false memory for trauma must also meet stringent criteria of proof: both that the person did have memories for the trauma in question and that the event actually did not happen. Indeed, many such case histories are available for abuse- and nonabuse-related trauma (see reviews in McNally 2003b; Schacter 1996, 2001), such as being kidnapped and held hostage (something that happened to classmates instead of oneself), being gang-raped by Satanists (although one's hymen remains intact), enduring the surgical removal of one's clitoris (contradicted by the patient's gynecologist), witnessing the sacrificial killing of persons later found alive, and even for having committed heinous crimes such as murders or sexual abuse (Henkel & Coffman 2004, Kassin 2006, Wright 1994).

While some claims have been proven factually inaccurate, others are simply impossible, such as detailed narrative memories of events occurring in the first days to six months of life (Arnold 1994, Usher & Neisser 1993). Although children do encode and remember such events during early life, these memories tend to be eventually lost, as adult reports of childhood memories rarely address events that occurred earlier than age 3 (see Bauer 2006).

Among the most frequently reported impossible false memories of trauma are those of abduction by space aliens. Although approximately 17% of Americans believe that aliens have abducted humans (and presumably returned them alive), we assume that memories of such events are clearly false. Nevertheless, large numbers of patients have reported memories of alien abductions that have largely developed in therapy and under hypnosis or while the patient was subject to other methods used in recovered memory therapy (see Mack 1994, McNally 2003b).

A large category of false memories concerns various forms of satanic ritual abuse reported by patients in recovered memory therapy. Alleged acts included gang rape; sacrifice of babies and other ritual murders; consumption of blood and

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

human waste; forced pregnancies or abortions; dancing, chanting, and other Satan worship activities; and brutal torture designed to cause victims to forget all they endured (see Loftus & Ketcham 1994, Noblitt & Perskin 2000, Ofshe & Watters 1994, Scott 2001). Such cases had so permeated our culture that as of 1991, the American Bar Association reported that 25% of prosecuting attorneys had handled cases involving satanic abuse (Qin et al. 1998). In a 1996 survey of clinicians from the American Psychological Association (Bottoms et al. 1996), 13% had cases involving children and 11% had cases involving adults with claims of satanic ritual abuse (with some therapists reporting more than 100 cases). None had obtained convincing verification of the abuse, nor have subsequent attempts to examine the validity of such claims found reliable evidence (see, e.g., La Fontaine 1998, Lanning 1991, Weir & Wheatscroft 1995). One recent survey found that satanic ritual abuse was reported in 19% of more than 1700 cases involving families who reported false allegations of abuse against a family member (McHugh et al. 2004). Across studies, 95% to 100% of patients had no recollection of abuse prior to therapy (McNally 2003b). Despite vigorous protests from recovered memory therapists (Terr 1994, Whitfield 1995), there are cogent reasons to believe that almost all such claims are false (see McNally 2003a,b).

In summary, although there is disagreement regarding the plausibility of some of the above instances of memory for trauma, there can be no doubt that "memories" for factually false as well as impossible or at least highly improbable horrific traumatic events were developed, particularly among persons subjected to suggestive memory recovery procedures. Some have viewed the prevalence of memories for satanic ritual abuse as the strongest evidence of real-life false memories of trauma (e.g., Ofshe & Watters 1994, Ross 1995).

Experiences of the "Retractors"

In the l990s, hundreds of individuals who had been persuaded that they had repressed and recovered memories of abuse began to realize their memories were false, and many sued their former therapists for planting false memories. Scores were studied by psychologists trying to gain insight into the processes by which the patients developed and later retracted their beliefs (see, e.g., De Rivera 1997; Lief & Fetkewicz 1995; Nelson & Simpson 1994; Ost et al. 2001, 2002). These studies revealed that the modal retractor first sought therapy for depression and then recovered "memories" of abuse during therapy, but later came to believe the "memories" were actually products of therapeutic suggestion. More than 90% recovered their memories in therapy; in one study (Lief & Fetkewicz 1995), 48% recovered memories of satanic ritual abuse and 38% recovered memories of witnessing murder. The vast majority had undergone suggestive procedures such as hypnosis. Retractors reported substantial pressure to recover memories, and noted that when they expressed doubts in their new memories, they were told that such doubt is common but not a sign of inaccuracy. Most reported that outside pressure played little to no role in their retractions (see, e.g., Ost et al. 2002). Instead, the retractions appeared to be based primarily on the experiential qualities of the

"memories" themselves. In essence, the memories did not seem truly "real," being either too clear and vivid (and increasing in vividness over time, rather than declining, as do most memories) or too vague and dreamlike.

Although many retractors have provided some insights into the processes they went through, these are case histories and are therefore subject to many of the limitations of case histories identified earlier (see entire *Psychological Inquiry,* 1997, Vol. 8, #4 for commentaries on the meaning of retractor reports). Not surprisingly, therapists have found these accounts unconvincing, arguing—ironically—that retractors are easily swayed by social pressure or that they are motivated by the potential of lawsuits against their therapists (see, e.g., Blume 1995, Brown et al. 1998). Notwithstanding such criticisms, there is little doubt that at least some people have developed clearly false memories that they later recognize as such.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Laboratory Research on the Malleability of Autobiographical Memory

EVIDENCE THAT FALSE AUTOBIOGRAPHICAL MEMORIES CAN BE CREATED Research in the past several decades has shown that it is relatively easy to change details of memories for previously experienced events (see reviews in Davis & Loftus 2006; Loftus 2005; entire *Handbook of Eyewitness Psychology*, *Vols. I and II*), but it is also possible to implant entirely false autobiographical memories, even of highly implausible or even impossible events. Using strong forms of suggestion in a paradigm known as the "familial informant false narrative procedure" or simply the "lost-in-the-mall" technique (Lindsay et al. 2004, Loftus 1993, Loftus & Pickrell 1995), people have been led to believe that, as children, they were lost in a shopping mall for an extended time, had an accident at a family wedding, were the victim of a vicious animal attack, nearly drowned and had to be rescued by a lifeguard, etc. These false memories can be planted by telling individuals that their relatives have provided the information and then suggestively interviewing the individuals to try to elicit memory reports.

Across many studies utilizing the lost-in-the-mall procedure, an average of approximately 30% of subjects have developed partial or complete false beliefs or memories (Lindsay et al. 2004), although these rates can vary from 0% with relatively implausible events (receiving a rectal enema; Pezdek et al. 1997) to more than 50% for more mundane events (a ride in a hot air balloon; Wade et al. 2002). Techniques such as those involving guided imagination (e.g., Libby 2003), suggestive dream interpretation, or exposure to doctored photographs have also led subjects to believe falsely that they experienced events in their distant and even in their recent past (Loftus 2003). Some develop false memories right away, whereas others begin with little memory but after several suggestive interviews begin to recall false events in great detail (Ost et al. 2005).

Implanted memories might be viewed as fleeting and unimportant. But even false beliefs implanted in laboratory studies have repercussions affecting later thoughts, behaviors, and intentions. In several studies, false memories of having

gotten sick after eating particular foods as children led to avoidance of the foods as adults (Bernstein et al. 2005).

Several criticisms have been lodged regarding the autobiographical memory implantation research. Most prominent are (*a*) that we often cannot know for sure (despite familial reports otherwise) that those who develop "false" memories did *not* experience the target event, and (*b*) that target events in such experiments are less traumatic and more plausible than those commonly "recovered" in therapy. In response to such criticisms, researchers have endeavored to implant both impossible and highly traumatic or "implausible" autobiographical memories.

To address the criticism of verification, for example, Braun et al. (2002) led subjects to believe the impossible event that they had met Bugs Bunny (a Warner Brothers character) at a Disney resort (after exposure to fake Disney ads featuring Bugs Bunny). These authors found that a single fake ad led 16% of subjects to claim they had met Bugs. Even higher rates of false belief were obtained by Braun-LaTour et al. (2004), and ads containing pictures of Bugs produced more false memories than those with only verbal mention of him. The criticism involving the degree of trauma has been more widely addressed. Although researchers have not attempted to plant memories of abuse, they have attempted to plant memories for relatively unpleasant, and in some cases fairly traumatic, events such as hospitalizations, medical procedures, near drowning, or vicious animal attacks. Finally, researchers have planted highly implausible memories for both mundane (e.g., rubbing chalk on one's head or kissing a plastic frog) and strange and dramatic (witnessing demonic possession as a child) events (Loftus 2003).

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Whereas it may be more difficult (or even sometimes impossible) to plant implausible memories in many circumstances, this difficulty can be overcome by changing the degree of implausibility first (Hyman & Loftus 2002). Just such a strategy was demonstrated by Mazzoni et al. (2001), who first exposed some participants to material designed to enhance the plausibility of demon possession and later attempted to plant memories of having personally witnessed such an event. Those exposed to the plausibility-enhancing manipulation later reported greater likelihood that they had personally witnessed demon possession.

A final criticism of the memory implantation studies is that although false beliefs that one has experienced the event may be planted in many subjects, detailed false memories, particularly for more implausible or more traumatic events, have been planted in relatively few individuals. However, one might argue that given the ease with which false memories can be planted in a short period of time, the rate of false memory development in long-term therapy might be substantially more.

RECONSTRUCTIVE INFLUENCES OF BELIEFS, GOALS, AND SELF-VIEWS ON AUTOBIOGRAPHICAL MEMORY Studies concerned with the retrospective bias have shown that reports of our own past attitudes or behaviors are biased by current self-views, goals, and beliefs (and vice versa; Dawes 1991, Wilson & Ross 2003). Similarly, recollections of one's own behavior tend to change to conform to newly acquired information about how one *should* behave, so that we believe we behaved in a

more consistent, sensible, or desirable way than we actually did. Given this reconstructive bias in autobiographical memory, what direct and indirect reconstructive effects might the sheer belief that one has been abused exert?

Even simply considering the possibility that one has been abused might exert such effects, but some patients may come to adopt a highly elaborate personal identity as an abuse survivor. Kihlstrom (1998) coined the term "false memory syndrome," which he described as "a condition in which a person's identity and interpersonal relationships are centered around a memory of traumatic experience which is objectively false but in which the person strongly believes . . . the syndrome may be diagnosed when the memory is so deeply ingrained that it orients the person's entire personality and lifestyle" (p. 16). Though some have objected to the use of the term "syndrome" (Pope 1996), clearly many patients do experience their status as abuse victims as the central or even the primary aspect of their identity. Self-definition as an abuse survivor is likely to exert reconstructive influence on autobiographical memory, which in turn is presumed to serve essentially as the basis of one's identity. Pressures exist to conform memories to current identity as well as to conform current identity to memories [see entire issue 11, no. 2, of *Memory* (2003) on autobiographical memory; Tafarodi et al. 2003 on self-esteem and memory].

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Summary

There is little doubt that abuse can be forgotten and later remembered, although ordinary forgetting and remembering seem more than adequate to account for this. Nor can there be doubt that false memories of abuse or other trauma can be confabulated. Doubt remains, however, regarding the base rates at which each occurs and the circumstances and persons for which each is most likely. Why do some people, and not others, develop false memories? How does this depend upon the social context in which memories are triggered? And fundamentally, how—if at all—are encoding, storage, and retrieval for traumatic or highly emotional content different? When traumatic material is inaccessible for a period of time, what are the processes responsible? Although progress has been made with respect to each question, much remains to be learned.

THERAPEUTIC PROCESS AND FALSE MEMORIES OF ABUSE

Bearing in mind that false memories can be created, we consider below how this occurs and what might be done to minimize the likelihood.

A Priori Assumptions Regarding Abuse

As most accounts of the "recovered memory" controversy have documented, a dramatic increase in awareness of sexual abuse began in the 1980s, accompanied by widespread media coverage of abuse and recovered memories as well as a number

of popular books on the topic. This situation has essentially "primed" the notion of abuse, including repressed memories of abuse, in the general population, and elevated awareness among therapists already acquainted with concepts of trauma and repression. Among the effects of priming particular "schemas" are selective attention to relevant information, biased interpretation of relevant information, and constructive and reconstructive memory processes that generally consist of confabulation of schema-consistent (but false) memories and distortion of memories of past events toward consistency with currently activated schemas (Davis & Follette 2001, Davis & Loftus 2006, Kunda 1999). A patient that has been exposed to accounts of repressed abuse, and with abuse fully primed in her mind, may present with a pre-existing idea that she may have been abused. Likewise, the therapist may expect a high rate of repressed abuse among patients, or particular types of patients, thereby setting in motion a biased assessment process—often followed with vigorous suggestive efforts to test and verify the abuse hypothesis.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Confirmation Biases and the Dangers of Specific Hypothesis Testing

Confirmatory biases are likely to manifest in initial interview and assessment processes, possibly in both patient and therapist. A patient already considering or convinced of abuse may offer information she perceives as relevant to abuse, perhaps even arguing its significance in terms of abuse. A therapist who perceives abuse as prevalent or likely may inquire about symptoms and facts seen as diagnostic, and if the patient has already brought up the possibility of abuse, rather than engage in a systematic differential diagnosis to examine and rule out alternatives, the therapist may jump directly to the conclusion that the patient in fact was abused. Such "premature cognitive commitment" (Pope & Brown 1996) is among common errors of clinical judgment that some clinicians warn of. Rather than conducting objective hypothesis testing, the therapist may embark upon a quest to discover abuse-consistent evidence (including reports of consistent information and memories from the patient), discounting any inconsistent evidence and doggedly pursuing the presumption of abuse notwithstanding protests and inconsistent evidence from the patient.

In recognition of these processes, critics of recovered memory therapy point to scientific literature documenting the dangers of the confirmation bias (the tendency to affirm the diagnosis one is considering) in clinical diagnosis and judgment (see, e.g., Garb 1998). Indeed, even clinicians largely in the recovered memory camp of the debate have warned of the dangers of the confirmation bias in diagnosis (e.g., Pope & Brown 1996). This bias has been documented even under circumstances in which clinicians are asked to review an unknown patient's file to evaluate whether the patient suffers a particular disorder without any contact with the patient, without any reason to favor the designated diagnosis, and in an effort to provide an unbiased assessment (e.g., Copeland & Snyder 1995; see Kassin & Gudjonsson 2004 and Meissner & Kassin 2004 for discussion of confirmation biases among interrogators).

The tendency to confirm clinical hypotheses results in part from biases in interpretation. However, consistent with the vast literature documenting expectancy confirmation processes in social interaction (e.g., Kirsch 1999), biased interviewing procedures also contribute in that questions are typically asked in a manner that tends to elicit apparently confirming information from the patient or interviewee (see, e.g., Fazio et al. 1981, Snyder 1984, Snyder & Thomsen 1988). The therapist may also suggest activities to the patient that would likewise tend to elicit (apparently) confirmatory information, such as reading survivor literature, completing various "homework" activities focused on abuse, or participating in survivor groups. These activities may serve to elicit more apparently abuse-consistent information from the patient. Furthermore, some therapists instruct patients to avoid exposure to contradictory information and those who might provide it rather than objectively testing the abuse hypothesis by seeking, and encouraging patients to seek, informative abuse-inconsistent information as well.

THE ROLE OF MOTIVATED COGNITION Many, if not most, patients enter therapy in search of an explanation for their problems. This very need for explanation may render patients vulnerable to accepting seemingly plausible potential causes. Believing a patient was abused, a therapist might directly suggest this hypothesis, as well as provide apparently confirmatory "evidence," such as the extent to which the patient's symptoms conform to those thought to be associated with abuse. Particularly when combined with other "evidence" gleaned from survivor literature, survivor groups, media, and other sources, the abuse hypothesis may seem a compelling explanation to patients who fail to realize their symptoms may be better explained in other ways. The very existence of a potential explanation may motivate some patients to prematurely seize and freeze on the abuse hypothesis, thereby causing them to engage in a strongly biased search for and interpretation of information and to defend against doubt. Some therapists may also be driven by motives—ranging from fierce victim advocacy to potential financial rewards of prolonged therapy—that tend to encourage confirmatory strategies.

Plausibility-Enhancing "Evidence"

Biased hypothesis-testing strategies are likely to elicit apparently confirmatory evidence. The therapist faced with any set of evidence (even were it not selective and biased) can succumb to a number of heuristic and other fallacies of reasoning that result in a tendency to confirm the abuse hypothesis (see Garb 1998), as can patients. Schematic processing can result in selective attention to abuse-relevant information, disregard for abuse-irrelevant information, and interpretative biases toward consistency with the abuse hypothesis, including explaining away apparently inconsistent information. This includes retrospective biases of interpretation such as the "hindsight" bias (Fischhoff 1975), whereby the past is interpreted as consistent with current knowledge.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Therapists may also fall prey to the "representativeness" heuristic (Kahneman & Tversky 1972), assuming that if a patient's symptoms fit those viewed as consistent with abuse, the patient must have been abused. This, of course, is the message of much of the survivor literature (e.g., Bass & Davis 1988). Such logic is fallacious, in that "If Abuse, then Symptom" does not logically imply "If Symptom, then Abuse." Abuse-related symptoms can result from abuse as well as from many other causes.

The problem is further complicated by a questionable assumption regarding the true association of various symptoms with abuse. "The phenomenon of recovered memories has been greatly confounded by the assumption made by some clinicians that repetitive behavioral patterns, special sensory reactivities, and unbidden ideation in the form of flashbacks or nightmares necessarily reflect implicit memory for unremembered events. . .. In short, without a corresponding explicit memory, the existence of past trauma cannot be conclusively inferred from any repetitive behaviors or intrusive ideation" (Brenneis 2000, p. 67). Notwithstanding these and other misunderstandings of what constitutes (or does not constitute) valid indicators of abuse, therapist and patient may each offer "evidence" to the other, in support of their own hypotheses. If the patient has not yet adopted the abuse hypothesis, the therapist may proceed with a number of leading and even coercive procedures designed to elicit confirmation, including persistent persuasion and efforts to elicit consistent information and failure to believe inconsistent information at all or to interpret it as actually inconsistent with repressed abuse.

