NORTHWESTERN UNIVERSITY

Predictive Mechanisms of Placebo Analgesia in Chronic Back Pain Patients

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Neuroscience

By

Sara Elizabeth Berger

EVANSTON, ILLINOIS

December 2016

ABSTRACT

Chronic pain is a prevalent and under-treated condition that remains a mystery to the medical system and a major social and economic problem. Unfortunately, there is no single treatment superior to others for relieving chronic pain. While recent scientific discoveries have provided us with functional and anatomical brain biomarkers of persistence and recovery, as well as identified changes in emotional processing and behavior indicative of underlying neuroplasticity, we still do not understand why some patients respond to certain kinds of medications while others do not, nor have we fully captured the characteristics – physiological and psychological – that make someone with chronic pain likely to respond to a treatment now or in the future. The placebo effect - which describes a psychobiological response to an inactive or sham treatment - provides a unique framework under which to examine these questions and elucidate some of the mechanisms governing treatment response. The 3 studies presented here all use the setting of a randomized control trial (RCT) to investigate contributions of brain biomarkers, personality, memory, and semantic language properties in predisposing chronic low back pain patients to placebo response. The main aim of this dissertation is to identify neurological, psychosocial, and linguistic parameters that predict and/or significantly contribute to clinical placebo analgesia in chronic pain.

In **Study 1**, resting state brain networks indicated that placebo responders and non-responders showed differential functional connectivity of lateral frontal regions with sensorimotor regions and the periaqueductal grey at baseline. Gray matter density, subcortical volumes, and cortical thickness also provided neuroanatomical pre-determinates of placebo response. Psychological profiling of patients using a battery of questionnaires indicated that placebo response in chronic pain patients depended on increased emotional awareness and decreased emotional suppression traits. This unique combination of imaging and personality variables dissociated responders from non-responders in a multivariate model and demonstrated that placebo analgesia in a clinical trial can be predicted from the functional circuitry of the lateral frontal cortex and key psychological characteristics of pain patients.

In **Study 2**, we utilized quantitative language measures to better understand the placebo response in chronic low back pain. As part of the same RCT, we conducted a semi-structured, open-

ended exit interview on all patients who finished the study. Using latent semantic analysis (LSA), we calculated the semantic similarity between each interview to 60 words of interest; 4 of these words explained over 68% of the variance in placebo response. 6 unique combinations of words significantly correlated to the brain measures and personality traits that predicted placebo propensity identified in **Study 1**, and when tested on their own, these language factors also identified novel functional connections corresponding to placebo analgesia. These results indicate that language can be used to identify psychosocial and neurological characteristics mediating placebo response propensity.

Finally, in **Study 3,** we investigated the mismatch between daily pain experience and the memory of this experience, a phenomenon that could influence the placebo response or reported treatment efficacy. As part of the same RCT, we examined the discrepancy between experienced chronic low back pain intensity (from daily ratings recorded using a smartphone application) and self-reported memory of pain over the same time period. The cause of this discrepancy was studied relative to psychometric properties, morphology of the hippocampus, and personality traits. The majority of patients exaggerated their remembered pain, which depended on their strongest experienced pain and their most recent mood. This bias remained stable over 1 year and generalized to both reward memory bias and loss aversion. Shape displacement of the left posterior hippocampus mediated the effects of loss aversion personality on this pain memory bias. In two control groups, morphology of the posterior hippocampus was also stable over 1 year and unperturbed by the development of persistent pain. Importantly, a multi-parameter model accurately predicted pain memory bias in a validation group of patients. These results imply that hard-wired hippocampal learning circuitry and reward-related personality traits determine individuals' exaggeration of their daily experience of pain.

We conclude that (1) the propensity for clinical placebo analgesia in chronic low back pain can be predicted using a combination of personality traits, anatomical brain biomarkers, and functional brain connectivity, (2) language from patient narratives can be used as a surrogate for these predictive markers and as a tool in identifying them, and (3) pain memory bias is a clinically significant phenomenon predetermined by hippocampal morphometry and personality that must be accounted for in clinical trials using retrospective pain assessments.

ACKNOWLEDGMENTS

There are many people I would like to acknowledge, as I could not have completed this dissertation (or graduate school, for that matter) without their guidance and support. First, I would like to thank Mr. Travis Coyle, my Human Anatomy and Physiology instructor in high school. Mr. Coyle's class helped me first identify that I was not only interested in science but also good at it, a surprise to almost everyone. This was because up until my junior year, I had been seen mostly as the "class artist", painting murals around the school and competing in all of the local art shows. It was during a project in Mr. Coyle's class where I discovered my interest in neuroscience specifically, having to write a report and make a presentation on Autism, and I distinctly remember Googling whether a neuroscientist was a "real" profession and how someone would go about becoming such a thing. I always knew I wanted to go to college (and I had a fascination with the human brain by the age of 10), but it was during this project that I was able to mentally connect my interests in biology and the human body with my passion for people and my unquenchable thirst for knowledge, and I decided then and there at age 17 that my focus in college would be something with the brain.

This decision led me to choose Macalester College, which was one of the few schools in the US to offer an undergraduate major in neuroscience. Which leads to the next person on my list of people to thank – Dr. Eric P. Wiertelak, my neuroscience mentor and PI at Macalester. I could write about 100 pages about all of the ways in which Eric inspired me and helped me to succeed, but due to time and space, I'll do my best to sum them here. Dr. Wiertelak taught Brain, Mind, and Behavior, which was my first class during my freshman year. He had a way of teaching the material that was spell-binding: it was interesting, sometimes funny, and most of all memorable, and I excelled in the class. Eric was also my first exposure to the field of pain, as he ran an alternative medicine and pain lab at Macalester. Although I was young and inexperienced with research, Eric saw something in me that was enough for him to provide me with a summer stipend from his NIH grant so that I could learn how to conduct animal research. He taught me that you cannot get what you don't ask for, that hard work pays off, and that my uniqueness was one of my greatest assets. Under Eric's guidance during my 4 undergraduate years, I flourished both in class and in lab. He gave me confidence in my abilities to think critically, freedom to

explore ideas, opportunities to precept and teach, and above all else, a solid neuroscience education which has been a cornerstone that I have returned too countless times in my journey. Additionally, I need to thank Eric in combination with Dr. Corie Hammers for believing in me enough to let me pursue my honors thesis, an endeavor that I was crazy enough to try on my own.

I am very appreciative of the many people I met and collaborated with at Northwestern University. I have to first start with the Apkarian lab and thank them for the 6.5 years of support. I am grateful for all the knowledge and advice I received either via the over-300 lab meetings and journal clubs I attended or through my insightful interactions with lab members and collaborators past and present: Taha Abdullah, Christina Arroyo, Marwan Baliki, Qiyuan Bao, Marivi Centeno, Pei-Ching Chang, Don Davis, Melissa Farmer, Colleen Fitzgerald, Mariam Ghantous, Javeria Hashmi, Kristi Hermann (Williams), Lejian Huang, Beth Hunt, Hridi Hussain, Rami Jabakhanji, Ali Mansour, Amelia Mutso, Elle Parks, Todd Parrish, Bogdan Petre, Sarah Pollema (Mays), Wenjie Ren, Thomas J. Schnitzer, Bangli Shen, Pascal Tétreault, Souraya Torbey, Étienne Vachon-Presseau, Xu-Hong Wei, Brian Williams Jr., Binbin Wu, Lily Yang, and Renita Yeasted. It has been a pleasure working with and learning from all of you. I specifically want to thank Marwan, Ali, Alex, and Étienne for always taking the time to answer my questions; Alex (and his wife, Janice), Elle, and Marivi for always being available for me when I needed them and for giving me sound life advice; Chris, Kristi, Amelia, and Bogdan for always being able to distract me and make me laugh; Taha for all of his dedication and persistence when finishing the study; Lejian for all of his involvement and hard work in our quality control and preprocessing pipelines (we are all indebted to him); Bogdan for the introducing me to good German beer, followed by late night or early morning work sessions; Rami for always being so kind to me and upbeat in general (you make the lab environment so enjoyable); Marivi for being the best office mate ever (M&Ms and Zumba, need I say more?); Tom and Renita for all of the clinical assistance and their unwavering support and belief in my capabilities as a coordinator; and of course, Étienne, who was the co-leader of these studies (a.k.a., my partner in crime), and one of the best mentors, colleagues, and friends I could have asked for during my graduate school experience. Most of all, I am extremely grateful for my advisor and PI, Vania Apkarian, who accepted me into his lab and gave me such an incredible and indispensible opportunity to learn and grow as a scientist. Without his guidance and support, I would not have been able to cultivate such a rich and diverse skill set

encompassing basic animal research, neuroimaging, qualitative interviewing, mathematical modeling, and clinical coordination. I leave his lab as both a stronger scientist and a stronger person because of this experience. On a more personal note, I also want to thank him for all of his book recommendations and re-giftings over the years. As a tradition that is part of his annual holiday party, lab members bring books to share with one another and Vania always partakes, sometimes buying a particular favorite for everyone who comes. From the 6 (going on 7) holiday parties attended during my graduate career, I have amassed quite the collection of books, ranging from Sartres to network theory, and if it weren't for Vania giving all of us a copy of Daniel Kahneman's "Thinking Fast and Slow", I wouldn't have had the inspiration for **Study 3** presented here. Dr. Apkarian's wide breadth of knowledge and appreciation for literature is truly a gift that keeps on giving (and I hope he continues on with this tradition in my absence).