PLAUSIBILITY, BELIEF, AND MEMORY Just as the memory implantation research reveals, information that serves to render previously implausible information subjectively more plausible can smooth the way for the development of false memories. Suggestive influences inside and outside of therapy are likely to enhance the plausibility of abuse, notwithstanding the absence of memories. Persuasive information can come from the media, survivor literature, survivor groups, therapist suggestions, and other sources. When apparently authoritative sources state unequivocally that particular symptoms are pathognomic of abuse, the plausibility that a person suffering from such symptoms could have been abused is enhanced.

Adopting and Confirming the Belief in Abuse

If a patient comes to believe that she may have been abused, efforts by both therapist and patient may ensue to confirm and defend the new survivor identity. These may include memory recovery procedures in and outside of therapy, participation in survivor groups, solicitation of consistent information from family, witnesses, and others—all with significant potential both to bias construction of historical narratives and to lead to confabulation of false memories.

MEMORY RECOVERY PROCEDURES Whether the patient has yet adopted the survivor identity or not, the therapist may suggest a variety of memory recovery procedures, both in and outside of therapy, such as hypnosis, age regression, dream

interpretation, guided abuse-related imagery, use of photographs to trigger memories, instructions to work at remembering (including through journaling or other homework), and interpreting physical symptoms as implicit memories (see, e.g., Poole et al. 1995). These and other procedures, and their potential to cause false memories, have been extensively discussed (e.g., Brainerd & Reyna 2005, Loftus & Ketcham 1994, McNally 2003a,b).

HYPNOSIS Prominent in the memory recovery arsenal is hypnosis. Interestingly, however, as reviewed by Mazzoni et al. (2006), both memory-enhancing and memory-distorting functions of hypnosis have been recognized and employed by therapists beginning with Freud, Janet, and other early psychotherapists. Therapists have used hypnotic memory retrieval in two opposite ways, without any apparent awareness of the implications that one use had for the other. That is, while they viewed hypnosis as an excellent memory recovery tool to recover accurate memories, they also deliberately used hypnosis as a suggestive memory confabulation tool to create "healing" positive pseudomemories to replace "true" traumatic memories previously "recovered" through hypnosis. Indeed, modern research has verified both functions. Hypnosis can lead to retrieval of greater numbers of or increased detail for accurate memories as well as to greater production of false memories. Persons under hypnosis have developed a number of bizarre or impossible memories, such as memories of satanic ritual abuse (described above), impossible memories from infancy (Spanos et al. 1999), memories from previous lives, sexual abuse during past lives (Stevenson 1994, Spanos et al. 1991), and even memories from one's own future (see reviews by Kihlstrom 1997, Mazzoni et al. 2006, McNally 2003b). Hypnotic age-regression, a procedure commonly employed by recovered memory therapists (see Poole et al. 1995), is subject to the same distortions as hypnotic memories of recent events (see Nash 1987). As Mazzoni et al. (2006) point out, if hypnosis is not a reliable means of recovering memories of recent events, there is no reason to expect it to be more effective for memories of the distant past or childhood. Nor is there any reason to expect that it can facilitate retrieval of memories beyond the veil of infantile amnesia.

GUIDED IMAGERY Therapists commonly employ various imaging activities in their sessions and in homework assignments for clients. Guided imagery, whereby a client is asked to actively try to imagine and create images of past events, is viewed by researchers as dangerous in that these vivid and elaborate images may later become confused with memories. Indeed, memory researchers have shown that imagining events tends to inflate perceptions of the likelihood they had actually occurred—an effect generally referred to as "imagination inflation" (Garry et al. 1996), and a host of studies have shown that active imagination/visualization of events, objects, or persons can lead to false memories of having actually seen, performed, or experienced them. Imagination has produced false memories for simple perceptions, such as having seen or heard objects or sounds, as well as for more complex recent personal actions (such as having said or done something,

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

both mundane and bizarre) and distant autobiographical memories for a range of events (see reviews by Davis & Loftus 2006; Johnson et al. 1993; Mazzoni & Memon 2003; Schacter 1996, 2001; Thomas & Loftus 2002). Even paraphrasing event descriptions or explaining how an event might have happened can produce inflation (see, e.g., Sharman et al. 2004, 2005). Techniques emphasizing imagination not only can generate false memories, but also can inflate confidence in those memories (see Arbuthnott et al. 2001 on imagery, Mazzoni et al. 2006 on hypnosis, Spanos et al. 1999 on age regression).

Presumably, like other implantation techniques, imagination works in a three-stage process whereby people first come to believe an event is plausible, next come to believe the event did actually occur, and finally reinterpret their narratives and images of the event as actual memories (Mazzoni et al. 2001). Imagery and imagination may contribute to all levels of this process. Imagery is crucial to plausibility and hence persuasion [see Green & Brock (2002) for evidence that narratives are persuasive to the extent they evoke imagery of their contents]. Research from the source-monitoring tradition has shown that images can be confused with real memories, particularly when they have many of the subjective characteristics of real memories. Johnson et al. (1988) reasoned that when real memories are vague and lacking in vivid detail, as when memories are from the distant past or were never encoded richly in the first place, it is easier to confuse imagined and real events.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

DREAM INTERPRETATION Another common tool of psychotherapy, dream interpretation, can lead some to develop false memories. Mazzoni and her colleagues (see, e.g., Mazzoni et al. 1999) studied direct suggestion in the form of a psychologist's bogus interpretation of dreams. For some participants, these bogus interpretations (i.e., the same interpretation given to all subjects, regardless of the dream reported, and with no reason to believe the interpretation applied to each subject) led to false memories for mildly traumatic suggested events. Dreams themselves may also be confused with actual experiences in some cases (Kemp et al. 2003).

FAMILY PHOTOS As an apparently sensible "context reinstatement" procedure, therapists may recommend that patients use family photos to trigger lost memories. But autobiographical memory can be distorted through exposure to photographs (Lindsay et al. 2004). Brainerd & Reyna (2005) suggest that this apparently sensible procedure may backfire because the photos are employed in the context of delayed repeated attempts to recall after previous attempts have failed.

REPEATED RECALL When "memories" are difficult to retrieve, a variety of memory recovery techniques may be used repeatedly over a long period. Even in the absence of other suggestive procedures and influences, evidence exists to show that (*a*) as time passes, both spontaneous false memories and false reports in response to suggestion increase (even for short delays involving hours, days, or weeks), (*b*) repeated attempts to recall increase the yield of false as well as true information,

(*c*) information recalled in later attempts is proportionately more likely to be false, (*d*) the previous two findings are particularly likely when there are long delays between repeated-recall tests and the index experience, and (*e*) these patterns obtain for autobiographical memory as well as for laboratory tasks (see review by Brainerd & Reyna 2005).

ABUSE-RELATED IMAGES AND THEIR MISATTRIBUTION AS MEMORIES As images appear through various efforts to retrieve memories they may combine with related inferences to develop into full narratives of abusive events. The person may then misunderstand the source of these images and narratives [an error of source monitoring (Johnson et al. 1993)], misattributing them to true memories of the abusive events. But how does this process of misattribution occur? Figure 1 depicts a model of this process.

Modern theories of the development of false memories (Brainerd & Reyna 2005) assume that remembering consists of subjectively experienced internal representations of an event combined with judgment criteria for determining whether these representations correspond to a previously experienced index event. Internal representations can be either verbatim traces (i.e., of the exact surface form and other specific information, much like seeing with the mind's eye) or gist traces (i.e., of the essential semantic meaning or generalized physical form of objects and events, for example, "going to the movie" versus specific visual images of the people, objects, and actions involved). Judgment criterion can also vary in specificity. At one extreme, the person may require vivid, elaborate verbatim memory traces (i.e., the ability to fully picture the event in the mind's eye or ear) in order to label the internal representations as a memory. At the other, the person may label even fuzzy, unelaborated fragmented gist traces as memories. Generally, the stronger the person's verbatim and gist traces, and the weaker (or more gist-based) the judgment criterion, the more likely a particular representation is to be judged as a memory. Therefore, therapeutic and nontherapeutic factors that influence the strength of either form of trace or of the nature of the judgment criterion can contribute to source-monitoring errors. To understand how this would occur, we must first address what contributes to the strength of both verbatim and gist traces.

Verbatim traces consist of vivid internal images of the index event. A number of factors influence the strength of such images, including the depth of original encoding, personal memorial abilities, and the passage of time. Unfortunately, vivid images may also be created independent of actual experience, such as through the various guided imagery procedures commonly practiced in recovered memory therapy. These highly elaborated internally generated images would possess the apparent verbatim trace representations of actual events and therefore pass even the stringent verbatim-match judgment criterion for assessing validity. Gist traces, on the other hand, may be created through both overlapping and dissimilar processes. Like verbatim representations, gist traces may be strengthened by depth of encoding or personal memorial abilities. However, whereas verbatim memory becomes relatively weaker over time, gist traces become relatively more dominant.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Repeated Cueing of Abuse Inferences
Rich Abuse-Related Memory Support
Abuse Narratives
Abuse Images
Imaginative Ability
Dissociative Tendencies
Psychopathology
Familiarity
Failures of Source Monitoring
Impaired Self-Regulation
Memory Judgment Criteria
Impaired Cognitive Resources
Belief in Abuse
Judgment Threshold
Social Influence
Schematic Processing
Priming Abuse

Figure 1 Causes of source monitoring failures/memory misattributions

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Also, whereas verbatim memory, by definition, can be contributed to only by the target event, gist memory may come to include overall meanings that represent the combined semantic understanding of the nature of the event. Hence, gist traces may be affected by one's overall event-relevant knowledge structure, such as when memory is affected by schematic processing. Thus, like verbatim traces, gist traces can be artificially created or strengthened through various imaging procedures, but can also be strengthened through other activities serving to develop abuse-related schemas, beliefs, and personal narratives.

When confronted with the need to judge a particular internal trace, the more existing memory support that can be retrieved to "verify" an experience (and the less contradictory information), the more likely the person will label the trace as a true memory. Such support can include the quality of verbatim and gist traces as well as related images, narratives within which the target event resides, semantic knowledge, beliefs, and other relevant traces. In other words, the more information of any kind that is available to increase the subjective likelihood that an internal trace represents an actual event, the more likely it will be judged as a memory. As depicted in Figure 1, repeated cueing of relevant information (such as through priming abuse, developing abuse narratives, discussing abuse) tends to create rich abuse-related memory support. One may also use familiarity as the judgment criteria, whereby things that simply feel sufficiently familiar are judged as memories. The feeling of familiarity may be increased through the same forms of memory "support" discussed above. Familiarity, or gist-based criteria generally, are more lax judgment criteria and are empirically associated with more reported false memories.

Given this background, we can now summarize how therapeutic and associated processes directly and indirectly promote source-monitoring errors. As shown in Figure 1, the social influences that keep abuse primed, promote schematic processing, and support belief in abuse exert persuasive influence on both verbatim and gist memory traces as well as on the judgment criteria applied to those traces. Above, we discussed the way in which verbatim and gist traces can be developed and/or elaborated through therapeutic procedures. Essentially, various imaging activities in and outside of therapy can produce vivid artificial images that resemble verbatim memories (particularly among those with substantial imaginative ability) and also contribute to gist traces. This abuse-related memory support is added to by the various suggestive procedures that develop rich abuse-related narratives, personal abuse-related identities, and abuse-supportive beliefs. These apparent verbatim and gist traces and semantic memory support enhance the plausibility of, and confer a sense of familiarity upon, abuse "memories."

As depicted in Figure 1, therapeutic and nontherapeutic processes also directly and indirectly affect the judgment criteria applied to target internal representations. Generally, the impaired cognitive processes that tend to be characteristic of recovered memory patients (see below) are known to be associated with source-monitoring errors, tendencies toward automatic (e.g., schematic) rather than controlled processing, susceptibility to social influence, and use of gist-based

judgment criteria (see Brainerd & Reyna 2005, Davis & Loftus 2005, Davis & O'Donohue 2004 for reviews). People who are already susceptible are then subject to other influences that affect judgment criteria.

Therapists often directly argue for lax criteria, suggesting that no "memory" should be doubted, and that even bodily sensations and other nonmental reactions should be interpreted as memories. Essentially, familiarity, gist-based, and other unique criteria are promoted. Further, all abuse-supportive beliefs (particularly those supporting specific episodic memories) can directly lower the criteria applied to internal representations. That is, the more the person simply believes an event occurred, the less episode-specific verbatim or gist trace is needed to support a conclusion of verity. The patient's own faith in the effectiveness of memory-recovery procedures can similarly directly lower the criteria. Once interpreted as real, of course, such "memories" serve to reinforce the belief in abuse and solidify the survivor identity, perhaps thereby encouraging development of additional false memories.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

THERAPY AND SOCIAL INFLUENCE

Social influence is, in a sense, the point of therapy. The patient hopes to find solutions to problems with the help of a presumably knowledgeable authority. If the relationship develops as desired, the therapist will possess the primary attributes known to promote social influence: likeability, credibility, and power (control over desired resources) (see Pratkanis & Aronson 2001). The patient is likely to feel strong emotional attachment to, great respect for, and even dependence on, the therapist, feelings which would render her more susceptible to believing information and adopting behavioral suggestions such as joining survivor groups, reading survivor literature, or engaging in memory-recovery activities at home. These same feelings would render the person more susceptible to biasing therapist suggestions in and outside of specific activities such as hypnosis or guided imagery. But influence in therapy is actually bidirectional. The abuse hypothesis neither is always first suggested nor is most strongly promoted by the therapist. Patients' own biases and preconceptions developed outside of therapy can exert a strong influence upon their own hypothesis-testing activities as well as those of their therapists.

Individual Vulnerability to False Memories

Given that therapeutic procedures possess considerable potential to create false beliefs and memories in some, it is important to consider which patients might be especially susceptible to social influence or memory distortion. Here, we focus upon the issue of social influence. (See Brainerd & Reyna 2005 for a discussion of individual differences in memory distortions.) One might expect patients to be generally more susceptible to influence, and particularly from therapists from

whom they are seeking answers. As Brainerd & Reyna (2005) note, "Confused and uncertain people are looking for information that will shed light in dark corners, and they may believe that it can be found in therapy. To find answers, however, they must adopt attitudes of openness, exploration, and discovery with respect to the events of their lives, and they must trust in the wisdom and experience of their therapists. Obviously, these latter characteristics are not commensurate with a narrow, reality-based perspective on memory. Rather, the perspective is a much broader one that searches for answers and solutions in the events of one's life, which is not precisely the same thing as searching for autobiographical facts" (pp. 386–387).

Notwithstanding elevated vulnerability within the general patient population, there will be notable individual differences. To understand who might be most vulnerable to social influence, it is first necessary to understand what is needed to resist social influence. Davis & O'Donohue (2004) provide an extensive analysis of the processes and abilities that promote or undermine resistance to influence with regard to resisting powerful influence in police interrogations (see Knowles & Linn 2004 on resistance to influence in other contexts). Sources of resistance that are most relevant here reduce to two general factors: (*a*) the ability to resist (consisting of the ability to understand and evaluate relevant information and the ability to exert one's will to refuse to comply) and (*b*) the motivation to resist.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Ability to Resist

The immediate antecedents of the ability to resist are the abilities to (*a*) analyze relevant information and (*b*) exert one's will in a particular direction. These two abilities are in turn affected by both chronic and acute individual differences.

CHRONIC AND ACUTE INTELLECTUAL ABILITIES To understand and evaluate relevant information adequately, one must possess adequate relevant knowledge and have adequate chronic intellectual abilities. To see flaws in abuse-relevant suggestions, for example, one would be aided both by clear existing autobiographical knowledge and memories and by domain-specific knowledge regarding abuse, therapy, memory, and other relevant facts. Given this knowledge, one would also need to possess the intellectual abilities to analyze incoming suggestions in light of this knowledge and to have the acute capacity to bring relevant knowledge and abilities to bear. The latter requires the intact self-regulatory resources needed to control cognitive processes, including attention (to attend to relevant incoming information, to access relevant information from long-term memory, and to exclude distracting information) and working memory (to hold relevant information in mind while assessing its implications).

Self-regulation is central to these abilities. Although many may think of self-regulation as relevant to control of overt behaviors, substantial research has demonstrated that depletion of self-regulatory resources impairs intellectual performance of all sorts, including resistance to persuasion. Furthermore, self-regulatory capacity varies between individuals and can be easily depleted through such means as

previous exertions of effort or will, physical depletion, emotional stress, difficult social interactions, and resisting temptation (Baumeister & Vohs 2004; see Davis & O'Donohue 2004 for self-regulation and interrogation). Recovered-memory therapy patients are likely to suffer depleted self-regulatory resources due to psychopathology, emotional distress, poor physical condition secondary to distress (and such factors as sleep difficulties), difficult interactions with family, problems at work, and other factors. This depletion would impair intellectual functioning and ability to analyze relevant information as well as the ability to exert one's will.