I'd also like to thank my NUIN study group, who became my first friends in Chicago and helped me conquer two years of classes and the Qualifying Exam (Kristan Leech, Lisa Qu, Kenneth Weber, and Mike Schied). I am also incredibly grateful for Dr. David Schneeweis, Dr. C.J. Heckman, Dr. Mark Bevan, and Dr. Megan Crowley-Matoka for all of their support, connections, and advice over the last few years. I of course also want to thank my thesis committee members for their support and constructive feedback over the last 3 years: Dr. Jay Gottfried, Dr. Guillermo Cecchi, Dr. Rick McGee, and Dr. Jamie Griffith. I feel so lucky to have had such a brilliant and interdisciplinary team of scientists guiding me through this extensive collection of projects, and I think my dissertation was substantially enriched by their participation.

I have some of the coolest, smartest, and kindest people as best friends. Thank you to fellow neuroscience PhD students, Rachel Diamond and Dave Harris, for the commiseration sessions that allowed me to vent and remain motivated after leaving Mac. Thank you to Elissa Butler, Colin Gerber, Pros Seng, and Chloe Mirzayi for always being one Skype call away and giving me a word of encouragement when needed.

And of course, I need to thank my family who is my rock and my foundation – they put things in perspective and remind me what is important in life. I could not have done any of this without their unwavering love and support. Thank you to my cousin, Jessica Berger, and her partner, Ayesha Hassan, for the home-cooked meals and ice cream, movie nights, a different place to stay, and of course the best

distraction in the world – babysitting Rafi. Thank you to my 4 older siblings and their families for always checking in on me via text message, emails, or Line. Words cannot express how extremely grateful I am for my parents, who have never stopped believing in me, and who have done nothing but encourage me throughout this whole process. To my moms: thank you for being one of the initial inspirations for my decision to become a neuroscientist. If it weren't for the two of you coming out and being open about who you were, I am not sure I would have ever been so interested in understanding the brain (sexual orientation, nature versus nurture, genetics, choice and decision making, etc) from such an early age. I also thank you so much for giving me access to science books early on (even if they were above my reading level), for the high academic standards you set for me and my brothers and sisters, and for your financial support in college (and later graduate school) that helped me pursue my interests. I also, of course, thank you for your incessant support through weekly pep talks and/or listening to me vent for hours at a time. Also, thank you for being the strongest source of stability in my life – whether I need a shoulder to cry on, someone to laugh with, or someone to sit in silence with, you both have always been there. To my dad: thank you for being one of my biggest cheerleaders (hearing the excitement or pride in your voice every time I reached a milestone or made progress kept me motivated and is something I will always remember and cherish). Thank you as well for always taking the time to listen to me express my worries or frustrations, and thanks as well for genuinely helping me be open to any possible outcome during this journey without fear of judgment or disappointment. Thank you for reminding me of Aunt Susan's words of wisdom, including "Trust the process" and "It doesn't matter". Most of all, thank you for all the laughs – your chuckle could brighten my day no matter what else was going on. You have been through so much over the last couple of years, but despite this, I never once felt like you weren't on the sidelines, telling me to "run past the line". Finally, there is my life partner, Evan, the first person I met in Chicago. To Evan: Thank you for sticking with me through thick and thin, ups and downs, and good and bad times throughout this entire process. Thank you for listening to my "Sara rants", for indulging and even participating in my nerd-out sessions, for calming me down, and for giving me the space necessary to complete the journey while also providing me with a source of distraction when I needed it – basically, for letting me be me. Most of all, thank you for helping me become a better person in the last 6.5 years and for being patient as I underwent a lot of changes as part of being in graduate school and going

through my mid 20s (oie vey). We have had our substantial challenges, but through it all, there is no one else I'd rather have gone through this with than you.

Last but not least, I want to thank my furry sidekick, Haraway, who has been my constant companion throughout this journey. From a scraggly little kitten in North Dakota to an opinionated fluff ball in Chicago, I don't know how I would have managed to stay sane without her non-judgmental and unconditional love, endless supply of laughs, much needed playtime distractions, and perfectly timed kitty cuddles. She was with me for this entire journey, and she will join me for the next adventure.

LIST OF ABBREVIATIONS

amy(g)	amygdala
App	application; short for smartphone application
BOLD	blood-oxygen level dependent
CBP	chronic back pain
CFS	chronic fatigue syndrome
CSF	cerebral spinal fluid
CON	controls; another term for healthy participants
CRPS	complex regional pain syndrome
DA	dopamine
(d)ACC	(dorsal) anterior cingulate cortex
DMN	default mode network
DTI	diffusion tensor imaging
EEG	electroencephalography
FM	fibromyalgia
fMRI	functional magnetic resonance imaging
GM(D)	grey matter (density)
hipp	hippocampus
IBS	irritable bowel syndrome
ICA	independent component analysis
IGT	Iowa Gambling Task
LFC	lateral frontal cortex
LOOCV	leave-one-out cross validation
LSA	latent semantic analysis
MEG	magnetoencephalography
mPFC	medial prefrontal cortex
NAc	nucleus accumbens
NB	Naïve Bayes(ian)
NSAIDs	non-steroidal anti-inflammatory drugs
OA	osteoarthritis
OFC	orbitofrontal cortex
PAG	periaqueductal grey
PCA	principal component analysis
PET	positron emission tomography
PROs	patient reported outcomes
QOL	quality of life
RCT	randomized control trial
RF	reticular formation
RS(N)	resting state (network)
RVM	rostral ventromedial medulla
S1M1	primary sensorimotor regions
SBP(p)	subacute back pain (persisting)
SVD	singular value decomposition
SVM	support vector machine
UCPPS	urological chronic pelvic pain syndromes (sometimes called CPP)
VPL	ventral posteriolateral nucleus
WDR	wide-dynamic range
5-HT	serotonin

TABLE OF CONTENTS

LIST OF STUDIES

This thesis is based on the following studies, which are referred to by their respective numbers:

1. **The neurobiological and psychosocial mechanisms predicting clinical placebo response in chronic pain.** Etienne Vachon-Presseau*, Sara E. Berger*, Taha B. Abdullah, Bogdan Petre, Lejian Huang, James W. Griffith, Thomas J. Schnitzer, A. Vania Apkarian. (*in progress; to be submitted to Science, November 2016).*
2. **Semantic language properties underlying the placebo response in chronic pain**. Sara E. Berger, Guillermo Cecchi, Etienne Vachon-Presseau, Taha B. Abdullah, A. Vania Apkarian. *(in progress, not yet submitted)*.
3. **Memory of chronic pain is biased by left posterior hippocampus morphology.** Sara E. Berger*, Etienne Vachon-Presseau*, Taha B. Abdullah, Alex T. Baria, Thomas J. Schnitzer, A. Vania Apkarian. *Submitted to Current Biology (November 2016)*.

* = co-first authors (contributed to study equally)

Note: One first-author publication is required as part of the Northwestern University Interdepartmental Neuroscience (NUIN) program's official requirements for the degree. I have met this requirement, but this already published paper was not a main component of my dissertation and thus is not explained in detail. It is, however, referenced throughout the thesis, and I provide a citation here if anyone is interested:

Berger SE*, Baria AT*, Baliki MN, Mansour A, Herrmann KM, Torbey S, Huang L, Parks EL, Schnitzer TJ, Apkarian AV. (2014). **Risky monetary behavior in chronic back pain is associated with altered modular connectivity of the nucleus accumbens**. *BMC Research Notes*, 7: 739. doi: 10.1186/1756-0500-7-739.

LIST OF FIGURES

LIST OF TABLES

PREAMBLE

"One of the principal qualities of pain is that it demands an explanation."
~ Ann Carson [1].

This dissertation is structured in the following manner. **Chapter 1** is an introduction, the goal of which is to provide the reader with a general understanding of the key concepts discussed in the thesis, as well as a broader sense of the current gaps in knowledge in the fields of pain and placebo analgesia. From this understanding, I hope to convey the significance of the questions asked and the importance of the aims addressed in this thesis, as I believe all three studies discussed have the potential to impact the neuroscience community at many levels, from basic science research to real-world clinical applications. **Chapter 2** provides an overview of the methods utilized; although three separate research agendas (studies) are covered in this thesis, data from all three studies come from the same clinical trial and thus largely overlap in methodology. However, in the instances where there are distinctions, I have indicated the study number where applicable. **Chapter 2** also describes how we analyzed the data; this section of the chapter is organized by study number; some redundancy should be expected as similar neuroimaging analyses and model development occurred between studies. **Chapters 3-5** provide the key findings of each of the studies and a discussion about their meaning; where appropriate, brief descriptions of the main questions and analyses used are given in order to remind the reader of the primary aims of the thesis. Finally, in **Chapter 6,** I discuss some of the limitations of the current research, the ethical implications of the research, and future directions that could be explored given the results presented here.

CHAPTER 1: INTRODUCTION

Section 1: What makes pain?

This thesis's main focus is on chronic low back pain and placebo-mediated relief of this pain. However, in order to understand these more complex phenomena, one must first be familiar with the definition of pain and the biology of the nociceptive and pain systems. The International Association for the Study of Pain (IASP) define pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage"[2]. This definition is significant because it highlights that pain is simultaneously physiological and psychological, automatically implicating the important and active role of the brain in pain perception. For hundreds of years, we have surmised at the basic pathways underlying the pain experience. Descartes' description of pain transmission in his treatise De l'homme [3], although inaccurate, depicts the simple premise that the sensation of pain is directional and transmitted from the body to the brain (**Figure 1**). This idea helped shape early research on the topic and aided in guiding later seminal investigations mapping the peripheral nervous system, spinal cord, brainstem, and cerebral cortex. The sections below briefly describe the main components of these systems and pathways, followed by a section that highlights some of the caveats in the ways in which these systems have been investigated and framed. Unless otherwise noted, all anatomical and physiological descriptions have been adapted from Nolte's "The Human Brain" [4] and Netter's "Atlas of Neuroscience"[5].

Figure 1: Descartes' treatise De l'homme. Descartes explained his theory of pain as follows: "If for example fire comes near the foot, minute particles of this fire, which you know move at great velocity, have the power to set in motion the spot of skin on the foot which they touch, and by this means pulling on the delicate thread which is attached to the spot of the skin, they open up at the same instant the pore against which the delicate thread ends, just as by pulling on one end of a rope one makes to strike at the same instant a bell which hangs at the end."

1.1 Nociception

The term "nociception" was first coined in the 1900s by Sherrington, who discovered that there were sensory receptors in the periphery (called nociceptors) that were activated by noxious or potentially noxious stimuli. Most nociceptors are free nerve endings, meaning that unlike other specialized receptors that exist throughout subcutenous and deep tissue (such as Merkel disks or Pacinian corpuscles), they are un-encapsulated and exhibit polymodality. Polymodality indicates that they respond to a wide variety of peripheral perturbations, such as suprathreshold temperatures, suprathreshold mechanical stimuli (like pressure, touch, and stretching), and various chemosensations (including inflammatory chemicals released locally). Nociceptors can be divided into two classifications based on their response modality and their rate of conduction. Aδ fibers respond primarily to extreme pressures and temperatures; they are smaller in diameter and thinly myelinated, enabling them to conduct signals relatively quickly at around 5-30 m/sec. In contrast, C fibers respond even more broadly to high-intensity mechanical, thermal, or chemical stimuli. These fibers are also small in diameter but they are unmyelinated, meaning that they are not able to conduct signals as quickly as Aδ fibers, getting to speeds only around 1 m/sec. The difference in conduction velocity is important because it is what gives rise to perceived first and second pain. First pain from Aδ fibers is the immediate sensation felt after an injury that often is described as "sharp"; second pain from C fibers is the sensation that comes a few seconds later, described as "throbbing", "burning", or "aching".

1.2 Spinal cord transmission and ascending pathways

The afferent fibers from peripheral nociceptors enter the spinal cord via the ipsilateral dorsal root (lateral division). From here, they terminate primarily in marginal layer (lamina I), substantia gelatinosa (lamina II), and the neck (lamina V) of the dorsal horn. Many of the neurons in lamina I are nociceptive-

specific, meaning that they only respond to noxious stimuli, while other second-order cells are wide-dynamic-range (WDR) neurons that respond in a graded fashion to both non-noxious and noxious stimuli. In lamina II, some of the nociceptor afferents synapse with excitatory and inhibitory interneurons, which convey information to other cells in other laminae and help regulate nociceptive input to third-order projection neurons. In lamina V, Aδ fibers often synapse with WDR cells and are thought to contribute, in part, to the phenomenon of pain referral [6]. Projection neurons from these 3 laminae decussate the midline, ascend a couple of segments in Lissauer's tract, and then synapse in the contralateral side of the spinal cord. From here, they ascend to the brain via different pathways in a large collection of tracts collectively referred to as the anterolateral pathway. As these tracts ascend, new fibers join on their medial edges so that much of the system becomes and remains somatotopically organized.

In the anterolateral system, there are three major ascending tracts along which nociceptive input can travel: spinoreticular, spinothalamic, and spinomesencephalic pathways. Each pathway has different conduction velocities and different areas of termination, distinguishing their functions. The spinoreticular pathway, which has fibers originating mainly from the intermediate gray matter, ascends in the ventrolateral spinal cord and synapses primarily on reticular cells in the medulla without a somatotopic arrangement; it also terminates in parts of the thalamus (intralaminar and other regions). These medulla reticular cells receive input from other sensory systems and project to reticular formation, thalamus, and limbic system. This tract is thought to contribute to changes in levels of attention in response to painful stimuli. Next is the spinothalamic tract, which is viewed as the principal pathway for somatosensation. It has neurons originating from lamina I and V, and as the name suggests, its projections end in the thalamus - more specifically, the ventral posteriolateral nucleus (VPL). These fibers are thought to play a special role in conscious awareness of the quality of a stimulus and its location. Projections from the VPL are widespread, terminating in the postcentral gyrus (primary somatosensory area), the insula, and other cortical areas. Finally, the spinomesencephalic tract carries information mainly from lamina I and V and terminates in both the reticular formation and the periaqueductal gray (PAG); this pathway is thought to have a role in intrinsic pain-control mechanisms. Some projections from this tract also go through the spinoparabrachial tract, synapsing on parabrachial nuclei which then send information to the amygdala,

potentially contributing to the emotional component of the pain experience. All of the fibers within this anterolateral system are intermingled with or adjacent to each other in the cord spinal cord.

In addition to the anterolateral system, there are other pathways that contribute to nociceptive transduction and processing, including the spinohypothalamic, spinocervical, and spinocerebellar tracts; however, they are not relevant for this the topic of this dissertation and are therefore not discussed here.