Substantial evidence supports the importance of intellectual and self-regulatory capacities for suggestibility. Such factors as IQ, age-related intellectual abilities, physical and emotional status, and acute and chronic self-regulatory capacities have been shown to influence suggestibility and persuasion, including development of false memories and other memory distortions (see reviews by Bruck & Melnyk 2004, Davis & Loftus 2005, Davis & O'Donohue 2004, Gudjonsson 2003, Kassin & Gudjonsson 2004). The inclination to use these capacities is also relevant. Some are more inclined to carefully analyze incoming information (referred to as systematic or central route processing*)* whereas others rely on heuristic cues such as source attractiveness or credibility to assess likely accuracy (referred to as heuristic or peripheral route processing; see Cacioppo et al. 1996). The latter are particularly prone to rely on those seen as credible authorities without careful consideration of the basis of their opinions.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

THE ABILITY TO EXERT ONE'S WILL The ability to exert one's will in the direction of resistance to suggestion is also crucial. The most coercive form of recovered memory therapy will greatly resemble a coercive interrogation. That is, the suggestion will be relentless. Just as interrogators relentlessly pursue a particular preconceived version of the target crime and the suspect's guilt, the therapist will make many varied and repeated suggestions regarding abuse, will not recognize arguments or evidence provided by the patient against the abuse hypothesis as valid, and will reinterpret apparently inconsistent evidence as actually irrelevant or supportive and all consistent evidence as confirming evidence. In other words, in the strongest form of recovered memory therapy, the patient will be faced with consistent powerful external forces toward acceptance and compliance, thereby requiring strong and intact self-regulatory capacity to resist.

RESISTANCE IN THE CONTEXT OF STRONG PRIMING AND DIRECTED ATTENTION Even if a person suffers no impairments of self-regulation, knowledge, or cognitive resources, access to relevant information may be impaired by factors that selectively direct attention to confirming information. If the patient is in a situation where abuse has been suggested, where therapeutic interactions and procedures focus on abuse-related content, where suggestions to read survivor literature or to join survivor discussion groups have been adopted, and there has been a focus on efforts to remember abuse, the full situation will surely selectively direct attention to abuse-consistent content. The therapist is likely to ask questions that will tend to elicit

abuse-consistent content. Hence, even if the patient possesses knowledge (personal historical knowledge or other) that would contradict the abuse hypothesis, abuse and relevant consistent information may be so strongly primed and cognitively available that the contradictory information is effectively kept out of awareness.

THE ROLE OF DIRECTED INTERPRETATION The same therapeutic activities that selectively direct attention to abuse-related content will also tend to control the interpretation of information that is considered. As noted by researchers studying influence in the interrogation room, interrogators essentially persuade suspects to confess (including to confess falsely) by controlling both the information that is attended to and considered and the interpretation given to that information (e.g., Davis & O'Donohue 2004, Ofshe & Leo 1997).

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Motivation to Resist

One may see flaws in a suggestion and be able to resist but nevertheless have no desire to do so. Motivation can also affect the inclination to carefully scrutinize suggestions and therefore the initial detection of any flaws. Motivation to resist may be directly diffused by attraction and emotional attachment to the source, by his assumed expertise and authority, and by perceived helplessness in the face of his or her power or excessive dependence on the source (see Pratkanis & Aronson 2001 for reviews regarding source characteristics and influence). Individual differences in tendencies to trust authorities, to view them as credible or powerful, or to depend upon them are associated with greater deference and compliance (see Gudjonnson 2003, Kassin & Gudjonnson 2004).

Further, as thoroughly documented in the persuasion literature, resistance is fueled by motivated commitment to current beliefs and by pre-existing inconsistent attitudes and beliefs. In the context of therapy, motivation to resist may be undermined by the need to understand the source of one's problems (as discussed above) and by abuse-consistent beliefs developed prior to or in the process of therapy. Finally, external influences such as friends, other accusers, and family may directly encourage or discourage abuse-related assumptions. Generally, however, the therapeutic context is likely to undermine resistance in many patients, often progressively so as more abuse-consistent beliefs develop throughout the process. This is perhaps most likely to occur in dispositionally suggestible people who are also most confused, most motivated to find an explanation for their problems, and who possess pre-existing beliefs concerning themselves, repression, and the accused that would support developing abuse narratives.

RECOVERED MEMORIES AND PUBLIC POLICY

The controversies surrounding allegedly repressed memories have created unfortunate tensions among professionals. But out of this process have come useful discussions by clinicians and nonclinicians writing together about changes in practices that would minimize the problems that false accusations can bring to all

involved. We found many instances where clinicians and scientists have provided useful advice that stemmed from the research we have reviewed. But we give one of the last words to a clinician, Sarnoff (2001), who has worried that a focus on "believing the victim" (p. 169) has essentially eliminated healthy skepticism as a quality to be encouraged in all who encounter questionable claims. We particularly resonate to her concerns, having learned this truth about memory: Just because a "memory" report is detailed, just because a person expresses it with confidence and emotion, does not mean that the event actually happened. Keeping that truth in mind may help to minimize harm to the many victims of the "memory wars": the patients who were misdiagnosed, the innocents who were falsely accused, the good therapists who suffered damaged reputations, and the genuine victims of abuse whose experiences were trivialized by the dubious claims of others.

Many clinicians will be aware that prominent professional organizations such as the American Medical Association, the American Psychological Association, the Australian Psychological Society, and the British Royal College of Psychiatrists have issued strong warnings against practices associated with recovered memory therapy (see the collected statements in Brainerd & Reyna 2005). It is our hope that this review will provide greater understanding of the basis of these recommendations and of the processes that contribute to the development of false memories of all kinds. In light of the biasing potential inherent to clinical diagnosis and therapy, it is essential for clinicians of the future to be well trained in the dangers of subjectivity and suggestive procedures such as those covered here.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

ACKNOWLEDGMENTS

Gratitude is expressed to the Grawemeyer Prize in Psychology for funding used to carry out this research. DD has testified as an expert witness on memory but has not done so specifically with respect to recovered memories. EL occasionally testifies as an expert witness in trials where recovered memories are an issue.

The *Annual Review of Clinical Psychology* is online at http://clinpsy.annualreviews.org

LITERATURE CITED

Alpert JL. 1995. *Sexual Abuse Recalled*. Northvale, NJ: Jason Aronson

Arbuthnott KD, Arbuthnott DW, Rossiter L. 2001. Guided imagery and memory: implications for psychotherapists. *J. Couns. Psychol.* 48:123–32

Arnold R. 1994. *My Lives*. New York: Ballantine

Bass E, Davis L. 1988. *The Courage to Heal*. New York: Harper & Row

Bauer PJ. 2006. *Remembering the Times of Our Lives: Memory in Infancy and Beyond*. Mahwah, NJ: Erlbaum. In press

Baumeister RF, Vohs KD, eds. 2004. *Handbook of Self-Regulation: Research, Theory, and Applications*. New York: Guilford

Bernstein DM, Laney C, Morris EK, Loftus EF. 2005. False beliefs about fattening foods can have healthy consequences. *Proc. Natl. Acad. Sci. USA* 102:13724–31

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Blume ES. 1995. *The Ownership of Truth*. New York: Wiley

Bottoms BL, Shaver PR, Goodman GS. 1996. An analysis of ritualistic and religion-related child abuse allegations. *Law Hum. Behav.* 20:1–34

Braun KA, Ellis R, Loftus EF. 2002. Make my memory: how advertising can change our memories of the past. *Psychol. Mark.* 19:1–23

Braun-LaTour KA, LaTour MS, Pickrell J, Loftus EF. 2004. How (and when) advertising can influence memory for consumer experience. *J. Advert*. 33:7–25

Brainerd CJ, Reyna VF. 2005. *The Science of False Memory*. New York: Oxford Univ. Press

Brenneis CB. 2000. Evaluating the evidence: Can we find authenticated recovered memory? *J. Am. Psychoanal. Assoc.* 17:61–77

Briere J, Conte J. 1993. Self-reported amnesia for abuse in adults molested as children. *J. Trauma. Stress* 6:21–31

Brown D, Scheflin AW, Hammond DC. 1998. *Memory, Trauma, Treatment, and the Law*. New York: Norton

Bruck M, Melnyk L. 2004. Individual differences in children's suggestibility: a review and synthesis. *Appl. Cogn. Psychol.* 18:947–96

Burgess AW, Hartman CR, Baker T. 1995. Memory presentations of childhood sexual abuse. *J. Psychosoc. Nurs.* 33:9–16

Cacioppo JT, Petty RE, Feinstein J, Jarvis B. 1996. Dispositional differences in cognitive motivation: the life and times of individuals low versus high in need for cognition. *Psychol. Bull.* 119:197–253

Clancy S. 2005. *Abducted: How People Come to Believe They Were Kidnapped By Aliens*. Cambridge, MA: Harvard Univ. Press

Copeland J, Snyder M. 1995. When counselors confirm: a functional analysis. *Personal. Soc. Psychol. Bull.* 21:1210–20

Corwin DL, Olafson E. 1997. Videotaped discovery of a reportedly unrecallable memory of child sexual abuse: comparison with a childhood interview videotaped 11 years before. *Child Maltreat.* 2:91–112

Crews FC. 2004. The trauma trap. *NY Rev. Books* 51(4)

Davis D, Follette WC. 2001. Foibles of witness memory for traumatic/high-profile events. *J. Air Law Commerce* 66:1421–549

Davis D, Loftus EF. 2005. Age and functioning in the legal system: victims, witnesses and jurors. In *Handbook of Human Factors in Litigation*, ed. YI Noy, W Karwowski, pp. 11–1–11–53. New York: CRC Press

Davis D, Loftus EF. 2006. Internal and external sources of distortion in adult witness memory. In *Handbook of Eyewitness Memory (Vol. 1)*: *Memory for Events*, ed. MP Toglia, JD Read, DR Ross, RCL Lindsay. Mahwah, NJ: Erlbaum. In press

Davis D, Loftus EF, Follette WC. 2001. How, when, and whether to use informed consent for recovered memory therapy. *J. Am. Acad. Psychiatry Law* 29:148–59

Davis D, O'Donohue WT. 2004. The road to perdition: extreme influence tactics in the interrogation room. In *Handbook of Forensic Psychology*, ed. WT O'Donohue, E Levensky, pp. 897–996. New York: Elsevier

Dawes RM. 1991. Biases of retrospection. *Issues Child Abuse Accusations* 1:25–28

Della Femina D, Yeager C, Lewis DO. 1990. Child abuse: adolescent records vs. adult recall. *Child Abuse Negl.* 14:227–31

De Rivera J. 1997. The construction of false memory syndrome: the experience of retractors. *Psychol. Inq.* 8:271–92

Fazio RH, Effrein EA, Falender VJ. 1981. Self-perceptions following social interaction. *J. Personal. Soc. Psychol.* 41:232–42

Fischhoff B. 1975. Hindsight does not equal foresight: the effect of outcome knowledge on judgment under uncertainty. *J. Exp. Psychol. Hum. Percept. Perform.* 1:288–99

Fivush R, Edwards VJ. 2004. Remembering and forgetting childhood sexual abuse. *J. Child Sex. Abuse* 13:1–19

Ganaway GK. 1989. Historical versus narrative truth: clarifying the role of exogenous trauma

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

in the etiology of MPD and its variants. *Dissociation* 2:205–20

Garb HN. 1998. Clinical judgment. In *Studying the Clinician: Judgment Research and Psychological Assessment*, ed. HN Garb, pp. 173–206. Washington, DC: Am. Psychol. Assoc.

Garry M, Manning CG, Loftus EF. 1996. Imagination inflation: imagining a childhood event inflates confidence that it occurred. *Psychon. Bull. Rev.* 3:208–14

Gleaves DH, Smith SM, Butler LD, Speigel D. 2004. False and recovered memories in the laboratory and clinic: a review of experimental and clinical evidence. *Clin. Psychol. Sci. Pract.* 11:3–28

Goodman GS, Ghetti S, Quas JA, Edelstein RS, Alexander KW, et al. 2003. A prospective study of memory for child sexual abuse: new findings relevant to the repressed-memory controversy. *Psychol. Sci.* 14:113–18

Green MC, Brock TC. 2002. In the mind's eye: transportation-imagery model of narrative persuasion. In *Narrative Impact: Social and Cognitive Foundations*, ed. MC Green, JJ Strange, TC Brock, pp. 315–41. Mahwah, NJ: Erlbaum

Gudjonsson GH. 2003. *The Psychology of Interrogations and Confessions: A Handbook*. West Sussex, UK: Wiley

Henkel LA, Coffman KJ. 2004. Memory distortions in coerced false confessions: a source monitoring framework analysis. *Appl. Cogn. Psychol.* 18:567–88

Herman JL, Schatzow E. 1987. Recovery and verification of memories of childhood sexual trauma. *Psychoanal. Psychol.* 4:1–14

Holmes D. 1990. The evidence for repression: an examination of sixty years of research. In *Repression and Dissociation*, ed. J Singer, pp. 85–102. Chicago: Univ. Chicago Press

Hyman IE, Loftus EF. 2002. False childhood memories and eyewitness memory errors. In *Memory and Suggestibility in the Forensic Interview*, ed. ML Eisen, JA Quas, GS Goodman, pp. 63–84. Mahwah, NJ: Erlbaum

Johnson MK, Foley MA, Suengas AG, Raye CL. 1988. Phenomenal characteristics of memories for perceived and imagined autobiographical events. *J. Exp. Psychol. Gen.* 117:371–76

Johnson MK, Hashtroudi S, Lindsay DS. 1993. Source monitoring. *Psychol. Bull.* 114:3–28

Kahneman D, Tversky A. 1972. Subjective probability: a judgment of representativeness. *Cogn. Psychol.* 3:430–54

Kassin SM. 2006. Coerced internalized false confession. In *Handbook of Eyewitness Psychology (Vol. 1)*: *Memory for Events*, ed. MP Toglia, JD Read, DR Ross, RCL Lindsay. Mahwah, NJ: Erlbaum. In press

Kassin SM, Gudjonsson GH. 2004. The psychology of confessions. *Psychol. Sci. Public Int.* 5:33–67

Kassin SM, Tubb VA, Hosch HM, Memon A. 2001. On the "general acceptance" of eyewitness testimony research. *Am. Psychol.* 56:405–16

Kemp S, Burt CDB, Sheen M. 2003. Remembering dreamt and actual experiences. *Appl. Cogn. Psychol.* 17:577–91

Kihlstrom JF. 1997. Hypnosis, memory and amnesia. *Philos. Trans. R. Soc. Biol. Sci.* 352:1727–32

Kihlstrom JF. 1998. Exhumed memory. In *Truth in Memory*, ed. KM McConkey, pp. 3–31. New York: Guilford

Kihlstrom JF. 2005. Dissociative disorders. *Annu. Rev. Clin. Psychol.* 1:227–53

Kirsch I. 1999. *How Expectancies Shape Experiences*. Washington, DC: Am. Psychol. Assoc.

Kluft RP. 1997. The argument for the reality of delayed recall of trauma. In *Trauma and Memory: Clinical and Legal Controversies*, ed. PS Appelbaum, LA Uyehara, MR Elin, pp. 25–27. New York: Oxford Univ. Press

Knowles ES, Linn JA. 2004. Approach-avoidance model of persuasion: alpha and omega strategies for change. In *Resistance and Persuasion*, ed. ES Knowles, JA Linn, pp. 117–48. Mahwah, NJ: Erlbaum

Kunda Z. 1999. *Social Cognition: Making Sense of People*. Cambridge, MA: MIT Press

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

La Fontaine JS. 1998. *Speak of the Devil: Tales of Satanic Abuse in Contemporary England*. Cambridge: Cambridge Univ. Press

Lanning KV. 1991. Ritual abuse: a law enforcement view or perspective. *Child Abuse Negl.* 15:171–73

Libby LK. 2003. Imagery perspective and source monitoring in imagination inflation. *Mem. Cogn.* 7:1072–81

Lief HI, Fetkewicz J. 1995. Retractors of false memories: The evolution of pseudomemories. *J. Psychiatry Law* 23:411–35

Lindsay DS, Hagen L, Read JD, Wade KA, Garry M. 2004. True photographs and false memories. *Psychol. Sci.* 15:149–54

Lindsay DS, Read JD. 1994. Psychotherapy and memories of childhood sexual abuse: a cognitive perspective. *Appl. Cogn. Psychol.* 8:281–338

Loftus EF. 1993. The reality of repressed memories. *Am. Psychol.* 48:518–37

Loftus EF. 2003. Make-believe memories. *Am. Psychol.* 58:867–73

Loftus EF. 2005. Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn. Mem.* 12:361–66

Loftus EF, Garry M, Feldman J. 1994. Forgetting sexual trauma. *J. Consult. Clin. Psychol.* 62:1177–81

Loftus EF, Guyer M. 2002a. Who abused Jane Doe? The hazards of the single case history: part I. *Skeptical Inquirer* 26:24–32

Loftus EF, Guyer MJ. 2002b. Who abused Jane Doe? Part II. *Skeptical Inquirer* 26:37–40

Loftus EF, Ketcham K. 1994. *The Myth of Repressed Memory*. New York: St. Martin's

Loftus EF, Pickrell JE. 1995. The formation of false memories. *Psychiatr. Ann.* 25:720–25

Loftus EF, Polonsky S, Fullilove MT. 1994. Memories of childhood sexual abuse: remembering and repressing. *Psychol. Women Q.* 18:67–84

Mack JE. 1994. *Abduction: Human Encounters With Aliens*. New York: Ballantine. Rev. ed.