1.3 Descending modulatory pathways

There are also special circuits that control pain from the top-down (brain to spinal cord). The major hub in these pathways is the PAG, which receives input from higher brain areas including the thalamus, hypothalamus, amygdala, and cortex. In 1976, Mayer and Price provided evidence of PAG's involvement in pain modulation, showing that the stimulation of this region produced profound and selective analgesia[7]. PAG neurons make excitatory connections with nuclei in the rostral ventromedial medulla (RVM), whose cells then project back down to the spinal cord, creating inhibitory connections on neurons in laminae I, II, and V of the dorsal horn, which in turn disrupt ascending nociceptive signals. Importantly, some of these mechanisms are driven by the various endogenous systems, including seritonergic, noradrenergic, and opioid pathways. Stimulation of either the PAG or RVM causes the release of serotonin (5-HT) in the spinal cord and increased levels of norepinephrine in the cerebral spinal fluid (CSF), both eliciting antinociception. Regarding opioid involvement, the stimulation of PAG stimulation, as well as analgesic doses of morphine (a opioid agonist), has been shown to cause a subtype of RVM neurons called "off cells" to fire continuously; these cells block pain perception and inhibit spinal withdrawal reflexes away from noxious peripheral input [8, 9]. Additionally, activation of inhibitory interneurons in lamina II of the spinal cord causes release of enkephalin or dynorphin (endogenous opioid neurotransmitters) which bind to receptors on the incoming C and Aδ fibers carrying information from the periphery, modulating or inhibiting the nociceptive signal [10]. These descending modulatory pathways are thought to be crucial in placebo response mechanisms in pain, with the idea being that the emotional and internal state of a person can influence incoming nociception perception through these top-

down processes. Opioid involvement in particular is thought to play a role, as placebo analgesia has been shown to to be blocked by naloxone, an opioid antagonist [11, 12].

1.4 The role of the brain: from brainstem to thalamus to cortex

Thus far, we have only covered nociception, which is not the same as pain. In order for pain – a strictly conscious experience – to be perceived, the brain must be involved. Projections from the spinal cord primarily (though not exclusively) terminate in the brainstem or the thalamus, both of which are responsible for integrating nociceptive inputs. As mentioned above, the reticular formation (RF) is a key player in nociceptive and aversive drive, and it is seen as the central core of the brain stem. The RF is diffusely organized, meaning that its pattern of connectivity is characterized by both convergence and divergence, and cells within it can respond to a several sensory modalities or to stimuli applied all over the body. Importantly, the RF is thought to be important in regulating and integrating inputs from both the arousal system and the autonomic system, particularly through connections with the thalamus and the limbic system (including the amygdala).

Head [13] first postulated that the pain center of the brain was the thalamus. Although we now know that pain processing is spread throughout the brain, indeed the thalamus is still the main relay station for nociceptive input. The thalamus has been divided into functional and anatomical subdivisions based on its connections with the spinal cord, and several nuclei are important for nociceptive signal processing, including the lateral and medial groups and those nuclei receiving input from the spinothalamic tract. Moreover, the thalamus has a diverse range of subcortical and cortical inputs (e.g., the cerebellum, basal ganglia, hippocampus, and parietal lobes) and outputs (e.g., cingulate cortex, motor regions, somatosensory cortex, insula, and prefrontal cortex); this interconnectedness makes it extremely important for nociceptive transmission and signaling of pain.

Beyond the thalamus, nociception and pain perception become even more complicated and delocalized. With the advancement of neuroimaging techniques spanning PET (positron emission tomography), EEG (electroencephalography), fMRI (functional magnetic resonance imaging), and MEG (magnetoencephalography) studies, we have begun to refine our knowledge of how pain is represented in the cortex. Common cortical areas found to be active during nociceptive stimulation and acute pain

include the primary and secondary somatosensory cortex (S1,S2), frontal regions, anterior cingulate cortex (ACC), and insula (**Figure 2**); other regions are also sometimes present depending upon the experiment (e.g., the NAc, amygdala, hippocampus, and posterior parietal regions) [14, 15]. Together, these areas make up a dynamic network modulated by psychological and physiological processes that shape the perception of pain.

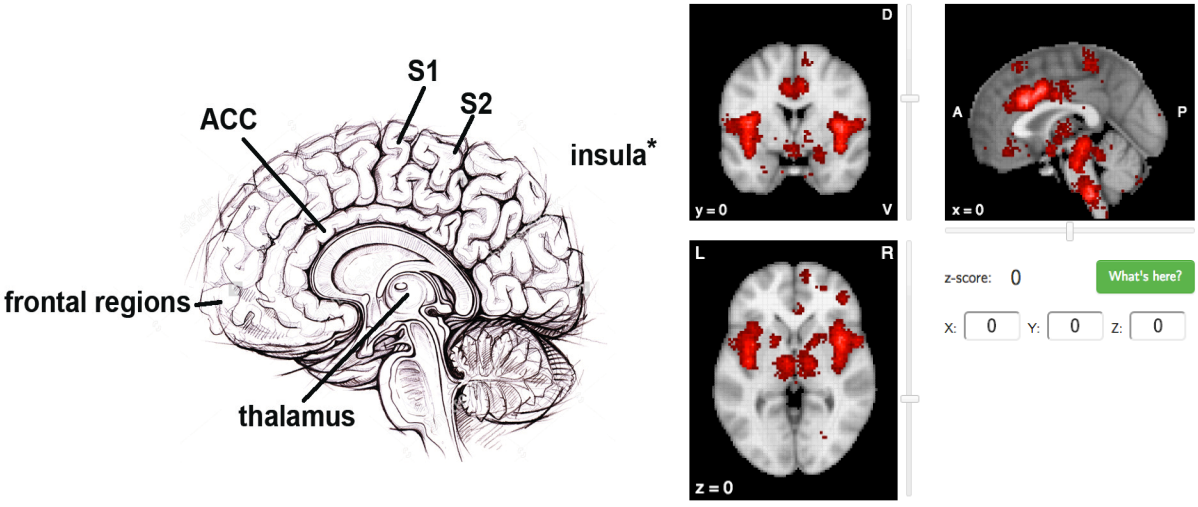


Figure 2: Brain regions associated with nociception and acute pain perception. Left: key brain areas identified by multiple studies to be involved in nociceptive and acute pain processing are shown. Insula (*) is also involved, but is not depicted here as it is lateral to the midline. Right: Neurosynth reverse-inference map is shown for the term "pain"; regions shown in red represent the regions consistently activated in response to acute painful stimuli across a meta-analysis of 420 studies.

Importantly, these areas are involved in the multi-dimensional conscious experience of pain that includes not only sensory components but also affective, motivational, and cognitive aspects as well [16]. The somatosensory cortex is the primary region involved in processing the sensory aspects of pain; both contralateral S1 and bilateral S2 receive input from the thalamus about noxious and non-noxious stimuli [17, 18]. Cells in S1 are thought to encode stimulus intensity and location, whereas cells in S2 are thought to aid in recognizing a stimulus as noxious or painful, in addition to encoding for spatial, temporal, and intensity aspects of the pain [19]. S2 cortex is consistently found in neuroimaging studies of pain perception, but S1 is not, perhaps because it is highly susceptible to cognitive and emotional interference [14, 20].

The frontal regions (better known as the frontal association cortices) are involved in executive control and higher-order cognitive processes in general. In acute pain, these areas are thought to influence expectations of pain and attentional drive, as well as modulate underlying mood and emotional states (which have indirect effects on pain perception [21]. Frontal regions have also been shown to be affected when acute pain becomes pathological in nature, a topic which will be addressed in a later section.

The ACC and insula are both considered part of the classic limbic system [22] and are regions that are highly involved in the affective motivational dimensions of pain. While both the ACC and insula are consistently activated during pain studies, their functions are diverse and not specific to pain. The ACC receives indirect nociceptive input from medial thalamic neurons and direct nociceptive input from the periphery. Activation to the painfullness and unpleasantness of noxious stimuli have been seen in many studies [16]. The insula receives nociceptive information through direct thalamocortical input and is hypothesized to encode sensory aspects of pain, as well as integrate sensation, body position, and environmental context [19]. It is also thought to be involved in general aspects of embodiment [23, 24] and pain anticipation [25].

All of these regions are involved in daily cognitive functioning, and because of these shared resources, interference between pain and cognitive processes may result in many of the areas listed above. This potential competition may be one of the reasons why chronic pain patients show emotional, working-memory, and other cognitive deficits (discussed in a later section).