MacLean HN. 1993. *Once Upon a Time: A True Story of Memory, Murder, and the Law*. New York: Harper Collins

Mazzoni GAL, Loftus EF, Kirsch I. 2001. Changing beliefs about implausible autobiographical events: a little plausibility goes a long way. *J. Exp. Psychol. Appl.* 7:51–59

Mazzoni GAL, Loftus EF, Seitz A, Lynn SJ. 1999. Creating a new childhood: changing beliefs and memories through dream interpretation. *Appl. Cogn. Psychol.* 13:125–44

Mazzoni GAL, Lynn SJ, Kirsch I. 2006. Using hypnosis in eyewitness memory: past and current issues. In *Handbook of Eyewitness Memory (Vol. 1)*: *Memory for Events*, ed. MP Toglia, JD Read, DR Ross, RCL Lindsay. Mahwah, NJ: Erlbaum. In press

Mazzoni GAL, Memon A. 2003. Imagination can create false autobiographical memories. *Psychol. Sci.* 14:186–88

McNally RJ. 2003a. Progress and controversy in the study of posttraumatic stress disorder. *Annu. Rev. Psychol.* 54:229–52

McNally RJ. 2003b. *Remembering Trauma.* Cambridge, MA: Harvard Univ. Press

McNally RJ. 2004. Is traumatic amnesia nothing but psychiatric folklore? *Cogn. Behav. Ther.* 33:97–101

McNally RJ, Ristuccia CS, Perlman CA. 2005. Forgetting of trauma cues in adults reporting continuous or recovered memories of childhood sexual abuse. *Psychol. Sci.* 16:336–40

Meissner CA, Kassin SM. 2004. "You're guilty, so just confess!" Cognitive and behavioral confirmation biases in the interrogation room. In *Interrogations, Confessions and Entrapment*, ed. GD Lassiter, pp. 85–106. New York: Kluwer Acad.

Melchert TP. 1996. Childhood memory and a history of different forms of abuse. *Prof. Psychol. Res. Pract.* 27:438–46

Nash MR. 1987. What, if anything, is regressed about hypnotic age regression? A review of the empirical literature. *Psychol. Bull.* 102:42–52

Nelson EL, Simpson P. 1994. First glimpse: an initial examination of subjects who have rejected their recovered visualizations as false memories. *Issues Child Abuse Accusations* 6:123–33

Noblitt JR, Perskin PS. 2000. *Cult and Ritual*

Abuse: Its History, Anthropology, and Recent Discovery in Contemporary America. Westport, CT: Praeger. Rev. ed.
Ofshe RJ, Leo RA. 1997. The decision to confess falsely: rational choice and irrational action. *Denver Univ. Law Rev.* 74:979–1122
Ofshe R, Watters E. 1994. *Making Monsters: False Memories, Psychotherapy, and Sexual Hysteria*. New York: Scribner's
Ost J. 2003. Seeking the middle ground in the "memory wars." *Br. J. Psychol.* 94:125–39
Ost J, Costall A, Bull R. 2001. False confessions and false memories: a model for understanding retractors' experiences. *J. Forensic Psychiatry* 12:549–79
Ost J, Costall A, Bull R. 2002. A perfect symmetry? A study of retractor' experiences of making and then repudiating claims of early sexual abuse. *Psychol. Crime Law* 8:155–81
Ost J, Foster S, Costall A, Bull R. 2005. False reports in appropriate interviews. *Memory* 7:700–10
Pezdek K, Finger K, Hodge D. 1997. Planting false childhood memories: the role of event plausibility. *Psychol. Sci.* 8:437–41
Piper A Jr, Pope HG Jr, Borowiecki JJI. 2000. Custer's last stand: Brown, Scheflin, and Whitfield's latest attempt to salvage "dissociative amnesia." *J. Psychiatry Law* 28:149–213
Poole DA, Lindsay DS, Memon A, Bull R. 1995. Psychotherapy and the recovery of memories of childhood sexual abuse: U.S. and British practitioners' opinions, practices, and experiences. *J. Consult. Clin. Psychol.* 63:426–37
Pope HG Jr, Hudson JI. 1995. Can individuals "repress" memory of childhood sexual abuse? An examination of the evidence. *Psychiatr. Ann.* 25:715–19
Pope HG Jr, Hudson JI, Bodkin JA, Oliva P. 1998. Questionable validity of "dissociative amnesia" in trauma victims. *Br. J. Psychiatry* 172:210–15
Pope KS. 1996. Memory, abuse, and science. *Am. Psychol.* 51:957–74
Pope KS. 1998. Pseudoscience, cross-examination, and scientific evidence in the recovered memory controversy. *Psychol. Public Policy Law* 4:1160–81
Pope KS, Brown LS. 1996. *Recovered Memories of Abuse*. Washington, DC: Am. Psychol. Assoc.
Pratkanis AR, Aronson E. 2001. *Age of Propaganda*. New York: Freeman
Qin J, Goodman GS, Bottoms BL, Shaver PR. 1998. Repressed memories of ritualistic and religion-related child abuse. In *Truth in Memory*, ed. SJ Lynn, KM McConkey, pp. 163–89. New York: Guilford
Roe CM, Schwartz MF. 1996. Characteristics of previously forgotten memories of sexual abuse: a descriptive study. *J. Psychiatry Law* 24:189–206
Rogers ML, ed. 1992. Satanic ritual abuse. *J. Psychol. Theol.*: *Special Issue*. 20(3):entire issue
Ross CA. 1995. *Satanic Ritual Abuse: Principles of Treatment*. Toronto: Univ. Toronto Press
Sarnoff SK. 2001. *Sanctified Snake Oil: The Effect of Junk Science on Public Policy*. Westport, CT: Praeger
Schacter DL. 1996. *Searching for Memory: The Brain, the Mind, and the Past*. New York: Basic Books
Schacter DL. 2001. *The Seven Sins of Memory: How the Mind Forgets and Remembers*. New York: Houghton Mifflin
Schacter DL, Koutstall W, Johnson MK, Gross MS, Angell KE. 1997. False recollection induced by photographs: a comparison of older and younger adults. *Psychol. Aging* 12:203–15
Schooler JW, Bendiksen M, Ambadar Z. 1997. Taking the middle line: Can we accommodate both fabricated and recovered memories of sexual abuse? In *Recovered Memories and False Memories*, ed. MA Conway, pp. 251–92. Oxford: Oxford Univ. Press
Scott S. 2001. *The Politics and Experience of Ritual Abuse: Beyond Disbelief*. Buckingham, UK: Open Univ. Press
Sharman SJ, Garry M, Beuke CJ. 2004. Imagination or exposure causes imagination inflation. *Am. J. Psychol.* 117:157–69

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Sharman SJ, Manning CG, Garry M. 2005. Explain this: explaining childhood events inflates confidence for those events. *Appl. Cogn. Psychol.* 19:67–74

Snyder M. 1984. When belief creates reality. In *Advances in Experimental Social Psychology*, ed. L Berkowitz, pp. 224–305. New York: Academic

Snyder M, Thomsen CJ. 1988. Interactions between therapists and clients: hypothesis testing and behavioral confirmation. In *Reasoning, Inference, and Judgment in Clinical Psychology*, ed. DC Turk, P Salovey, pp. 124–52. New York: Free Press

Spanos NP, Burgess CA, Burgess MF, Samuels C, Blois WO. 1999. Creating false memories of infancy with hypnotic and non-hypnotic procedures. *Appl. Cogn. Psychol.* 13:201–18

Spanos NP, Menary E, Gabora NJ, DuBreuil SC, Dewhirst B. 1991. Secondary identity enactments during hypnotic past-life regression: a sociocognitive perspective. *J. Personal. Soc. Psychol.* 61:308–20

Stevenson I. 1994. A case of the Psychotherapist's Fallacy: hypnotic regression to "previous lives." *Am. J. Clin. Hypn.* 26:188–93

Tafarodi RW, Marshall TC, Milne AB. 2003. Self-esteem and memory. *J. Personal. Soc. Psychol.* 84:29–45

Tavris C. 1993. Beware the incest-survivor machine. *NY Times Book Rev.* Jan. 3, pp. 1, 16–17

Taylor SE, Crocker J. 1981. Schematic bases of social information processing. In *Social Cognition: The Ontario Symposium*, ed. ET Higgins, CP Herman, MP Zanna, pp. 89–134. Hillsdale, NJ: Erlbaum

Terr L. 1994. *Unchained Memories*. New York: Basic Books

Thomas AK, Loftus EF. 2002. Creating bizarre false memories through imagination. *Mem. Cogn.* 30:423–31

Usher JA, Neisser U. 1993. Childhood amnesia and the beginnings of memory for four early life events. *J. Exp. Psychol.: Gen.* 122:155–65

van der Kolk BA, Fisler R. 1995. Dissociation and the fragmentary nature of traumatic memories: overview and exploratory study. *J. Trauma. Stress* 8:505–25

Wade KA, Garry M, Read JD, Lindsay DS. 2002. A picture is worth a thousand lies: using false photographs to create false childhood memories. *Psychon. Bull. Rev.* 9:597–603

Wagenaar WAC, Crombag H. 2005. *The Popular Policeman and Other Cases*. Chicago/Amsterdam: Univ. Chicago Press/Amsterdam Univ. Press

Weir IK, Wheatcroft MS. 1995. Allegations of children's involvement in ritual sexual abuse: clinical experience of 20 cases. *Child Abuse Negl.* 19:491–505

Whitfield CL. 1995. *Memory and Abuse*. Deerfield Beach, FL: Health Commun.

Williams LM. 1994. Recall of childhood trauma: a prospective study of women's memories of child sexual abuse. *J. Consult. Clin. Psychol.* 62:1167–76

Wilson AE, Ross M. 2003. The identity function of autobiographical memory: Time is on our side. *Memory* 11:137–49

Wright L. 1994. *Remembering Satan*. New York: Knopf

Annual Review of Clinical Psychology
Volume 2, 2006

Contents

The History and Empirical Status of Key Psychoanalytic Concepts, *Lester Luborsky and Marna S. Barrett* 1

Doctoral Training in Clinical Psychology, *Richard M. McFall* 21

Methodological and Conceptual Issues in Functional Magnetic Resonance Imaging: Applications to Schizophrenia Research, *Gregory G. Brown and Lisa T. Eyler* 51

The Use of Structural Analysis of Social Behavior (SASB) as an Assessment Tool, *Lorna Smith Benjamin, Jeffrey Conrad Rothweiler, and Kenneth L. Critchfield* 83

Reinterpreting Comorbidity: A Model-Based Approach to Understanding and Classifying Psychopathology, *Robert F. Krueger and Kristian E. Markon* 111

Women's Mental Health Research: The Emergence of a Biomedical Field, *Mary C. Blehar* 135

Posttraumatic Stress Disorder: Etiology, Epidemiology, and Treatment Outcome, *Terence M. Keane, Amy D. Marshall, and Casey T. Taft* 161

The Psychopathology and Treatment of Bipolar Disorder, *David J. Miklowitz and Sheri L. Johnson* 199

Attempted and Completed Suicide in Adolescence, *Anthony Spirito and Christianne Esposito-Smythers* 237

Endophenotypes in the Genetic Analyses of Mental Disorders, *Tyrone D. Cannon and Matthew C. Keller* 267

Schizotypal Personality: Neurodevelopmental and Psychosocial Trajectories, *Adrian Raine* 291

Autism from Developmental and Neuropsychological Perspectives, *Marian Sigman, Sarah J. Spence, and A. Ting Wang* 327

Obesity, *Anthony N. Fabricatore and Thomas A. Wadden* 357

Mild Cognitive Impairment and Dementia, *Marilyn S. Albert and Deborah Blacker* 379

Annu. Rev. Clin. Psychol. 2006.2:469-498. Downloaded from arjournals.annualreviews.org
by UNIVERSITY OF FLORIDA on 08/22/07. For personal use only.

Journal of Applied Research in Memory and Cognition 8 (2019) 403–410

Contents lists available at ScienceDirect

Journal of Applied Research in Memory and Cognition

journal homepage: www.elsevier.com/locate/jarmac

Commentary

Title IX and "Trauma-Focused" Investigations: The Good, The Bad, and the Ugly

Deborah Davis*
University of Nevada, Reno, United States

Elizabeth F. Loftus
University of California, Irvine, United States

Beginning in earnest during the Obama era, campus, state, and federal authorities have struggled to find Title IX rules, regulations, and investigatory procedures that would balance the needs and rights of those who allege sexual harassment or assault against those of the accused. The "dear colleague" letter issued in 2011 (US Department of Education, 2011) reminded campuses of requirements under Title IX and issued guidelines for their enforcement, including those concerning the nature of campus policies, the operation of Title IX offices and officials, and the process of resolution of complaints. Although the letter specified—as does Title IX itself—that "equitable" procedures should be used, it did not require that due process protections be adopted (such as allowing lawyers to participate, access to all evidence), but merely stated that if allowed at all, both parties must be able to use them. It specified that due process must be provided for the accused, but at the same time, these protections must not restrict or unnecessarily delay Title IX protections for the complainant. It also required that the relatively lax "preponderance of the evidence" standard of proof be employed to assess any claim, thereby making it easier to find in favor of the complainant.

In September of 2017, Education Secretary Betsy DeVos rescinded Obama era regulations. In November of 2018 DeVos proposed changes designed to eliminate restrictions in the investigations, to bolster the rights of the accused, to encourage more equitable investigations, and to allow the option to choose a higher standard of proof for allegations (such as "clear and convincing" rather than "preponderance of evidence"). Changes included options that either had not been mentioned or not required in Title IX regulations or DOE guidance, or not adopted or permitted in specific campus regulations. These included the previously absent or restricted right to cross-examine the witnesses during mandatory live hearings, equal opportunity to present witnesses and to examine evidence, and separation of those who investigate the complaint from those who make the ultimate finding. Even as DeVos proposed such changes, some courts had begun to order colleges to offer due process protections for the accused, such as the right to question accusers (Watanabe, 2018).

Title IX itself, the "dear colleague" guidance, and subsequent efforts to revise relevant regulations and procedures have evoked considerable controversy and criticism. Reflecting this dissatisfaction, our society has seen a growing number of individual and class action lawsuits brought against universities by alleged perpetrators who claimed that their rights were violated by unfair campus regulations and investigatory procedures, and by biased execution of these procedures (Watanabe, 2019). While some are ongoing, nearly half of these plaintiffs have won their suits or settled their claims with the schools (Gersen, 2019). Apparently, the DOE (2011) Title IX guidance has had the unintended effect of spawning procedures that violated many of its own dictates regarding equity and fairness. We regard the widespread discussion, airing of divergent views, and efforts to continually revise and improve the relevant regulations and processes as "the good" associated with Title IX. But despite these efforts, as we write this article, Title IX campus investigatory rules and procedures are still hotly debated.

Campus Title IX offices have scrambled to comply with new guidance, regulations, and court rulings, and to offer increasing due process protections to defendants while still offering support and protection to alleged victims. On the one hand, there is no question that alleged victims of sexual assault have

Author Note.
Deborah Davis, University of Nevada, Reno, NV, United States.
Elizabeth F. Loftus, University of California, Irvine, United States.

* Correspondence concerning this article should be addressed to Deborah Davis, University of Nevada, Reno, United States. Contact: debdavis@unr.edu.

historically faced extraordinary scrutiny and doubt of their claims, greater than that faced by victims of any other crime (see Rerick, Livingston, & Davis, in press, for review). This fact has been brought home once again by a recent lengthy investigation documenting the widespread, and—in some jurisdictions—almost complete failure of law enforcement to take claims of rape seriously. "Seriously" means enough even to test rape kits or to otherwise pursue thorough investigation of reported rapes (Hagerty, 2019). Hagerty suggested such failures reflect an "epidemic of disbelief" of victims of rape that has led to an overwhelming societal failure to catch and convict rapists. Likewise, the social science literature has documented the many factors contributing to failure of victims to report rape, disbelief of victims who do report, and failure to convict perpetrators (e.g., Allison & Wrightsman, 1993; Reddington & Kreisel, 2017; Ward, 1995). Here we do not question these realities, nor do we suggest that the problems facing real victims of sexual assault have ceased.

On the other hand, it is also clear that in an effort to counter these problems for the victim, Title IX investigatory rules and regulations—and how they are enacted—have placed many accused in jeopardy. In practice, there is a presumption of truth in claims of rape, and an adoption of investigatory training and procedures that may bias findings in favor of the complainant and against the accused, to the point that some schools' investigators or adjudicators have been trained to "start by believing" the complainant (Gersen, 2019). While we recognize the many challenges facing real victims of sexual assault in having their claims taken seriously and prosecuted fairly, we also suggest that a system that tends to presume guilt of the perpetrator is no better than one that refuses to recognize valid claims of the victim. And, as Meissner and Lyle (2019) note, any presumption of guilt or other bias toward one party can set in motion confirmatory processes leading to biased collection and interpretation of evidence.

In the sections to come, we discuss the basis of these concerns, with focus on Title IX investigations, including (a) some flawed assumptions that may directly promote judgments of guilt; (b) investigatory procedures that can produce evidence biased in favor of complainants; (c) incorrect assumptions regarding memory and behavior that encourage *interpretation* of available evidence as favoring the complainant (but nevertheless may produce biases against them in some cases); and (d) omission of information in training and in collection of evidence in investigations that would importantly inform judgments. We devote considerable attention to the nature of trauma-focused interviewing and the way in which these four categories of concerns are reflected in the training for such interviews and in the interviews themselves.

Title IX Investigations and the Presumption of Guilt

There has long been an argument among prosecutors and victim advocates that the base rate of guilt among those accused of sexual assault is extraordinarily high. That is, many in the legal system make the assumption that almost no allegations are false, and trial testimony by purported experts citing statistics to this effect is common (despite fundamental methodological flaws in studies attempting to establish rates of false allegations). In effect, just as Carl Sagan famously asserted that extraordinary claims require extraordinary evidence, (Sagan & Druyan, 1997), prosecutors invite jurors to presume a false allegation to be extremely unlikely (extraordinary), and therefore to require extraordinary evidence of innocence to vote to acquit.

Notwithstanding admonitions to conduct fair and equitable investigations, we suggest that Title IX regulations and procedures reflect this assumption of low base rate innocence among the accused, and that they may well infuse this assumption into the minds of those who must investigate and judge rape or sexual harassment complaints. There are numerous reasons for such a claim. First, the renewed emphasis on Title IX concerns was inspired in part by widely publicized statistics on campus rapes. The DOE "dear colleague" letter (2011) stated at the outset that "the statistics on sexual violence are both deeply troubling and a call to action for the nation" and proceeded to report statistics indicating that 20% of college women and 6% of college males have experienced sexual assault (p. 2). Though the methodologies for assessing rates of sexual assault and the accuracy of such rates have been contested (see Krause et al., 2018; Muehlenhard, Peterson, Humphreys, & Jozkowski, 2017 for reviews), they continue to be presented in multiple contexts, including in Title IX training.

The evolution of Title IX regulations and procedures has continued in the context of the #metoo movement. The movement emphasizes the pervasive nature of sexual harassment and sexual assault and the extent to which reporting these actions has long been discouraged, and disbelieved when reported. Widespread media coverage and discussion related to #metoo encourages a cultural zeitgeist suggesting claims are to be believed and perpetrators brought to a long overdue reckoning.

Second, though even the Obama era DOE letter (2011) emphasized the importance of fair and equitable procedures at one level, it and subsequent regulations have simultaneously emphasized the importance of student safety on campus, and in doing so have disadvantaged the accused. Their specific statements regarding safety and associated rules and procedures concern the safety of alleged victims, not of the accused. For example, an accused may be subject to exclusion from any contact with the accuser (including being removed from dormitories and classes, or being suspended) even prior to any determination of guilt. Due process protections for the accused must not interfere with provision of safety for the accuser. Active measures must be taken to prevent revenge of the accused against the accuser. Moreover, remedies for the complainant can include escorts to ensure safety from the accused, counseling or medical services, academic support services such as tutoring, ensuring that class withdrawals do not affect the complainant's academic record, and reviewing the complainant's disciplinary history to assess whether any problems may have been caused by the actions of the accused. But none of these apply the other way around (DOE, 2011). These reflect an assumption that the accuser is indeed a victim, though it is clearly also possible that the accused may be the victim of a false allegation.

Third, we suggest that a presumption of guilt is reflected in the very notion of "trauma focused" or "trauma informed" interviews and investigations. As Meissner and Lyle (2019) note, the *First Report of the White House Task Force to Protect Students From Sexual Assault* (2014) tasked the Justice Department's Center for Campus Public Safety with the development of a "trauma-informed" training program for investigation of allegations of sexual misconduct. This focus on trauma was further encouraged during the Obama administration and later the Task Force's second report in 2017. This focus is reflected in the Forensic Experiential Trauma Interview (Strand & Heitman, 2017) and in many statements and training materials for law enforcement, statements posted on campus websites, those of victim advocate organizations, and others (e.g., Webb et al., 2018). Though there is much "good" about the recommended process of the interview and many accuracies in portrayal of memory processes, the "bad" consists of additional incorrect and unstated assumptions and specific assertions about how the nature of interviewee memory reports informs judgments of the reality of their claims, and the "ugly" consists of the way in which these fallacies can mislead judgments: mostly favoring the complainant, but in some cases inappropriately favoring the accused. It is worth noting at this stage that training for FETI encourages the assumption that the accuser is traumatized (and therefore was raped), and encourages the interpretation of all responses as consistent with that trauma.

Problematic Investigatory Tools: The Case of FETI (Forensic Experiential Trauma Interview)

As Meissner and Lyle (2019) discuss, the *procedures* of FETI are largely empirically supported, in that they essentially adopt elements of the widely tested Cognitive Interview, developed by Fisher and Geiselman (1992; 2010) and since shown to be effective in eliciting true information and minimizing false. These include developing rapport and demonstrating empathy, developing the interviewee's interest in the interview, use of open-ended prompts, active listening, and avoidance of leading or suggestive questions, for example. Differences lie primarily in FETI's emphasis on asking about emotions and sensory memories, though such questions are also used with the Cognitive Interview.

We find little fault with the recommended interview procedure itself, though we do note that there is much less attention paid in campus procedures and by promoters of FETI to what is the appropriate way to interview the accused. Instead, we focus here on several problematic assumptions and on questionable assertions concerning the meaning of responses obtained through use of FETI and the way in which they may mislead judgment, primarily in favor of the complainant.

Problematic Assumption 1: Sexual Assault Necessarily Produces Trauma Sufficient to Disable Cognition

FETI training suggests that sexual assault will be experienced as severely traumatic (see Strand & Heitman, 2017; and presentation of Strand posted at https://vimeo.com/117832921), as does the general exhortation of Title IX guidance to conduct trauma-informed investigations. From this assumption, the training goes on to discuss how memory works when experiencing trauma and what this will mean for the nature of traumatized persons' accounts. Derivations concerning memory and the meaning of memory reports are based on the presumption that cognition will be profoundly affected. Such claims raise the question of the extent to which the experiences of all or most victims of sexual assault (nevermind sexual harassment) include negative emotions rising to the level presumed in FETI training. That is, the training materials present effects of extreme trauma on stress hormones, cognition, behavior, and memory rising to the level of disabling frontal lobe executive functions and exerting debilitating effects on memory formation. If the event in question does not create such extreme emotions, what does this mean for the remainder of the training specifying what to expect victim accounts to look like and for the implications of either conforming or not conforming to those expectations?

It is highly unlikely that all sexual assaults result in such extreme emotional reactions (particularly in many disputed cases of acquaintance rape) and even more unlikely that all sexual harassment does so. The level of negative emotion that will disable cognition to the extent proposed by FETI training is a relatively high standard. Many of the same considerations raised in the repressed memory literature are relevant in cases of rape as well (e.g., Clancy, 2009; McNally, 2003). How extreme were the emotions *at the time it occurred* (even when blunted by alcohol), versus experienced later upon reflection, for example? When do negative emotions cross the line to become sufficient to disable cognition to the extent FETI training suggests? We suggest that the extreme trauma assumption itself and related theories regarding effects of trauma on memory and behavior are both problematic, and, as we explore below, can lead to inappropriate inculpation of the accused as well as inappropriate disbelief of the alleged victim.

Problematic Assumption 2: Trauma Is a One-Way Street

Generally, Title IX investigations are tasked with being "victim-centered" and "trauma-informed." This exhortation directs much of the effort toward care and handling of the alleged victim (including greater victim-focus in interview training). The theory and training of FETI, for example, focuses on how to interview the alleged victim. Where the alleged assailant is mentioned in the training, it is to contrast the presumed status of the victim and assailant brains and processing status (disabled by emotion vs. calm and rational), and therefore the likely types of memory reports they will be able to offer (disorganized, fragmentary and difficult to access vs. organized and more accessible). The import of the discussion is to suggest that the assailant will not likely experience intense emotions and therefore will not experience impairments of cognition or memory—at least not impairments approaching those experienced by the alleged victim. Purportedly, the perpetrator will be calm and unemotional, in part because most are repeat offenders who plan assaults and find them habitual and enjoyable in most instances (https://www.bwjp.

org/resource-center/resource-results/the-forensic-experiential-trauma-interview-feti.html).

But in a fair process, shouldn't consideration of emotion and trauma extend to the accused? If a person is innocent, strong negative emotions are not likely relevant until an accusation is made. But once an accused becomes aware of the accusation, they are undoubtedly common. One of us (Loftus) worked on the famous case of Cardinal Bernardin of Chicago, who himself was accused of sexual misconduct by a former seminarian, Stephen Cook. Writing in his memoir, the Cardinal talked about his initial reaction: He was "startled and devastated" (Bernardin, 1997, p. 19). He would also tell others that the false accusation and the cancer that ultimately killed him were the worst things that had happened to him in his life. Fortunately, before Bernardin died Stephen Cook dropped the case, making clear by his comments that he was convinced of Bernardin's innocence. But how might the terrible emotional devastation Bernardin felt because of the accusation have affected his reports during relevant interviews?

Strong emotions can affect the encoding of events. But they also impair retrieval. FETI training notes that the potential for intense emotions and re-traumatization is present for alleged victims, even during a supportive FETI interview—and these may affect memory reports. A more everyday experience is when people cannot remember something while nervous in front of an audience or while taking a test, but remember as soon as they leave the stressful situation. Such processes likely also affect memory reports for respondents. In addition to effects of strong emotion, an innocent accused may also be distracted by the struggle for answers. He may be confused by the unexpected accusation and struggle to understand why the accusation was made, or how his behaviors might have been misreported or misinterpreted.

To the extent that a FETI interviewer expects the accused to be able to provide a linear "who, what, where, when, and how" account and the accuser to provide a disorganized, incomplete and sometimes inconsistent account (as suggested by FETI training: e.g., Strand webinar presentation: https://www.bwjp.org/resource-center/resource-results/the-forensic-experiential-trauma-interview-feti.html), that interviewer may inappropriately interpret normal failures of memory as indicators of deception by the accused, while interpreting similar disorganization or mistakes as indicators of truth for the accuser. In short, there are many potential explanations for failures of frontal lobe executive functions, and for memory disorganization, errors and omissions of the sort anticipated by FETI for victims. These include alcohol or drug impairment, stress during recall, and lying. Such report characteristics cannot be assumed to reflect the validity of either party's account.

Other inequities regarding interviews or their interpretation are also important. For example, given that the accused may also be traumatized, the same considerations of safety and acceptance should apply as for accusers. Moreover, as FETI trainers suggest, asking open-ended questions about what the person was thinking or feeling during an event can trigger important information for the investigation. These techniques should be used to give respondents the best chance to retrieve potentially exculpatory information, just as they are recommended to give the accuser the chance to retrieve accusation supportive information.

Problematic Assumption 3: Reliable Differences Exist in Characteristics of Memory Reports for Traumatic versus Non-Traumatic Events

FETI materials outline the way in which trauma is expected to affect the manner of encoding, and therefore memory reports, for traumatic events. Space does not permit full exploration of the validity of all such claims. However, many are correct: for example, claims that intense emotion can impair frontal lobe function, reduce control of attention, focus attention more strongly on the perceiver's central concerns or the most salient aspects of the event, reduce attention to peripheral concerns, and others. As a result, the accounts of victims of trauma will purportedly be disorganized, focus on feelings and sensations, be inaccurate regarding timing, order, and other contextual details, and be inconsistent within and across tellings. Normal pathways of retrieval are expected to be impaired due to peritraumatic dissociation, which prevents normal associative pathways between elements of the event from being formed and makes retrieval of relevant memories more difficult. In contrast, the accused is expected to provide much more organized accounts, better able to satisfy demands of investigators for "who, what, when, where, how" information.

As Meissner and Lyle (2019) review, however, evidence does not support the predicted stark differences in accounts of traumatic versus non-traumatic memories, or between accused and accuser. The FETI training fails to acknowledge the many pathways to any given failure or characteristic of memory. Moreover, there is almost no mention in FETI training of the way in which alcohol may alter the emotional experience of sexual assault or the interaction of alcohol and emotion on memory processes.

Problematic Assumption 4: Characteristics of Memory Reports Can Be Taken as "Evidence" of Whether Trauma Did Occur

An issue of considerable importance is that of whether memory reports taken with FETI procedures and judged under FETI claims regarding traumatic and non-traumatic memories and victim-perpetrator differences in memory characteristics will lead to more accurate conclusions regarding the truth of the claims. Several issues are relevant to this question.

Is there really a "profile" of a true report? The unfortunately "ugly" result of FETI-related mistaken assumptions of whose memory reports should look how and under what conditions is the risk of mistaken judgments. A particularly ugly feature of FETI training is that it specifically suggests that if memory reports of alleged victims fit the "profile" of those expected from a trauma victim, this fit should serve as evidence that the report is true: "good solid neurobiological science routinely demonstrates that, when a person is stressed or traumatized, inconsistent statements are not only the norm, but sometimes strong evidence that the memory was encoded in the context of severe stress and trauma" (Strand & Heitman, p. 2). Clearly, given that evidence does not show that trauma is the

unitary cause for such memory reports (see next section), such an assumption poses considerable risk of an overly confident attribution of truth to an accusation.

FETI developer Strand made other such claims in his webinar (https://www.bwjp.org/resource-center/resource-results/the-forensic-experiential-trauma-interview-feti.html). One such claim might disadvantage actual victims. That is, Strand stated that he would be suspicious of alleged victims who were able to remember too much peripheral detail (because trauma should narrow attentional focus to central aspects of the event). It is notable, though, that he also made the contradictory claim that because victims often dissociate during rape, they may be focusing on peripheral details to avoid focus on what is happening. Strand also claimed that when alleged victims report expected emotional/behavioral reactions (e.g., terror, freezing), these reactions (which he clearly presumes true) can provide support for required elements of proof for prosecution of rape claims, such as fear, force, or nonconsent. In other words, the victim's reported reactions are considered proof that nonconsensual or forcible sex did occur.

What other explanations exist for memory features FETI training attributes to trauma? The claims above adopt logic such as the following: If men have four appendages, then all creatures with four appendages must be men! Just as there are many other creatures with four appendages, there are also many factors potentially responsible for the types of memory accounts that FETI training offers as evidence of trauma (or against). Prominent among them is intoxication, which, at high levels tends to produce fragmentary memories, myopic focus impairing memory for periphery, and other effects similar to those specified by FETI training (see Davis & Loftus, 2016, for review). Generally, one might ask how the purported memory effects of trauma might be different from what happens when one is asked about things (or is trying to remember things) that were never encoded clearly for any reason, or when one is lying. But even truth tellers may provide less coherent or complete reports when under the cognitive load imposed by the stresses of an accusation, the investigation, interviews, and worry over consequences. Modern studies of lie detection have focused, as Meissner and Lyle (2019) review, on what happens when one imposes a cognitive load on would-be deceivers. A liar will have trouble with peripheral details, and with unusual requests for information or unusual manners of retrieval (such as reverse order; see Vrij, 2019; Vrij, Hartwig, & Granhag, 2019). Combined with the effects of other stressors, the cognitive load imposed by efforts to lie successfully can produce memory reports sharing many features with those FETI training attributes to trauma.

Does sex-related trauma = sexual assault? One thing to keep in mind while considering the issue of "traumatic" memories is that a person may experience fear, high stress, or "trauma" during a sexual encounter, even though objective judgment would not suggest the encounter should be viewed as sexual assault. This scenario might occur, for example, if the alleged victim found the encounter highly aversive, but provided no overt indicators of nonconsent, or engaged in overtly voluntary, but actually unwanted, sex. In such circumstances, the advice inherent to the FETI training (Strand & Heitman, 2017) to consider evidence of trauma (in the form of memory characteristics and reports of emotions) as evidence of truth of the accusation can lead to misleading inferences that these feelings indicate assault.

One of us (Davis) served as expert witness in a case illustrating perfectly the problems with such an assumption. The participants, who we will call Jane and John, were both inexperienced: she a virgin and he a near virgin. They were both interested in one another, and arranged a late meeting at his apartment to watch a movie. Jane had told John early in the visit that she wasn't ready for sex yet (the primary argument for nonconsent). Yet, as the evening progressed she engaged in many behaviors that suggest consent. Her cross-examination at trial consisted in essence of the following: Did you get in bed with John? [Yes.] Did you make out with him with clothes on? [Yes.] At some point, did he begin to remove your shirt [bra]? [Yes.] Did you say no or try to stop him? [No.] At some point, did he being to remove your shorts [underwear]? [Yes.] Did you say no, tell him to stop, or try to physically stop him? [No] Did you in fact raise your buttocks as he tried to remove them? [Yes.] Did he then move to position himself between your legs? [Yes.] Did you spread your legs voluntarily? [Yes.] Did you at any time tell him not to insert his penis, say no, or in any way try to physically stop him? [No.] Nevertheless, Jane immediately reported the incident as rape to authorities. She made a recorded phone call at their behest to attempt to get John to admit to the rape on the record. During the call she talked about her feelings and the fact that she had early on told him she wasn't ready for sex, and tried to get him to admit that he knew she didn't want it. For his part, John was obviously excited at first to hear from her, talked about when they could get together again, and clearly assumed the previous night was the beginning of a relationship. But as Jane disclosed her feelings, cried, and talked about how he had to have known she didn't want it, John exclaimed "Jane! Oh my god! Jane! I raped you! Oh my god! I didn't realize! Jane! I'm so sorry. I'm so sorry. What can I do? I didn't want to hurt you!"

Nothing could have been clearer from their accounts and their trial testimony. Jane really hadn't wanted to have sex. She experienced intense negative emotions, arguably trauma, as the result of the interaction. She exhibited strong distress, and cried often, in her initial report, hearings, and trial, and upon news of John's acquittal. And yet, John clearly had no clue that she felt this way until the call the day after their encounter. His surprise was clear. Jane had not overtly made clear her very real desire not to have sex.

FETI trainers and others may assume that Jane experienced intense fear as it became clear to her that John might attempt to have sex with her, and as a result experienced "tonic immobility" and the inability to marshal resistance, and indeed she reported that she wanted to resist and didn't know why she didn't or couldn't. But this does not justify a finding of rape when it was unambiguous that John received no cues of resistance, and he made a genuine "mistake of fact" regarding consent. Confusion, lack of experience, and a poorly developed repertoire for negotiating potentially sexual interactions among naïve young people can produce many similar scenarios among college students.

Problematic Assumption 5: FETI's Focus on Emotion Has Only Positive Effects

The story of John and Jane raises another question regarding FETI. That is, great emphasis is placed on asking the accuser about the emotions and sensations she experienced during the event. The assumption is that such emotions and sensory memories are stronger than memory for "who, what, when, where, and why" details. Emotion and sensory-focused questions purportedly build rapport with the interviewee and can also trigger associative pathways by which the traumatized victim may be able to retrieve memories for other aspects of the event (Strand & Heitman, 2017; https://www.bwjp.org/resource-center/resource-results/the-forensic-experiential-trauma-interview-feti.html). We agree. But emotion focus, specifically, is likely to have other effects as well, raising the question of whether FETI's emotion focus may be a double-edged sword. In particular, two important issues deserve consideration.