1.5 Types of pain

In general, the literature differentiates 3 kinds of pain experiences based on mode of activation: nociceptive, inflammatory, and neuropathic. Nociceptive pain is pain caused from directly stimulating nociceptors in the periphery and is seen as a key early warning system of the body. Activation of these nociceptors does not necessarily mean that an injury has taken place; however, in the event that tissue damage has occured, the body switches from a state of guarding against harm to a state of healing. This is when we feel inflammatory pain, which promotes the healing process via changing peripheral sensitivity. Immune system cells, glial cells, and surrounding tissues all release various chemicals and

proteins into the surrounding milieu; this “inflammatory soup” includes histamine, prostaglandins, acetylcholine, substance P, and pro-inflammatory cytokines [26] that all contribute to allodynia and hyperalgesia by sensitization of nociceptors. Because these inflammatory chemicals degrade quickly, inflammatory pain requires constant stimulation to be sustained; this is why its presence results in healing because it promotes decreased movement and tactile contact of the injured region. Once the tissue has healed, normal inflammatory pain will cease. Finally, neuropathic pain is pain caused from direct injury to peripheral nerves via diseases, cancer, infections, traumatic injury, or certain kinds of medication (e.g., chemotherapy). Unlike nociceptive and inflammatory pain, neuropathic pain does not always disappear after the inciting event, nor does it readily respond to traditional analgesics. Importantly, both prolonged neuropathic and inflammatory pain can give rise to pathological chronic pain disorders that severely disable individuals and disrupt their quality of life (for example, diabetic neuropathy and rheumatoid arthritis).

1.6 Caveats

Although peripheral, nociceptive, spinal, and brain components of the pain system are all vitally important, the significance of each of these in a person’s actual pain experience remains unknown and a source of contention. More specifically, the idea that nociception gives rise to pain through the pathways mentioned above is somewhat debatable, especially to the extent that that nociception and pain do not necessitate that one always occurs in the presence of the other. As mentioned at the start of this section, the IASP definition of pain emphasizes subjectivity and conscious experience. But interestingly, nociceptors can be (and usually are) active in the absence of pain. In fact, for the large majority of our lives, nociceptors are continuously working to correct our behaviors or postures and subvert damage so that we do not experience pain more often; such control is largely unconscious or at the very least subconscious[27]. Additionally, there is ample evidence that conscious acute pain is highly malleable by mood and context such that nociceptive barrage does not always translate to a certain kind or level of brain activity or to a stereotypical type of pain perception. The value and level of pain can change depending on the motivational state of the organism. As one example, many soldiers with severe wounds from fighting do not report pain or suffering until hours or days later despite considerable peripheral

nociceptive activity [27]. Although no one would argue that conscious pain is usually driven by nociceptive inputs in most cases, the stimulation of nociceptive afferents in the periphery does not necessitate the experience of pain or the production of a pain behavior. Likewise, experience of pain does not necessitate currently active nociceptors; in some chronic conditions (such as fibromyalgia and chronic fatigue syndrome, CFS) pain is reported with little evidence of nociceptive involvement in the peripheral system [28], and in other cases, pain is idiopathic or spontaneous and has no inciting injury or underlying systemic disease. Moreover, activity in any one of the brain regions mentioned above does not mean that there is pain or nociception present, since many of these areas are multimodal or involved in a diverse range of functions.

Thus the idea that pain is "directional" or caused from nociceptive activity in the peripheral nervous system, whose signals make their way up the spinal cord to the brain, is a narrow viewpoint at best. In addition to ignoring the roles of emotion, memory, motivation, and attention in gating present and incoming pain and nociceptive signals, this viewpoint also does not account for cultural and context mediations of pain experience [27], which also likely influence the expression of pain and other pain behaviors. These caveats are important to keep in mind for the thesis for two reasons. First, its is very apparent that the brain in chronic pain is not equivalent to the brain in acute pain, and therefore studies which investigate neural activity associated with acute pain are not necessarily comparable to those studying chronic pain. Second. If the mechanisms underlying chronic pain are complex and not fully understood, it follows that alleviation of chronic pain is also likely to be complicated and must involve modulation of circuitry that is not only important for acute pain but also circuitry critical in the transition to and stability of a chronic state.

Section 2: The problem of chronic pain

The mechanisms above are primarily relevant for acute pain (which is short in duration and dissipates as part of the healing process). While acute pain is almost always initially advantageous to an organism (in that it signals actual or potential tissue damage and aids in important avoidance learning), over time it can be maladaptive at both an individual and societal level. Today, pain remains one of the primary reasons why people seek healthcare. It is estimated that around 50 million Americans are either

partially or totally disabled by some type of pain on any given day. This pain, if left untreated, can become long-lasting or permanent, and it can lead to other deleterious co-morbidities such as depression, insomnia, lowered immune function, impaired cognitive functions, and decreases in mobility, all of which can seriously impact quality of life. The resulting condition of chronic pain is defined by IASP as pain that persists for more than 3 months [29], with an operation definition extending up to 6 months depending on the condition [30], and chronic pain's long-term suffering and consequent changes in behavior represent an even larger and more complex problem to society. The Institute of Medicine of the National Academies (IOM) released a comprehensive report regarding chronic pain a few years ago (www.iom.edu, released on 6/29/2011). They summarize: "Chronic pain affects at least 116 million American adults – more than the total affected by heart disease, cancer, and diabetes combined. Pain also costs the nation up to 635 billion dollars each year in medical treatment and lost productivity." Worldwide, chronic pain is estimated to affect around 15% of the population [31].

In addition to its prevalence and cost, chronic pain also remains inadequately treated. While pain that is the consequence of an acute injury or inflammatory process can usually be readily alleviated via simple anti-inflammatory analgesics, the same cannot be said of chronic pain – despite years of research, no consistently and generally effective therapies for chronic pain have been identified. Those medications that are currently available often have long-term side effects and/or addictive properties [32], or only provide modest improvements that are not sufficient to achieve clinical meaningful amelioration of disability [33]. Additionally, the majority of these analgesics simply do not work. Over a third of individuals with chronic pain define their pain as severe [34] and over 40% of those suffering from chronic pain are not satisfied with their current care [35]. One report estimated that drugs on the market are inefficient in about 70% of patients in pooled analyses of placebo-controlled trials [29]. These findings highlight the fact that pain treatments created today are usually symptomatic- rather than disease-modifying, which is probably one of the reasons why they fail [36].

2.1 Chronic low back pain

One of the most prevalent persistent pain conditions is chronic low back pain (CBP), estimated to affect 15-45% of Americans at any given time [37] and between 70-85% of adults during their lifetimes

[38]. CBP is the most common cause of limited activity in people under the age of 45, and it is the second-most frequent reason for physician visits [39]. Additionally, over 149 million days of work per year are lost due to this condition alone [40]. It was initially suggested that 95% of individuals recovered after an initial back pain episode [41], but more recent evidence suggests that this number is closer to 50% [42].

Unfortunately, research shows that the prevalence of CBP is increasing in the US population, and will likely continue to rise as the population ages [43]. And although treatment with the standard of care - non-steroidal anti-inflammatory drugs (NSAIDs) - or antidepressants has shown some alleviation of early-stage (acute to subacute) low back pain, these modest improvements usually do not translate into clinically effective pain relief for individuals with late-stage (chronic) low back pain. In fact, the World Health Organization Advisory Panel has concluded that there is no single treatment superior to others for relieving CBP [44].

2.1.1. Neurobiology of CBP

CBP is characterized by predisposing neuroanatomical and neurophysiological characteristics, and the actual transition from acute to chronic back pain also causes dynamic changes to functional brain properties. Over the last decade, researchers have demonstrated that back pain that persists past an acute or sub-acute stage involves systemic alterations of circuitry between frontal cortical areas (involved in higher functions such as executive control, learning, and appraisal) and limbic regions (involved in emotional processing and memory consolidation), including the nucleus accumbens (NAc), amygdala (amyg), and hippocampus (hipp). These changes include (1) enhanced functional connectivity between medial prefrontal cortex (mPFC) and NAc [42], (2) increased functional connections within the dorsomedial PFC-amyg-NAc circuit [45], (3) decreases in the strength of functional connections between the hipp. and mPFC [46], (4) distinct changes in the frequency and phase dynamics of the default-mode network (DMN) [47], (5) and an overall shift in activity from nociceptive and sensory processing regions to emotional processing regions over the time [48]. **Figure 3** summarizes some of these key findings.