What of the effects of emotion priming? Emotion related questions and prompts to relive emotions and sensations clearly serve a priming function. For a complex event, such priming could lead to preferential retrieval of emotion-consistent information at the expense of the contradictory, as shown by mood consistent retrieval effects. Moreover, as shown in the affective priming literature, emotion at retrieval can serve as context for the information that is retrieved, causing it to be interpreted in an emotion-consistent fashion (e.g. Bower & Forgas, 2001; Forgas, 2008; Gibbons, Seib-Pfeifer, Koppelheie-Gossel, & Schnuerch, 2018).

Finally, fuzzy trace theory (e.g., Brainerd & Reyna, 2005, p. 83) would predict that strong emotion memories, particularly in the absence of clear verbatim memories, would likely lead to constructive memory errors consistent with the emotion. The person might remember things that would be consistent with the emotion, even if they didn't happen (such as attempts to resist, or coercive actions by the accused), and fail to remember, or reinterpret actions that were inconsistent with the emotion. More generally, fuzzy trace theory suggests that when memories are unclear, one's general knowledge of what is likely in such circumstances will lead to memory errors consistent with general knowledge or expectations. In this way, if FETI training is correct regarding the lack of clarity in traumatic memories, this lack of clarity leaves open greater opportunity for memory to be distorted in the direction of expectations. Or, as Davis and Loftus (2016) put it, we tend to remember based on "who we think we are and what we think we did." John and Jane were not intoxicated during their encounter. But how much more opportunity for expectation-based errors is imposed by alcohol, or other impairments to memory clarity?

How good is memory for emotion itself? The FETI theory underlying its recommended procedures implies that emotion memory will be accurate and strong. These emotion memories can be used as pathways to retrieve accurate event-related information. But what if the person remembers the emotions incorrectly? A substantial literature exists to document inconsistencies in memories for emotion over time, and sources of distortion in memory for emotion comparable to those for other aspects of event memory. Notably, like other memories, memories for emotions are "functional," and alter over time in ways to facilitate one's current needs or goals, and they generally change to be consistent with current beliefs and appraisals (see Levine, Lench, & Safer, 2009, for review). Given that emotion memories are malleable, and particularly toward consistency with current goals, it is quite possible that when interviewees are asked to start with how they felt during an event, the emotions they report (and that serve as associative cues or as primes) may not reflect those experienced during the event, and as such, will not serve to prompt accurate memories or interpretations of the event.

Sexual Behavior and Sexual Consent Communications: A Glaring Area of Omission in Title IX Training

Our story of John and Jane raises other issues that are crucial to judgment of sexual assault. That is, there is a large scientific literature on how sexual consent is conveyed and interpreted, gender differences in perception of the meaning of behaviors that might or might not indicate consent, sources of misunderstanding of consent, effects of alcohol use on consent processes and perceptions, and other individual differences in these respects (e.g., see Davis & Loftus, 2016; Davis & Villalobos, 2014; Rerick et al., in press; Villalobos, Davis, & Leo, 2016; Wood, Rikkonen, & Davis, in press, for reviews).

Whereas the case of John and Jane is relatively clear regarding whether Jane displayed cues of nonconsent once the making out began, and many others are similar, it is also often the case that the clarity of consent is more difficult to judge, and correspondingly, the defense of reasonable mistake of fact. In this respect, greater training concerning norms of how consent or nonconsent tends to be communicated and interpreted in practice would be very useful. Did the accuser fail to convey cues of nonconsent widely recognized among students as such, or did the accused fail to recognize them if they occurred? Did the accuser engage in behaviors that she felt had nothing to do with consent, but that are widely considered to indicate consent? To our knowledge such issues are not covered in Title IX trainings.

Unfortunately, issues of interpretation complicate the task of those who must judge even further. FETI advocates Strand and Heitman (2017) noted that "What many in the criminal justice field have been educated to believe people do when they lie (e.g., changes in body language, affect, ah-filled pauses, lack of eye contact, etc.) actually occur naturally when human beings are highly stressed or traumatized" (p. 2). It is indeed clear that there are many confusions regarding both what may or may not indicate accuracy, subjective truth, or lying, as well as how perceivers understand and use such cues. The issue of consent communications adds to this the task of judging whether such communications conveyed each person's intentions clearly and whether each interpreted the other correctly. It is truly a difficult and error-fraught enterprise to judge who interpreted events correctly in the first place, who remembers accurately, who is telling the truth as they know it, and who is lying.

Conclusions and Caveats

Meissner and Lyle (2019) clearly lay out the case that much of what is taught to Title IX investigators—whether basic "facts," recommended procedures, or the theory underlying such procedures—lacks empirical tests or is unsupported or directly contradicted by existing research. We agree.

We have pointed to some problems with the training of Title IX investigators and the specific procedure of FETI and other "trauma-focused" interviewing. This discussion leaves open the question of how the investigations tend to be performed in practice. As Meissner and Lyle (2019) review, there are no formal training or minimum qualifications for those tasked with enacting recommended procedures, although such training may occur voluntarily, and many Title IX officials and investigators may possess important relevant qualifications. Assuredly, however, there will be significant variability in the manner and competence with which investigations are carried out. Such problems have been manifest in lawsuits against many colleges and universities based on failures of due process. More development of specific guidelines for *how* to conduct the investigations or recommended procedures (versus broad instructions such as to interview both parties) is needed.

It is also safe to assume that few investigators have adequate training in the many relevant areas of scientific knowledge such as detection of deception, interviewing, suggestion, memory, sexual behavior, sexual consent communications, and the effects of trauma on thinking, memory and behavior. This missing knowledge makes the task of making sense of the many reports and claims more fraught with error.

Meissner and Lyle (2019) point to the need to develop and employ evidence-based best practices for interviewing. We would add to this the need to provide more comprehensive education to those who must judge the complaints. Such education should cover evidence-based recommendations about how to interpret the information elicited through improved interviewing practices. Additionally, such education should convey accurate information about trauma and memory, as well as provide needed information about sexual behavior, sexual consent communications, and other topics.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The authors discussed this and conceived it together.

Keywords: Sexual assault, Memory, Title IX, Forensic Experiential Trauma Interview, FETI, Rape

References

Allison, J. A., & Wrightsman, L. S. (1993). *Rape: The misunderstood crime*. New York, NY: Sage.

Bernardin, J. C. (1997). *The gift of peace*. Chicago: Loyola Press.

Bower, G. H., & Forgas, J. P. (2001). Mood and social memory. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 95–120). Mahwah, NJ: Lawrence Erlbaum Associates.

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. NY: Oxford University Press.

Clancy, S. (2009). The trauma myth: The truth about the sexual abuse of children–and its aftermath. In *Basic Books*.

Davis, D., & Loftus, E. F. (2016). Remembering disputed sexual encounters: A new frontier for witness memory research. *Journal of Criminal Law and Criminology*, *105*, 811–851.

Davis, D., & Villalobos, J. G. (2014). Language and the law: Illustrations from cases of disputed sexual consent. In T. Holtgraves (Ed.), *Handbook of language and social psychology*. Oxford: Oxford University Press.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield: Charles C Thomas, Publisher.

Fisher, R. P., & Geiselman, R. E. (2010). The cognitive interview method of conducting police interviews: Eliciting extensive information and promoting therapeutic jurisprudence. *International Journal of Law & Psychiatry*, *33*, 321–328.

Forgas, J. P. (2008). Affect, cognition, and social behavior: The effects of mood on memory, social judgments, and social interaction. In M. A. Gluck, J. R. Anderson, & S. M. Kosslyn (Eds.), *Memory and mind: A festschrift for Gordon H. Bower* (pp. 261–279). Mahwah, NJ: Lawrence Erlbaum Associates.

Gersen, J. S. (2019, February 1). Assessing Betsy Devos's proposed rules on Title IX and sexual assault. *The New Yorker*. https://www.newyorker.com/news/our-columnists/assessing-betsy-devos-proposed-rules-on-title-ix-and-sexual-assault

Gibbons, H., Seib-Pfeifer, L.-E., Koppelheie-Gossel, J., & Schnuerch, R. (2018). Affective priming and cognitive load: Event-related potentials suggest and interplay of implicit affect misattribution and strategic inhibition. *Psychophysiology*, *55*, 1–19. http://dx.doi.org/10.1111/psyp.13009

Hagerty, B. B. (2019, August). An epidemic of disbelief: What new research reveals about sexual predators, and why police fail to catch them. *Atlantic Magazine*. https://www.theatlantic.com/magazine/archive/2019/08/an-epidemic-of-disbelief/592807/

Krause, K. H., Woofter, R., Haardörfer, R., Windle, M., Sales, J. M., & Yount, K. M. (2018, October 1). Measuring campus sexual assault and culture: A systematic review of campus climate surveys. *Psychology of Violence*, http://dx.doi.org/10.1037/vio0000209. Advance online publication

Levine, L. J., Lench, H. C., & Safer, M. A. (2009). Functions of remembering and misremembering emotion. *Applied Cognitive Psychology*, *23*, 1059–1075. http://dx.doi.org/10.1002/acp.1610

Nally Mc., R. M. (2003). *Remembering trauma*. Belknap Press.

Meissner, C. A., & Lyles, A. M. (2019). Title IX investigations: The importance of training investigators in evidence-based approaches to interviewing. *Journal of Applied Research in Memory and Cognition*, *8*, 387–397.

Muehlenhard, E. L., Peterson, Z. D., Humphreys, T. P., & Jozkowski, K. N. (2017). Evaluating the one-in-five statistic: Women's risk of sexual assault while in college. *The Journal of Sex Research*, *54*, 549–576. http://dx.doi.org/10.1080/00224499.2017.1295014

Reddington, F. P., & Kreisel, B. W. (2017). *Sexual assault: The victim, the perpetrator, and the criminal justice system*. Durham, NC: Carolina Academic Press.

Rerick, P. O., Livingston, T. N., & Davis, D. (in press). Rape and the jury. In W.O. O'Donohue, C. Cummings, & P. Schewe (Eds.), *Handbook of sexual assault and sexual assault prevention.* New York, NY: Springer.

Sagan, C., & Druyan, A. (1997). *The demon-haunted world: Science as a candle in the dark*. Location: Ballantine Books.

Strand & Heitman (2017). The Forensic Experiential Trauma Interview (FETI). Downloaded on July 28th, 2019, from https://www.slideshare.net/LoriDHeitman/feti-public-description-jan-2017

Villalobos, J. G., Davis, D., & Leo, R. A. (2016). His story, Her story: Sexual miscommunication, motivated remembering, and intoxication as pathways to honest false testimony regarding sexual consent. In R. Burnett (Ed.), *Vilified: Wrongful allegations of person abuse* (pp. 129–142). Oxford: Oxford University Press.

Vrij, A. (2019). Deception and truth detection when analyzing nonverbal and verbal cues. *Applied Cognitive Psychology*, *33*, 160–167.

Vrij, A., Hartwig, M., & Granhag, P. A. (2019). Reading lies: Nonverbal communication and deception. *Annual Review of Psychology*, *70*, 295–317.

Ward, C. A. (1995). *Attitudes toward rape: Feminist and social psychological perspectives*. London, UK: Sage.

Watanabe, T. (2018, November 16). DeVos moves to strengthen the rights of the accused in campus sexual misconduct cases. *Los Angeles Times*. https://www.latimes.com/local/education/la-me-edu-devos-sexual-misconduct-rules-20181116-story.html

Watanabe, T. (2019, August 3). UC student sues system in sexual harassment case. *Los Angeles Times*, p B1 and B4. (Also: Students accused of sexual harassment sue California universities. https://www.latimes.com/california/story/2019-08-02/california-universities-face-class-action-suits-by-students-accused-of-sexual-harassment)

Webb, K., Wyandt-Hiebert, M. A., Hanenberg, S., Beck, D., Claypool, T., Hoch, A., . . ., & Stewart, D. (2018). *Addressing sexual and relationship violence: A trauma-informed approach*. ACHA (American College Health Association).

Wood, E. G., Rikkonen, K., & Davis, D. (in press). Defining, communicating and interpreting sexual consent. In W.O. O'Donohue, C. Cummings, & P. Schewe (Eds.), *Handbook of sexual assault and sexual assault prevention.* New York, NY: Springer.

Received 11 August 2019;
accepted 11 August 2019

Moderate effects

Small effects in opposite directions

Table 1. Effect of Trigger Warnings on Anticipatory and Response Anxiety

Study	*N*	Source	Trauma exposure	Stimuli	Outcome	Anticipatory anxiety (*d*)	Response anxiety (*d*)
Bellet et al. (2018)	270	Crowd-sourced	No	Literature passages	Self-reported anxiety		0.06 [−0.18, 0.30]
Gainsburg & Earl (2018)[a]	276; 979	Crowd-sourced	Mixed	Essay	Negative affect (SAM)	0.75** [0.58, 0.92]; 0.26** [0.10, 0.42]	−0.17* [−0.33, −0.01]
Bellet et al. (2020)	462	Students	No	Literature passages	Self-reported anxiety		0.20* [0.02, 0.38]
Sanson et al. (2019)[b]	1,880	Students/crowd-sourced	Mixed	Story/film clip	Negative affect (PANAS)		0.02 [−0.08, 0.13]
Bridgland et al. (2019)[b,c]	1,600	Crowd-sourced	Mixed	Photos	State anxiety (STAI)	1.36** [0.99, 1.74]	0.07 [−0.03, 0.16]
Current study	451	Crowd-sourced	Yes	Literature passages	Self-reported anxiety		0.08 [−0.11, 0.26]

Note: Positive Cohen's *d* type effect sizes indicate an increase in anxiety. Values in brackets are 95% confidence intervals. SAM = Self-Assessment Manikin; PANAS = Positive and Negative Affect Scale; STAI = State Trait Anxiety Inventory. Cells are left blank for studies that did not measure anticipatory anxiety.
[a]The two anticipatory-anxiety effects are from Studies 2 and 3, respectively ($N = 276$, $N = 979$), and are based on reported t values; response anxiety is from Study 3. Confidence intervals are estimated on the basis of incomplete information. [b]Results are internal meta-analyses across all experiments. [c]Response anxiety is from our meta-analysis of the effects reported in Table 5.
$^{*}p < .05$. $^{**}p < .01$.

Adapted from Jones, P. J., Bellet, B. W., & McNally, R. J. (2020). Helping or harming? The effect of trigger warnings on individuals with trauma histories. Clinical Psychological Science, 8(5), 905-917.

The purpose of this study was to investigate how university students reacted to a passage from literature containing depictions of physical and sexual assault. In particular, the study was designed to investigate how one's trauma history and current symptoms affect an individual's responses to such material and how that might change over time.

In this study, the vast majority of participants were willing to read potentially triggering material. Right before reading the passage, all participants who were assigned to the triggering passage were given the offer, made in writing, to read an alternative passage that did not have any content related to physical or sexual assault. At the conclusion of Day 1, participants were asked to indicate, by checking a box, whether they read the assigned or alternate passage. It was clear to the participants that there was no penalty for making this change nor, given the anonymity of the study, was there any way to know if they made the change.

The majority of student participants decided to read the triggering passage even when they were aware of the content. Ninety-six percent of the students who were assigned the triggering passage decided to read it. These percentages were nearly identical for those who had identified themselves as having experienced a triggering trauma (96.9%) or those with a provisional PTSD diagnosis (97.6%). In the qualitative responses, only one participant specifically said they read the alternative passage because "I was afraid that I would be triggered emotionally by the material." Three indicated that "I preferred not to read the more difficult passage because I expected it to be unpleasant". One endorsed both of the previous responses. The most frequently endorsed reason participants chose the alternative passage was it "seemed more interesting". One read the alternative passage because "I could not understand which was the assigned reading", one read the alternative passage because they "didn't want to read some- thing depressing or sad", and two gave no reason for reading the alternative passage.

Kimble, M., Flack, W., Koide, J., Bennion, K., Brenneman, M., & Meyersburg, C. (2021). Student reactions to traumatic material in literature: Implications for trigger warnings. PLoS one, 16(3), e0247579.

Arch Sex Behav (2016) 45:1185–1193
DOI 10.1007/s10508-015-0666-2

ORIGINAL PAPER

Assessing Police Classifications of Sexual Assault Reports: A Meta-Analysis of False Reporting Rates

Claire E. Ferguson[1] · **John M. Malouff**[2]

Received: 2 February 2015 / Revised: 5 November 2015 / Accepted: 8 November 2015 / Published online: 17 December 2015
© Springer Science+Business Media New York 2015

Abstract The objective of the study was to determine, through meta-analysis, the rate of confirmed false reports of sexual assault to police. The meta-analysis initially involved a search for relevant articles. The search identified seven studies where researchers or their trained helpers evaluated reported sexual assault cases to determine the rate of confirmed false reports. The meta-analysis calculated an overall rate and tested for possible moderators of effect size. The meta-analytic rate of false reports of sexual assault was .052 (95 % CI .030, .089). The rates for the individual studies were heterogeneous, suggesting the possibility of moderators of rate. However, the four possible moderators examined—year of publication, whether the data set used had information in addition to police reports, whether the study was completed in the U.S. or elsewhere, and whether inter-rater reliabilities were reported—were all not significant. The meta-analysis of seven relevant studies shows that confirmed false allegations of sexual assault made to police occur at a significant rate. The total false reporting rate, including both confirmed and equivocal cases, would be greater than the 5 % rate found here.

Keywords Sexual assault · Rape · False allegations · Meta-analysis

✉ Claire E. Ferguson
claire.ferguson@qut.edu.au

1 School of Justice, Faculty of Law, Queensland University of Technology, GPO Box 2434, Brisbane, QLD 4000, Australia

2 School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale, NSW, Australia

Introduction

False allegations of sexual assault have the potential to ruin reputations, cost the accused their livelihoods, and waste both time and resources in the criminal justice system (CJS) (Levitt, 2013). Several studies have shown that some of the public, and some working in the CJS, believe false allegations of sex crimes are rife (Heenan & Murray, 2006; Jordan, 2004; Venema, 2014). This perception has been highlighted and perhaps exacerbated by recent media focus on accusations of sexual assault made on college campuses (e.g., Botelho & Gray, 2014). The view that false reports of sexual assault are rampant is at odds with most research findings in the area (Clark & Lewis, 1977; Grace, Lloyd, & Smith, 1992; Harris & Grace, 1999; Heenan & Murray, 2006; Kelly, Lovett, & Regan, 2005; Lisak, Gardinier, Nicksa, & Cote, 2010; Lonsway, Archambault, & Lisak, 2009; McCahill, Meyer, & Fischmann, 1979; Spohn, White, & Tellis, 2014). This view is problematic in that it may lead to self-fulfilling prophecies, especially when held by investigators determining whether allegations are legitimate and should proceed (Clark & Lewis, 1977; Heenan & Murray, 2006; Jordan, 2004; Kelly et al., 2005; Lea, Lanvers, & Shaw, 2003; Lisak et al., 2010). On the other hand, the perception that people do not lie about being sexually victimized is equally challenging. Either misperception can make reacting to claims of victimization much more difficult, and may contribute to the high levels of attrition presently occurring with sexual assault and rape reports internationally (Grace et al., 1992; Gregory & Lees, 1996; Harris & Grace, 1999; Kelly et al., 2005).

Why these misperceptions exist is a complex question and not the focus of the current discussion. This project instead focused on determining the frequency of confirmed false allegations of sexual assault made to police. This is necessary to explore, since allegations of sexual assault are extremely easy

to make, very difficult to refute, and can lead to severe damage to the accused.

Why Do We Need to Study False Allegations?

When estimating how often false reports of sexual assault or rape are made, different research methodologies yield results with marked variability. In a review of research findings, Rumney (2006) examined 20 different studies, with estimated rates of false allegations between 1.5 and 90 %. This broad range in estimates has allowed proponents to argue, ostensibly, for any conclusion desired. Depending on their specific agenda, some commentators report that false allegations of rape are basically non-existent or at least no more common than false allegations of other violent crime types (Theilade & Thomsen, 1986). Others maintain that large portions of sexual assault reports are false and therefore each should be treated skeptically (Firth, 1975; Kanin, 1994; MacDonald, 1973; Maclean, 1979; McDowell, 1990; Stewart, 1981). Both of these stances have some level of research support, making truly evidence-based conclusions difficult to draw. The result of this lack of consensus in the literature is a confused effect on legal doctrine, legislation, and police policy in many jurisdictions, with some advancements protecting complainants and making reporting to police easier, while others protect the accused and potentially make reporting much more difficult for false reporters. Moreover, inconsistent research findings may support the maintenance of problematic stereotypes on both sides of the argument, including that large numbers of individuals who were allegedly raped are fabricating their complaints, or that virtually zero individuals falsely report sex offenses.

Some of the variability in these rates is explained by to whom reports are made, with some studies looking at only complaints made to police and others examining individuals who present to medical professionals, crisis centeres, campus counselors, and so on. It is arguable that reports made to someone other than police (such as to college or university staff) would involve a higher false allegation rate, since they are easily made, there are fewer consequences, and sometimes greater benefits for the accuser. However much of the variability in rates reported in the literature is not explained by these differences in data sources, but other, more complex issues.

What Makes Studying False Allegations Difficult?

Unclear and Varying Standards for Deciding an Accusation is False

One of the reasons behind such a broad range of false reporting rates in the literature is the inherent difficulty of studying false accusations and interpreting results once studies have been conducted. Any examination of false reports is complicated by several issues, not the least of which is that it is never completely certain whether a report is true or false. That is, rape accusations are often made by one person against another, who denies the accusation, and there is little evidence other than the accusation to determine whether it is a true report. Even after the demonstrably false cases have been discovered, many more equivocal cases exist which cannot be confirmed or denied, and even recanted accusations may, in fact, be true. Researchers rarely address this problem or state what level of certainty they applied in deciding that a report was confirmed to be false. Additionally, after rates of false reporting are given, few researchers discuss the many other cases that were in doubt, but not proven or confirmed to be false.

Defining what constitutes a false report, and thus what to include in study samples, has historically been problematic for many researchers, leading to wide variations in samples and methodologies. In fact, difficulties defining false reports adequately have led some to forgo setting a definition at all. Instead, these researchers rely on the perceptions, policies, and classifications of investigative decision-makers, such as police or prosecutors.

In its simplest terms, most who choose to define a false allegation (also referred to as a false accusation or report) refer to it as an allegation made to police, either directly or through a third party, which the complainant knows to be untrue. This definition suggests awareness that the allegation is false and also maliciousness on the part of the complainant (Lisak et al., 2010). Reports that are unknowingly untrue (lacking awareness) and those which are non-malicious (involving mistaken identity, for example) are not considered false under most definitions used in the contemporary literature. However, many reports which are not true may not have these elements of awareness and maliciousness. Complainants may not be aware of what exactly happened to them because of alcohol or drug use; they may be uninitiated as to what constitutes a rape or sexual assault (or what does not) under the law in that jurisdiction; or they may be mistaken as to what took place due to mental health conditions. Reports from individuals such as this, although sometimes technically false, would not be included under the definition above or in research samples on the subject. The presence of cases like this makes defining and studying false allegations much more difficult than imagined.

In order for a report to be deemed false, most researchers add that not only must the report meet the definition outlined above, but it must also have been thoroughly investigated. Recently, commentators have been explicit that allegations cannot be deemed false simply because evidence fails to prove an assault took place. The International Association of Chiefs of Police (IACP) (2005) recommended:

> The determination that a report of a sexual assault is false can be made only if the evidence establishes that no crime was committed or attempted. This determination can be made only after a thorough investigation.

> This should not be confused with an investigation that fails to prove a sexual assault occurred. In that case, the investigation would be labelled unsubstantiated. The determination that a report is false must be supported by evidence that the assault did not happen (pp. 12–13).

With this recommendation in mind, most researchers have expanded the definition of a false allegation to include only those reports made which were untrue, which involved maliciousness and consciousness of the untruth on the part of the complainant, and where evidence exists to prove the crime or attempted crime did not take place. Applying this conservative definition of false allegations limits the number of allegations which are deemed "false" to only those which are confirmed through evidence. Although limiting the sample, this is a necessary step as it prevents opening the floodgates to many equivocal cases that are suspected but not demonstrably false. It errs on the side of caution by not including cases in doubt, mistaken cases, or those claims made to anyone other than police. Use of such a conservative definition is not meant to imply that all other cases are true reports, but just that they cannot responsibly be deemed confirmed false.

Accepting Police Classifications at Face Value

In the past, some researchers have accepted police classifications of rape cases for the purpose of determining rates of false allegations (Grace et al., 1992; Gregory & Lees, 1996; Harris & Grace, 1999). These researchers rightly noted that some police officers and agencies are quite skilled at assessing accusations of sexual assault, due to many changes to policy as a result of pressure to improve dating back many years. However, there are also reasons to doubt this confidence in police assessments of cases. First, using police classifications (i.e., simply counting how many cases investigators deemed as false reports out of the total reported) does not allow researchers to measure which officers or agencies are especially accepting (or critical) of these accusations. It also does not allow a determination of whether police are adhering to their own definitions and guidelines for classifying reports as false. Using samples of this type, researchers are forced to assume that police classifications are considered and accurately reflect the definitions in their policies. This is an unsafe assumption according to the findings of several studies (Gregory & Lees, 1996; IACP, 2005; Kelly et al., 2005; Lisak et al., 2010; Rumney, 2006). Second, reliance on such classifications does not allow researchers to determine whether any biases are present in how these guidelines and definitions are applied to specific cases on either side of the issue. For example, there is strong evidence to suggest that individuals raped by an intimate partner, those who appear intoxicated, and those who delay reporting to police may be viewed as more likely to be lying about the sexual assault due to problematic "real rape" stereotypes and rape myth acceptance (Clark & Lewis, 1977; Heenan & Murray, 2006; Jordan, 2004; Kelly et al., 2005; Lea et al., 2003; Lisak et al., 2010). On the other hand, those individuals who report the offense quickly, evidence injuries, and report not knowing the offender are more likely to be viewed as credible for the same reasons (Clark & Lewis, 1977; Heenan & Murray, 2006; Jordan, 2004; Kelly et al., 2005; Lea et al., 2003; Lisak et al., 2010). Neither of these stereotypes is completely accurate and the presence and effect of these biases simply cannot be known when researchers fail to scrutinize case classifications made by law enforcement or prosecutors. Third, using a sample of unscrutinized cases does not allow for any errors in recording or classification to be rectified before reports are counted into any specific category. This is an issue highlighted by researchers as potentially adding a significant amount of bias to subsequent results (Gregory & Lees, 1996; Kelly et al., 2005; Lisak et al., 2010; Rumney, 2006; Spohn et al., 2014). Numerous studies have found classification errors common in several jurisdictions internationally (Gregory & Lees, 1996; HMCPSI/HMIC, 2002; Kelly et al., 2005; Lisak et al., 2010; Rumney, 2006), with law enforcement professionals sometimes erroneously combining false and baseless cases (cases that are truthfully recounted but do not meet the legal definition of sexual assault) into one category. This is an error of classification, where "not sexual assault" is erroneously equated with "false."

Given the serious difficulties with studying false rape allegations, many of the reported false report rates, both high and low, cannot be relied upon for an accurate assessment of how often false allegations occur. As stated by Lisak et al. (2010), "there are actually very few studies that provide meaningful data on the frequency of false reports." By avoiding some of the pitfalls outlined above, a small number of researchers have been able to offer balanced assessments of rates of false reports made to police. Each of these teams of researchers has independently scrutinized classifications of investigating officers to determine whether a case was correctly categorized as false; articulated clearly the process used to determine how cases would be classified; explicitly stated their definition of a false report; and explained the source of their data. Because of the difficulties in defining such cases, these assessments provide rates of confirmed false allegations made to police only. The rates are thus conservative, with potentially many more false report cases existing but not being proven or reported to someone other than police.

Recently, Lisak et al. (2010) published a summary of five of these conservative studies. As such, each study does not need to be discussed in great detail again. Interested readers may consult Lisak et al. for a comprehensive discussion of each study. Instead, we will briefly summarize these studies and two others to offer an introduction to what research exists in the area, the definitions and samples used, and potential limitations. The findings of each study will be presented in the results section.

Definitions

Because different jurisdictions were examined by each of the authors in these studies, different definitions existed for rape/sexual assault in each. Some included specific behaviors which were involved (sexual intercourse, oral sex, etc.), while others simple addressed the general types of actions included, such as coerced sexual activity. Rather than report each separate definition, we have combined them into a summary. For our purposes then, the term sexual assault is used to refer to rape as well as other sex offenses. The term refers to sexual intercourse, sexual activity, or intentional sexual touching, using force, threat of force, impersonation, or misrepresentation, against the victim's will or in violation of the law. Some studies included attempted sexual assaults in their samples while others did not.

Samples

The samples used by these seven teams of researchers were all law enforcement based. However, several research teams also added other data to complement that collected by police. McCahill et al. (1979) in the U.S. also used medical data and detailed interviews of individuals alleging sexual assault conducted by social workers; Spohn et al. (2014) used data from law enforcement, witness reports, and the results of medical examinations, as well as interviews with experienced detectives. In the most detailed sample to date, Kelly et al. (2005) assessed data from police, medical examinations, interviews with investigators and witnesses, and several other sources.

Limitations of Relevant Studies

Many of the authors whose studies were sampled here made note of the fact that missing data were a major issue. This is inherent with using a law enforcement sample, rather than one designed specifically for research purposes. Heenan and Murray (2006) reported that their dataset was incomplete for 341 of 850 cases. Lonsway (personal communication, August 5, 2014) also highlighted a potential problem with her unpublished results, where it was unclear whether unchecked boxes on proformas (filled out by trained police researchers for each case in the sample) indicated absence of that variable or missing information. In the McCahill et al. (1979) study, only half of reported cases also involved victim interviews with social workers, meaning that in 50 % of the sample police classifications could not be scrutinized. This effectively cut their sample in half, as only those cases that were scrutinized by researchers could be included in the current analysis.

Three of the seven studies did not include information on how many cases were in doubt in addition to those confirmed to be false. Another three studies did not include allegations of sexual assault or rapes made to police by males. None of the studies broke down their sub-sample of false allegations by age or sex. This limits these results to a discussion of the rate of confirmed false reports as opposed to the false reporters themselves.

Meta-Analysis

The main purpose of the present meta-analysis was to quantify the overall confirmed false report rate for studies where the researchers or their helpers scrutinized police records regarding false reports rather than accepting police conclusions without question. A second purpose was to identify moderators of false report rates.

Method

We searched for studies reporting rates of false allegations of sexual assault using the electronic databases PsycINFO, PubMed, Sage Journals, and Google Scholar. The search terms were false allegation and sexual assault; attrition and sexual assault; false accusation and sexual assault; and false reporter and sexual assault. We also searched reference lists of review articles and published empirical reports to identify further possible studies for inclusion in the meta-analysis. Several researchers were contacted personally, requesting access to unpublished results. We completed the search in July 2014. We reviewed publications for whether they presented empirical information on rates of reports of sexual assault made to police that were determined by researchers or their trained helpers to be false. See Table 1 for the number and flow of evaluated articles.

We coded each included study for the date of publication, country in which the study was done, whether the dataset used had information in addition to police reports, and whether inter-rater reliabilities were reported. As it turned out, the only country with more than one included study was the U.S., so we categorized that potential moderator as U.S. or other country. We report all effect sizes as a false report odds ratio. We each coded every study independently. In all cases, we agreed.

The Comprehensive Meta-Analysis Program (Borenstein, Hedges, Higgins, & Rothstein, 2005) calculated the overall weighted effect size, using a random effects model. The Q statistic assessed effect-size homogeneity across studies. The trim and fill method and fail-safe N assessed the impact of possibly missing studies. We used the Q statistic to evaluate potential categorical moderators and method-of-moments meta-regression to evaluate year of publication as a moderator.

Results

Out of the seven studies identified in the literature that assessed investigative classifications, rates of confirmed false reports ranged from 2.1 to 10.3 %. On the lower end of this range was an Australian study, conducted in the State of Victoria over

Table 1 Characteristics of studies included in meta-analysis

Study	Country	Data based on more than police reports	*N*	False report rate (95 % CI)
Clark and Lewis (1977)	Canada	No	116	.103 (.060, .173)
Heenan and Murray (2006)	Australia	No	850	.020 (.012, .032)
Kelly et al. (2005)	UK	Yes[a]	2643	.025 (.020, .032)
Lisak et al. (2010)	USA	No	136	.059 (.030, .113)
Lonsway (personal communication, August 5, 2014)	USA	No	1984	.070 (.059, .082)
McCahill et al. (1979)	USA	Yes[b]	709	.034 (.023, .050)
Spohn et al. (2014)	USA	Yes[c]	401	.137 (.107, .174)

[a] Additional data included results of medical examinations, interviews with investigators and several other sources of information. See Kelly et al. (2005) for more information

[b] Additional data included results from medical examinations and victim interviews conducted by social workers

[c] Additional data included results of medical examinations and interviews with experienced sex-crimes detectives

3 years in the early 1990s (Heenan & Murray, 2006). This study was commissioned by the State-wide Steering Committee to Reduce Sexual Violence (Victoria), and examined a random sample of reported rape cases using the Law Enforcement Assistance Program database. Heenan and Murray did not indicate whether attempted rape was included in their sample, although they did include reports made by men (7.5 % of total) and children. The study found 17 confirmed false reports of rape out of 850, and noted that several more ($n = 77$) were "in question" by investigators. None of the false allegations were made by men in this study.

On the upper end of the range, Clark and Lewis' (1977) study in Toronto, Canada sampled all 116 rape reports made to police in 1 year. The researchers discovered 12 cases which were confirmed, using their definition, to be false reports. This amounts to a rate of 10.3 %. They did not include reports of attempted rape nor those made by men or children under the age of 14. In this sample, an additional 62 cases could not be definitively classified as true or false, and were thus equivocal.

The remaining five studies found rates between 2.5 and 7 %. Kelly et al. (2005) found a confirmed false reporting rate of 2.5 %. This study sampled all sexual assault cases across six sites and 15 years. Kelly et al. did not indicate whether attempted sexual assaults were included nor whether all ages were included, although they did include males alleging sexual assault. Kelly et al. also did not report on how many additional reports were suspected of being false but not confirmed.

McCahill et al. (1979) found a rate of 3.4 % in their 3-year study in Philadelphia. Reports made by females of all ages were included, regardless of whether completed or attempted sexual assault was alleged. McCahill et al. added that in addition to the 24 cases which were confirmed to be false, another 86 were in doubt or equivocal.

Spohn et al. (2014) examined a random sample of sexual assault cases from 1 year in one city police department. They included only female victims aged 12 years and older. They found a confirmed false reporting rate of 4.5 %, with an additional 2.7 % (11 cases) in doubt or equivocal.

Lisak et al. (2010) studied reports of sexual assault made to one university police department over 10 years. They included reports made by males and females, and did not indicate if any age ranges were excluded. They found 5.9 % of allegations were confirmed to be false. These researchers did not indicate whether, or how many, additional cases were suspected but not confirmed false reports.