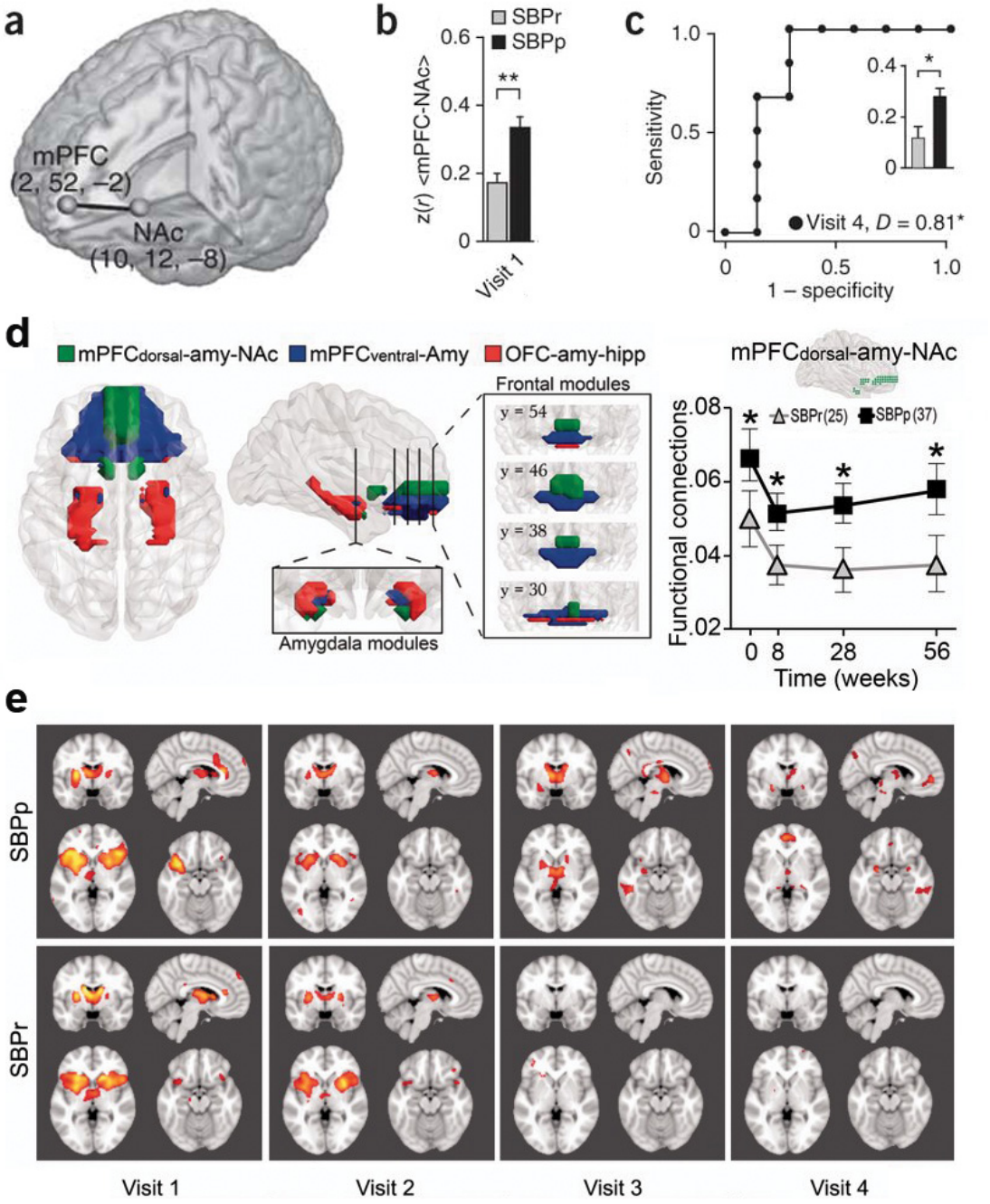


Figure 3: A summary of functional brain connectivity findings in chronic pain patients. The mechanisms underlying the transition from subacute (SBP) to chronic back pain was investigated in a previous longitudinal analysis. **A**. Functional connectivity between the medial prefrontal cortex (mPFC) and nucleus accumbens (NAc) is shown. **B.** This link was different between patients with subacute back pain who either recovered (SBPr) from their pain over the course of a year or persisted in their pain (SBPp) and became chronic over a year; mPFC-NAc functional connectivity was higher in SBPp than SBPr at baseline. **C.** Moreover, mPFC-NAc functional connectivity was also shown to be *predictive* of the transition from acute to chronic pain, predicting persistence 1 year later (visit 4) with 83% accuracy in a discovery group (not shown); this predictive capacity was validated in a separate group of SBP patients (n=13, shown here). D = area under the ROC curve (81% accuracy) [42]. **D.** Functional communities,

identified from white matter connections to limbic regions in SBP patients, are shown – a modularity analysis split them into 3 separate communities dorsal-medial PFC (dmPFC) – amygdala (amy) – accumbens (NAc); ventral-medial PFC (vmPFC) – amy; and orbitofrontal cortex (OFC) – amy – hippocampus (hipp). SBPp showed higher incidence of functional connections within the dmPFC-amy-NAc community than SBPr, a finding that was stable for 3 years [45]. **E.** Group average activation maps for SBPr and SBPp groups at 4 study visits spread out over 1 year. Both groups show activation within acute pain regions for the first 2 visits, encompassing bilateral insula, thalamus, and ACC. However, SBPr show no significant activity for the last 2 visits, whereas SBPp show increased activation in mPFC and amygdala, suggesting the transition to chronicity involves a shift from acute pain processing regions to those involved in emotional processing [48].

Anatomical predispositions within this same circuitry are also present, both in terms of the morphometry of these regions and regarding the physical connections between the limbic networks and frontal regions. Gray matter atrophy of the thalamus and dl-PFC [49], smaller amygdala volume [45], smaller hippocampal volume [50, 51], abnormal regional white matter integrity in structural connections between mPFC and NAc [52], and increased numbers of white matter tracks within the frontolimbic circuit [45] are all either attributes of CBP or risk factors in subacute back pain (SBP) patients who go on to develop CBP. These properties account for a large percentage of the variability in pain persistence, in some cases even predicting chronicity [27, 30, 53-55]. **Figure 4** summarizes some of these key findings.