Finally, Lonsway's unpublished results (personal communication, August 5, 2014) indicate a confirmed false report rate of 7 %. Lonsway studied 8 different communities over an 18–24-month period, examining all reports of sexual assault (both attempted and completed) made by individuals aged 12 years and up. Males were included in this sample. In addition to the 138 cases deemed to be false, they did not report how many more were equivocal.

Table 1 shows characteristics of the seven studies included in the meta-analysis. The meta-analytic false report ratio was .052 (95 % CI .030, .089). The effect sizes were heterogeneous, $Q(6) = 128, p < .001$. Analyses of the three categorical potential moderators evaluated showed all were not significant, as indicated in Table 2. Publication year was not significant in method-of-moments meta-regression, point estimate = .014 (95 % CIs −.016, .045), $p < .36$.

Discussion

Many of those writing about sexual assault have not adequately dealt with the issue of false allegations (Rumney, 2006). Some have chosen to ignore the problem altogether, whereas others have attempted to assess the rate of false reports using various, sometimes unreliable, methods. The meta-analytic false report rate of about 5 % indicates that a small but significant number of sexual assault reports made to police are confirmed to be false.

Table 2 Categorical moderator results

Moderator results	k	Ratio	95 % CI		Homogeneity analysis			Between categories	
			Lower	Upper	Q	df	p	Q	p
Dataset included information in addition to police reports								0.014	.90
No	4	.054	.028	.102	90	3	<.001		
Yes	3	.050	.016	.149	94	2	<.001		
Inter-rater reliability reported								2.57	.11
No	5	.041	.023	.074	72	4	<.001		
Yes	2	.096	.041	.208	6	1	.017		
Country of study								1.51	.22
USA	4	.068	.039	.117	39	3	<.001		
Other country	3	.037	.016	.082	23	2	<.001		

The statistical combining of these studies, using a conservative definition and only measuring cases confirmed to be false, reveals a rate much lower than that proposed by others (e.g., Kanin, 1994; Stewart, 1981).

Recent evidence suggests, despite efforts to change these views, some police officers still believe the rate of false reports for sexual assault to be higher than what can be demonstrated through careful examination. This was recently evidenced in the work of Heenan and Murray (2006), Jordan (2004), and Venema (2014). Conversely, others may have gone too far in the opposite direction, being very sensitive to these accusations and hesitant to deem a claim false even when convincing evidence exists to support them. Both of these errors may influence some police investigators to be biased in their work (Lonsway, personal communication, August 5, 2014; Spohn et al., 2014). Convincingly demonstrating that rates of false reporting differ from these perceptions may help dilute these views, possibly making reporting sexual assault less difficult for victims and more difficult for false reporters. In the future, training for sexual assault investigators should continue to utilize these findings (and others) to ease inaccurate stereotypes on both sides. Research and training should continue to focus on measuring and then combating myths about the characteristics of "real rapes" or "real victims" which are sometimes held by investigators. The present results may help explain some of the controversy about false accusation rates in a meaningful way.

Moderator Results

The meta-analysis had low power because only seven studies met the inclusion criteria. Hence, it is not surprising that the four potential moderators examined were not significant. The only trends in the direction of significance involved studies done in the U.S. showing higher false report rates than studies in other countries and those reporting inter-rater reliabilities showing higher false report rates.

Comparison to Other Studies

Studies using other methods and definitions have found false reporting rates above 40 % (Gregory & Lees, 1996; Kanin, 1994; Maclean, 1979; Stewart, 1981). The discrepancies between the current findings and the findings of these works are likely the result of using a conservative definition for false reports in the current study and requiring police classifications to be scrutinized by researchers. The definition we used requires awareness by the accuser of the falsity of the accusation, maliciousness, and evidence that the sexual assault did not take place. Hence, the current results provide a rate of confirmed, rather than suspected, false allegations. For example, Kanin's (1994) widely referenced 41 % rate of false allegations is based on a much less strict criterion, including what we would call "confirmed" and "equivocal" cases. His definition of falsity required only that a complainant state that the sexual assault did not happen after being asked to take a polygraph test. This procedure has been highly criticized and has even been legally prohibited in some states due to its potential to intimidate already hesitant complainants (IACP, 2005; Lisak et al., 2010; Rumney, 2006). Whether these reports were confirmed as false based on the evidence was also not assessed by Kanin directly and remains unknown. In our view, Kanin's result may be more aptly called a rate of "possible" false allegations. Given these very different methods and thresholds for determining falsity, it is unsurprising that such a discrepancy exists between results discovered using definitions like this and our more conservative definition.

Comparison to False Reporting Rates for Other Crimes

Are false reports more common with sexual assault than with other crimes? As with sex offenses, reliable research on rates of false allegations of other violent crime types is challenging to both find and interpret. Researchers sometimes cite a figure of

2 % for false allegations of non-sexual violent crimes and then (incorrectly) argue that rates are consistent across crime types (Russell & Bolen, 2000; Wells, 1985). This 2 % rate is taken from the Uniform Crime Report in the U.S., but, unfortunately, suffers the same problems which exist with classifying sexual assault allegations as false, such as sometimes systemic miscategorization of cases. This makes a false reporting rate of 2 % for non-sexual offenses dubious at best and greatly limits comparability across crime types. The same issues unfortunately exist with other research on false reporting rates for violent non-sexual offenses, with estimated rates ranging from 2 to 19 % depending on the methodology employed and crime type studied (Chambers & Millar, 1986; Sheridan & Blaauw, 2004; Theilade & Thomsen, 1986; Zona, Lane, & Moore, 1996).

However, rates of false reporting for other types of sex offenses, although also suffering some of the same issues, can provide a useful context for the current findings. Comparing false allegations of sexual assault to those of child sex abuse is most helpful because both types of allegations rely heavily on reports of the person making the allegations, and evidence indicating that the offense did or did not take place can be ambiguous since physical signs are often lacking. Adults, on behalf of children, can also make false reports of child sexual abuse, and the motivation may be similar to that in sexual abuse against older people (e.g., as revenge against an ex-partner or to gain access to services). Anthony and Watkeys (1991) found a false reporting rate for child sexual abuse of 8.5 % (in fully investigated cases) in the UK using a definition, including maliciousness, similar to ours. Goodwin, Sahd, and Rada (1985) and Peters (1976) found rates of false allegations of sexual abuse against children occurred at a rate of 8.5 and 6 %, respectively. All of these rates are higher than the 5 % confirmed false allegation rate found here for sexual assault more broadly. This may be due to additional, strong motives to claim child sex abuse by some parents, especially during custody disputes (Green, 1986). In support of this explanation, Lamb, Sternberg, Esplin, Hershkowitz, and Orbach (1997) found a false reporting rate of 3 % when studying only those claims of sexual abuse made by the children themselves. Of course, these comparison studies are not without limitations; however, the findings do provide some context for the current results.

In the end, it is very difficult to draw a conclusion as to whether false allegations of sexual assault are any more or less common than those of other crime types because of a lack of methodologically sound results and many potentially confounding issues. Tentatively, it seems as though false allegations of sex crimes (either against children or adults) happen more frequently than false allegations of other crime types. This is likely due to the ease with which these types of allegations can be made, the inherent difficulty of refuting them, and the level of acceptance and lack of skepticism they may have been afforded in the past. Examples such as the recovered memory therapy and day care abuse cases of the 1980s and 1990s in the U.S. indicate there is a need to be very cautious when examining claims of sexual abuse, both past and current.

Limitations

This meta-analysis was limited by a number of factors, most of which were a direct result of the limitations of the studies included. These limitations involve issues related to the locations sampled, the data accessed, the level of scrutiny given to police classifications, and errors in assessing sexual assault reports.

Generalizability Across Locations

Four out of seven studies sampled here used data from American sources. The most comprehensive of these examined data from 8 communities, whereas others looked at either one city police department or one university police department. The United Kingdom, Australia, and Canada were represented by one study each, including data taken from segments of the country, involving six regions, one state, and one city, respectively. Combining the findings of these studies into a meta-analysis means that, although four large Western democratic countries were represented, results may not be generalizable to all regions or cities in these countries. There is anecdotal evidence suggesting that many of the motives which are thought to cause false reporting of sexual assault, such as mental health issues in complainants, attention seeking, and revenge (McNamara, McDonald, & Lawrence, 2012), are similar the world over. However, this has not been demonstrated through empirical study. Various other factors, such as victim compensation incentives, might affect the rate of false reporting in any one place, thus further limiting generalizability to other locations.

Data Accessed/Level of Scrutiny

The findings of this meta-analysis may also be limited by the data that was accessed by the original authors and the level of scrutiny afforded to it. Unfortunately, it is unknown if any biases were present in the researchers who assessed these police data to determine which cases were actually false reports. Some safeguard against researcher bias lies in the fact that many of the studies sampled involved teams of researchers where cases were assessed by more than one independent person. However, it must be noted that determinations of truth or falsity were still made by researchers themselves, and their level of validity is therefore unknown. This is especially an issue in those studies where inter-rater reliability was not calculated.

Furthermore, the amount and type of data accessed by each of these teams may play a role in the independent

classification of cases, where those cases with less available data may have been more difficult to classify, and thus less likely to be deemed confirmed false based on requirement of a thorough investigation in our definition.

Errors in Assessing Sexual Assault Reports

A proportion of cases in the studies included here were equivocal, where the researchers simply could not determine whether the accusation was false. Certainly, there are likely to be cases where the accusation is indeed false, but the falsity could not be proven. The sometimes high number of cases which were in doubt but not confirmed as false in the studies included here supports this assertion. It is also possible that the researchers categorized some true sexual assault reports as false. It is impossible to guess to what extent either of these classification errors was present in our sample.

Conclusion

As Lonsway et al. (2009) noted, "The issue of false reporting may be one of the most important barriers to successfully investigating and prosecuting sexual assault." It is clear from the findings of the studies in the meta-analysis that some individuals lie about being sexually assaulted. Such false allegations may help trigger suspicion and cynicism directed at individuals who have been sexually assaulted, diversion of scarce resources to false cases, difficulties for police investigators in establishing the veracity of claims, and societal misperceptions about the rate at which people falsely accuse others. These conservative findings show that confirmed false reports of sexual assault occur at a rate of at least 5 %, meaning thousands of people are falsely accused annually around the world. Unfortunately, false reports wreak havoc on the innocent people involved, and often losses to their reputation, livelihood, and mental health are not recoverable even when the falsity of the claim is uncovered.

Future research should examine in more detail the equivocal cases that were not included here. Were there a valid way to test the veracity of these difficult claims, more accurate estimates could be made of the true rate of all false allegations of sexual assault. Other projects might examine rates of confirmed false reporting of sexual assault in areas of the world different from those covered in the meta-analysis, creating the possibility of comparing rates in different environments and cultures. Future studies could also examine change in rates over time or in response to variations in relevant laws or procedures. Calculating inter-rater agreement and using explicit standards for categorizations will help maximize the reliability and validity of judgments of whether reports are false.

References

Articles in meta-analysis are marked with *

Anthony, G., & Watkeys, J. (1991). False allegations in child sexual abuse: The pattern of referral in an area where reporting in not mandatory. *Children and Society, 5*, 111–122. doi:10.1111/j.1099-0860.1991.tb00377.x.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat.

Botelho, G., & Gray, M. (2014, December 6). Rolling Stone apologizes over account of UVA gang rape. *CNN*. Retrieved from http://edition.cnn.com/2014/12/05/us/rolling-stone-uva-apology/.

Chambers, G., & Millar, A. (1986). *Investigating sexual assault: A Scottish office social research study*. Edinburgh: Her Majesty's Stationary Office.

*Clark, L. M., & Lewis, D. J. (1977). *Rape: The price of coercive sexuality*. Toronto: Women's Press.

Firth, A. (1975). Interrogation. *Police Review, 28*, 1507–1508.

Goodwin, J., Sahd, D., & Rada, R. T. (1985). False accusations and false denials in incest: Clinical myths and clinical realities. In J. Garden (Ed.), *Sexual abuse: Incest victims and their families* (pp. 17–26). Boston: John Wright.

Grace, S., Lloyd, C., & Smith, L. (1992). *Rape: From recording to conviction* (Research and Planning Unit paper, No. 71). London, England: Home Office.

Green, A. H. (1986). True and false allegations in sexual abuse in child custody disputes. *Journal of the American Academy of Child Psychiatry, 25*, 449–456. doi:10.1016/S0002-7138(10)60001-5.

Gregory, J., & Lees, S. (1996). Attrition in rape and sexual assault cases. *British Journal of Criminology, 36*, 1–17. Retrieved from http://bjc.oxfordjournals.org.ezp01.library.qut.edu.au/content/36/1/1.long.

Harris, J., & Grace, S. (1999). *A question of evidence? Investigating and prosecuting rape in the 1990s* (Home Office Research Study, No. 196). London, England: Home Office.

*Heenan, M., & Murray, S. (2006). *Study of reported rapes in Victoria 2000–2003: Summary research report*. Retrieved from http://www.dhs.vic.gov.au/__data/assets/pdf_file/0004/644152/StudyofReportedRapes.pdf.

Her Majesty's Crown Prosecution Service Inspectorate/Her Majesty's Inspectorate of Constabulary. (2002). *A report on the joint inspection into the investigation and prosecution of cases involving allegations of rape*. Retrieved from https://www.justiceinspectorates.gov.uk/cjji/wp-content/uploads/sites/2/2014/04/ARC_20020401.pdf.

International Association of Chiefs of Police. (2005). *Investigating sexual assaults: Model policy*. Alexandria, VA: Author. Retrieved from http://www.theiacp.org/LinkClick.aspx?fileticket=iGMHlNOfZzo%3d&tabid=87.

Jordan, J. (2004). Beyond belief? Police, rape and women's credibility. *Criminal Justice, 4*, 29–59. doi:10.1177/1466802504042222.

Kanin, E. J. (1994). False rape allegations. *Archives of Sexual Behavior, 23*, 81–92.

*Kelly, L., Lovett, J., & Regan, L. (2005). *A gap or a chasm? Attrition in reported rape cases* (Home Office Research Study No. 293). London, England: Home Office Research, Development and Statistics Directorate. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=F865334C28EC100143D0BA30405F7130?doi=10.1.1.131.8325&rep=rep1&type=pdf.

Lamb, M. E., Sternberg, K. J., Esplin, P. W., Hershkowitz, I., & Orbach, Y. (1997). Assessing the credibility of children's allegations of sexual abuse: A survey of recent research. *Learning and Individual Differences, 9*, 175–194. doi:10.1016/S1041-6080(97)90005-4.

Lea, S. J., Lanvers, U., & Shaw, S. (2003). Attrition in rape cases: Developing a profile and identifying relevant factors. *British Journal of Criminology, 43*, 583–599. doi:10.1093/bjc/43.3.583.

Levitt, A., & The Crown Prosecution Service Equality and Diversity Unit. (2013). *Charging perverting the course of justice and wasting police time in cases involving allegedly false rape and domestic violence allegations*. Retrieved from http://www.cps.gov.uk/publications/research/perverting_course_of_justice_march_2013.pdf,

*Lisak, D., Gardinier, L., Nicksa, S. C., & Cote, A. M. (2010). False allegations of sexual assault: An analysis of ten years of reported cases. *Violence Against Women, 16*, 1318–1334. doi:10.1177/1077801210387747.

Lonsway, K. A., Archambault, J., & Lisak, D. (2009). False reports: Moving beyond the issue to successfully investigate and prosecute non-stranger sexual assault. *The Voice, 3*(1), 1–11. Retrieved from http://www.ndaa.org/pdf/the_voice_vol_3_no_1_2009.pdf.

MacDonald, J. M. (1973). False accusations of rape. *Medical Aspects of Human Sexuality, 7*(5), 170–193.

Maclean, N. (1979). Rape and false accusations of rape. *Police Surgeon, 29*, 30–38.

*McCahill, T. W., Meyer, L. C., & Fischmann, A. M. (1979). *The aftermath of rape*. Lexington, MA: Lexington Books.

McDowell, C. P. (1990). 'False alligators' and fuzzy data: A new look at crime analysis. *The Police Chief, 57*, 44–45.

McNamara, J. J., McDonald, S., & Lawrence, J. M. (2012). Characteristics of false allegation adult crimes. *Journal of Forensic Sciences, 57*, 643–646. doi:10.1111/j.1556-4029.2011.02019.x.

Peters, J. J. (1976). Children who are the victims of sexual assault and the psychology of offenders. *American Journal of Psychotherapy, 30*, 398–421.

Rumney, P. N. (2006). False allegations of rape. *Cambridge Law Journal, 65*, 128–158. doi:10.1017/S0008197306007069.

Russell, D. E., & Bolen, R. M. (2000). *The epidemic of rape and child sexual abuse in the United States*. London, England: Sage Publications.

Sheridan, L. P., & Blaauw, E. (2004). Characteristics of false stalking reports. *Criminal Justice and Behavior, 31*, 55–72. doi:10.1177/0093854803259235.

*Spohn, C., White, C., & Tellis, K. (2014). Unfounding sexual assault: Examining the decision to unfound and identifying false reports. *Law and Society Review, 48*, 161–192. doi:10.1111/lasr.12060.

Stewart, C. H. (1981). A retrospective survey of alleged sexual assault cases. *Police Surgeon, 28*, 32–35.

Theilade, P., & Thomsen, J. L. (1986). False allegations of rape. *Police Surgeon, 30*, 17–22.

Venema, R. M. (2014). Police officers schema of sexual assault reports: Real rape, ambiguous cases, and false reports. *Journal of Interpersonal Violence*. doi:10.1177/0886260514556765.

Wells, C. (1985). Law reform, rape and ideology. *Journal of Law and Society, 12*, 63–75.

Zona, M. A., Lane, J., & Moore, M. (1996). *The psychology and behaviour of stalkers.* Paper presented at the meeting of the American Academy of Forensic Sciences Annual, Nashville, TN.