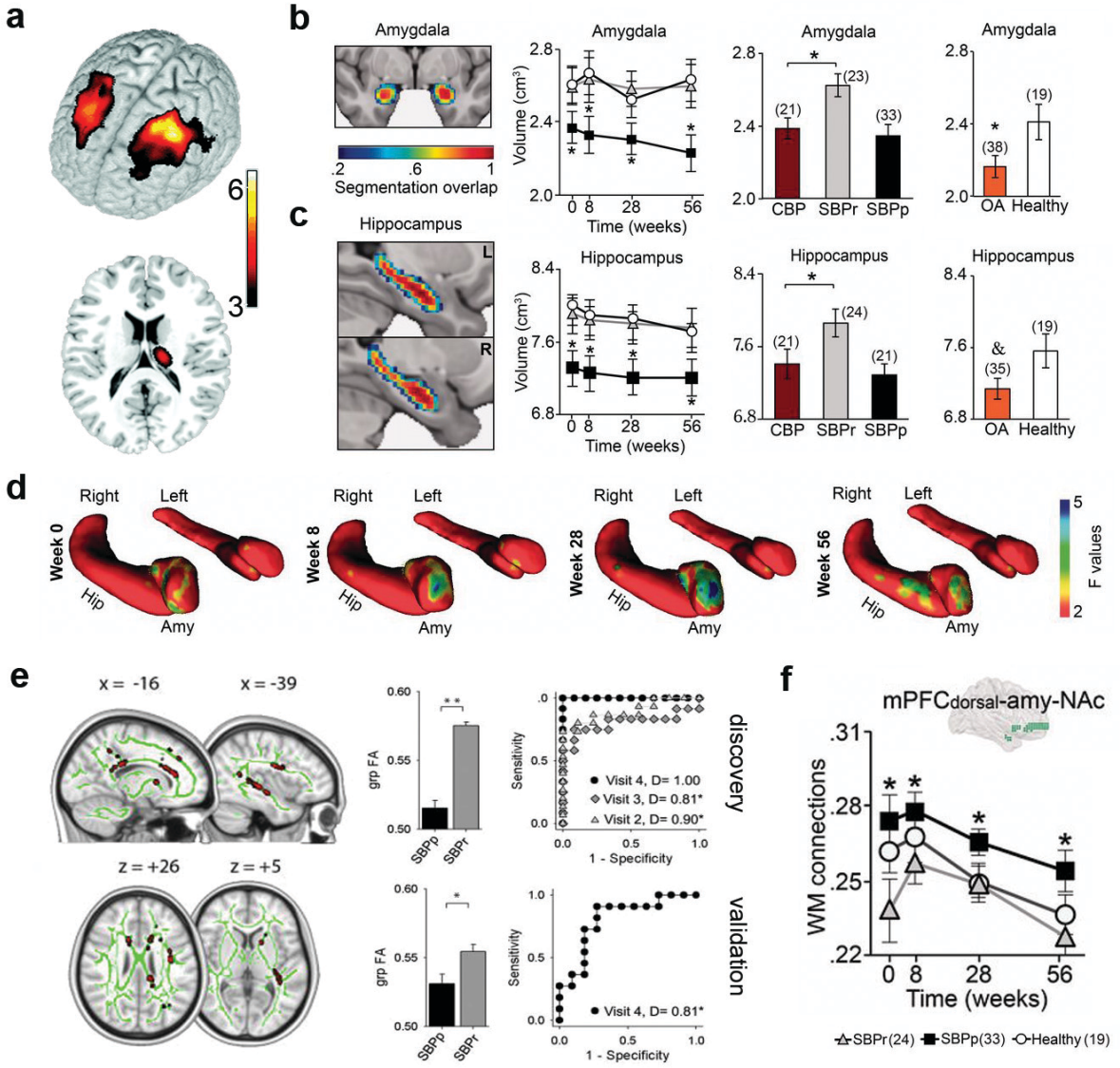


Figure 4: A summary of structural brain differences in chronic pain patients. **A**. Regional gray matter density (GMD) decreases in CBP subjects – top panel shows that GMD is reduced bilaterally in dLPFC; bottom panel shows that GMD is also reduced in the right anterior thalamus (bar graph shows t-values, where highly positive values indicate more reduction in gray matter) [49]. **B.** The amygdala is shown (heat maps display overlap of volume segmentation between SBPp, SBPr, and healthy controls (CON) at baseline. SBPp showed smaller amygdala volumes than SBPr and CON with no effect of time or groupXtime interactions. Similar findings were also seen in a different chronic pain cohort (osteoarthritis, OA, patients). **C.** The same results in the amygdala were also observed in the hippocampus for SBPp and OA. **D**. Vertex-wise shape displacement of the amygdala and hippocampus indicated that SBPp had thinner right amygdala across all visits. **E**. Whole-brain white matter functional anisotropy (FA, a measure of tract integrity) contrasted between groups (SBPr>SBPp) at baseline shows areas where tracts (green) are significantly lower in FA in SBPp patients (red). The average FA values across all white matter voxels differing between SBPr and SBPp (grp-FA) is shown in bar graphs; top panel = discovery group and bottom = validation group. In both groups, FA decreases in these regions significantly and accurately predicts persistence across time [52]. **F**. Percentage of white matter

connections from the same community reported in **Figure 3d** are shown in time over 3 years; SBPp showed more structural connections in this community than SBPr or CONs [45].

Importantly, many of the functional and anatomical alterations seen in CBP patients have also been demonstrated in other chronic pain conditions to various extents, including complex regional pain syndrome (CRPS) [56], knee OA [57, 58], fibromyalgia (FM) [59, 60], chronic vulvar pain [61], urologic chronic pelvic pain (UCPPS) [62, 63], irritable bowel syndrome [64], and headache [65]. These neurological changes, while somewhat overlapping between cohorts, have also been shown to be distinct between conditions [30, 66, 67] and partially reversible with pain relief [68, 69]. What these results indicate is that some of these neurobiological findings predetermine chronicity while others are the direct consequence of having long-term pain. Importantly, given that regions involved in CBP are similar to those in other pain conditions, treatments identified or developed to work on CBP might also show improvement in other chronic pain symptoms. Conversely, even if the anatomical regions involved are relatively consistent, the fact that different pain disorders have different and specific brain activity profiles might also mean that treatment response would not be homogenous.

2.1.2. Psychological and behavioral impact of CBP

Chronic back pain is also associated with myriad behaviors that point to abnormalities in the brain regions and networks specified above, particularly in the limbic system and especially in regards to the processing of emotionally-salient information. First, CBP is often co-morbid with various psychological disorders, including anxiety and depression [70-72]; although these illnesses are likely due, in part, to socioeconomic hardship and disability, they have also both been shown to affect the structure and function of limbic regions like the amygdala, hippocampus, and accumbens [73]. Second, CBP patients exhibit problems with learning and memory; these patients have shown problems in working memory and attentional tasks [74-76], as well as perspective memory tasks [77], many of which are dependent upon hippocampal and frontal mechanisms. Finally, chronic back pain also appears to impact the reward system. CBP patients have displayed atypical NAc activity during the encoding of acute aversive stimuli [78, 79], and studies have repeatedly shown that chronic pain patients display aberrant emotional decision-making [80, 81]. CBP patients consistently perform poorly across a variety of tasks that engage

reward-valuation and motivation circuitry, including choosing more often from disadvantageous decks in the Iowa Gambling Task (IGT, [82] **Figure 5a),** as well as displaying heightened gain sensitivity in a loss aversion task with accompanying alterations in functional modularity of the NAc ([83], **Figure 5b-f**). These differences may reflect cognitive deficits due to the underlying brain changes that take place in chronic pain and are associated with increased risk-taking and impulsivity within this patient cohort.

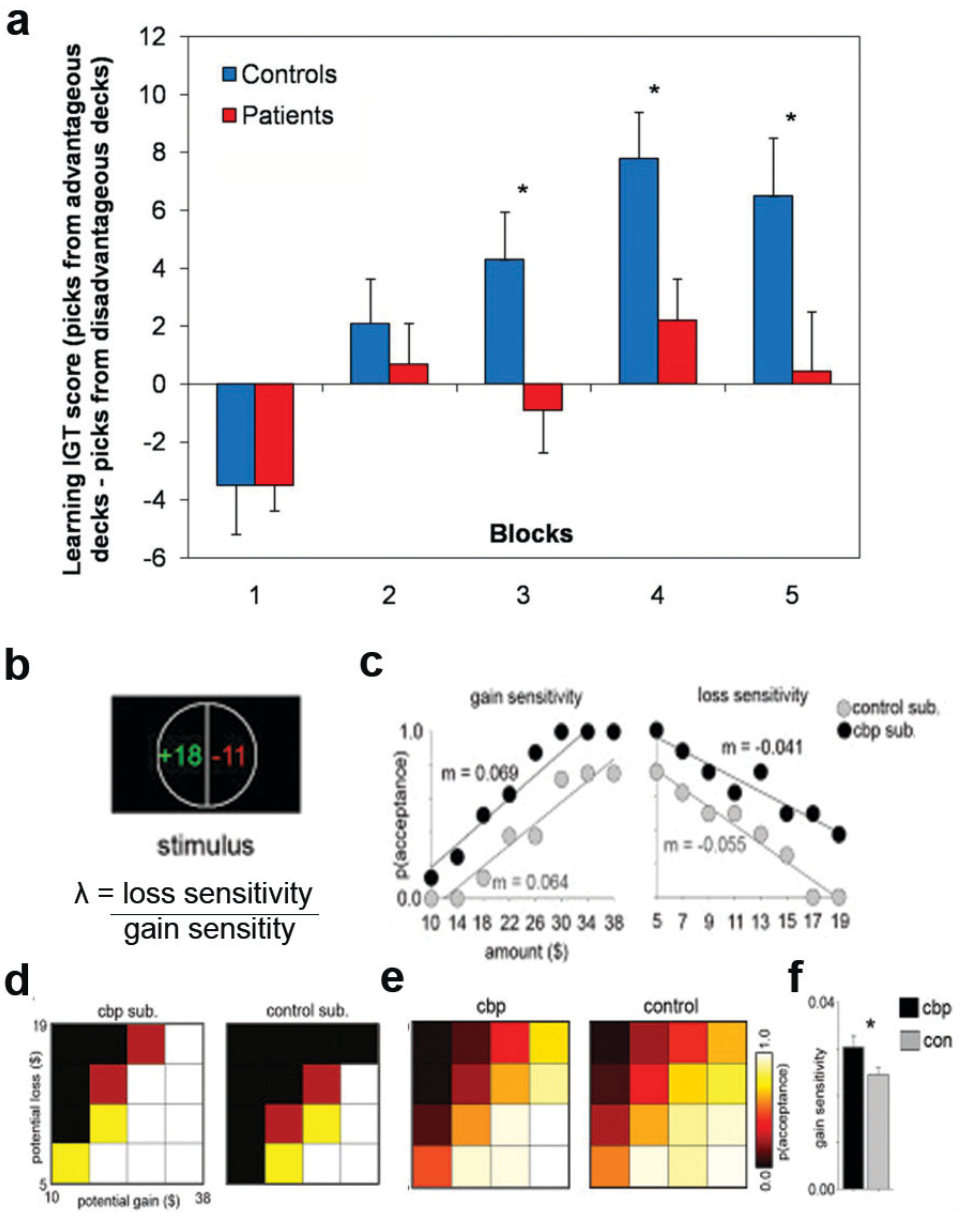


Figure 5: Examples of impaired emotional decision-making in chronic pain patients. The Iowa Gambling Task (IGT) measures decision-making strategies related to reward and punishment. Participants choose between 4 decks of cards; 2 sets of decks yield high immediate gain but larger future losses (disadvantageous) and the other 2 sets of decks yield lower immediate gain but smaller future losses (advantageous). Participants are instructed to maximize gains with minimal instructions about the underlying deck rules. Poor performance on this task indicates defective emotional decision-making and has been seen in patients with OFC damage. **A.** Healthy controls (CON) and patients with CBP and complex regional pain syndrome (CRPS) completed the IGT. Their performance (# choices of advantageous decks after subtracting # choices of disadvantageous decks) over time (blocks) is shown, with the performances of CBP and CRPS patients averaged into a patient group. While both groups started off with poor performance, CON eventually learned the implicit rules of the game and significantly improved in time; in contrast, chronic pain patients had difficulty learning the task and were significantly different from CONs in their performance over the last sets of trials [82]. **B.** Behavioral loss aversion describes the phenomenon by which losses have a larger hedonic impact than comparable gains; on average, healthy individuals are roughly 2 times as sensitive to losses as they are to gains (e.g., they would need a gain of at least $100 to make up for a potential loss of $50. Loss aversion is thus conceptualized as the ratio of losses/gains (lambda), such that a number >2 means more loss sensitive than expected and a number <2 means more gain sensitive than expected. We previously studied loss aversion in CBP patients using stimuli like what is shown in (B), where we asked patients to decide whether the possibility of winning the amount in green was worth possibly losing the amount in red (with gains ranging from $10-38 in increments of $4 and losses from $5-19 in increments of $2). **C**. Gain and loss sensitivity curves are shown for 2 subjects. Each point represents the probability of accepting an offer with each potential gain or loss. The number “m” is the slope of the fitted line, indicating their gain or loss sensitivity. **D.** After 64 trials of stimuli, decision matrices were generated by calculating the probability of accepting each individual $ offer out of the 8 presentations it was shown, creating an 8X8 matrix which was down-sampled to a 4X4 matrix for visualization. These matrices for the same two patients depicted in C are shown. Again, notice that the CBP patient was more likely to accept a monetary offer than the CON patient. **E**. Decision matrices averaged across each group are shown. F. Bar plots show each group’s mean gain sensitivity scores (denominator of lambda); CBP patients exhibited significantly higher gain sensitivity than CON [83].

2.1.3. Animal models of chronic pain

Finally, and importantly, many of these neurophysiological changes and anatomical predispositions have also been implicated and/or replicated in animal models of chronic pain post-neuropathic injury. These findings include: peripheral and spinal cord circuitry reorganization leading to peripheral and central sensitization as well as allodynia [84, 85], reorganization of resting state limbic networks [86], alterations in neuropeptide activity in the amygdala [87], changes in the functional connectivity, gene expression, and molecular excitability of the NAc [88, 89], changes in the excitation and activation of the ACC [90], abnormal hippocampal neurogenesis and short-term plasticity [91], and impaired hippocampal-based contextual fear extinction [50]. Animal models of chronic pain have also demonstrated behavioral aberrances comparable to human counterparts. In an operant task similar to the human IGT, rats with chronic inflammatory pain have been shown to prefer levers with larger but infrequent rewards (a riskier choice) compared to rats without pain [92], a response driven by decreased

dopaminergic levels and decreased neuronal firing of cells within the orbital frontal cortex (OFC) [93-95]. Additionally, animals with chronic pain also display impairments in both their short-term and recognition memories [89, 96]. These translational results indicate, in part, a conservation across species of maladaptive and dysfunctional neurological and behavioral responses in response to being in unavoidable, long-term suffering. Such overlap in findings between humans and animals provides evidence for a system whose mechanisms can be better mapped and in turn targeted by drugs or other therapies to alleviate chronic pain.

2.1.4. What remains unknown in chronic pain research

Although our understanding of the mechanisms and neural circuitry underlying the chronic pain state has substantially grown in the past decade, we have not yet been able to translate these efforts into meaningful treatments. Moreover, we still do not understand why some patients respond to certain kinds of medications while others do not, and likewise, we also have not fully captured the characteristics – neurophysiological and psychological, trait- and/or state-based – that make someone with chronic pain likely to respond to a treatment now or in the future. This information would be extremely valuable, not only for its obvious impact on personalized medicine initiatives, but also for its potential in making efficacy assessments in clinical trials of novel analgesics more accurate.

Section 3: The placebo effect

The placebo effect describes an improvement in symptoms caused by receiving a sham or inert treatment disguised to be indistinguishable from an active medical treatment. While we refer to the term placebo "response" to describe this phenomenon, it is important to note that patients are not responding to the placebo treatment itself, but rather to the language, caring, culture, history, and overall context surrounding the treatment administration [97]. While the use of placebos in today's medical industry is well known and viewed as a relatively mundane component of most randomized control trials (RCTs), the emergence of the placebo's clinical utility has a long and interesting history.

3.1. History of placebos: from funeral processions to dissertations

The word placebo is Latin in origin, meaning “I shall please”. The term was first used in the 14th century in reference to hired mourners at funerals who stood in for family members of the deceased; these professionals would cry and wail during the procession and burial, often beginning their chants with “Placebo Domino in regione vivorum”, the 9th verse of Psalm CXIV, which translates to “I shall please the Lord in the land of the living” [98]. Placebo wasn’t used as a medical term until the late 18th century, when the New Medical Dictionary described it as a “commonplace method or medicine” and a few years later, the Quincy’s Lexicon-Medicum defined it as “an epithet given to any medicine adapted more to please than to benefit the patient” [98]. For the next 100 years, the use of placebos in every-day medical practice was prevalent, and many physicians were taught to utilize everything from colored water and residual ash powder to bread and sugar pills, viewing these practices as endorsed and necessary forms of deception [98]. In the early 1900s, the first placebo-controlled medical trials took place, and in 1938, the word placebo was first applied in reference to being a control for an active treatment in a clinical trial investigating efficacy of cold vaccinations [98]. Interestingly, the trial ended up having negative results due to substantial improvement in the placebo control group, and the placebo effect as we now know it was born.

Although the power of placebos has been proven time and time again, the use of sham procedures and medicines has been shrouded in controversy, predominantly in regards to the physiological and psychological authenticity of a placebo response and the ethical implications of placebo administration. Even during their peak usage in the 1900s as either primary treatments or supplements to medicines, placebos were viewed as having no impact on physiology and were only thought to comfort or appease patients, in particular those deemed as having lower intelligence or higher emotional lability. These two sentiments remain in present day ideas of placebo effects, albeit more subtly. For example, some clinicians still view the presence of the placebo effect to be evidence that certain symptoms are not caused by “real” or “organic” diseases [99], with patients who report alleviation of symptoms often still judged to be faking their responses or illness (i.e., “malingerers”) [100]. This is unfortunately quite prevalent in studies of pain, not only because pain cannot be easily measured and remains a highly subjective, personal experience but also because chronic pain is often idiopathic, with no history of

disease or known cause of injury. However, inferences like these are only valid if placebo treatments have no actual effects on pain pathophysiology or experience, which until the early 1950s was still unknown.

Compounded with concerns of patient trust were those surrounding patient care and physician trust, issues that were deeply rooted in the history of placebos. After World War II, two key factors changed how physicians, researchers, and society understood the placebo effect. The first was the introduction of RCTs as the gold standard for clinical research, and along with this systemic change in medical testing came the association of the word placebo with sham interventions. Because of this association, people began viewing placebos as nothing more than controls for confounding factors in active treatment responses (in turn diminishing what little interest there was in explaining the biology of the placebo effect). Second, the rise in autonomy-based theories of medical ethics put pressure on the medical field to verify informed consent and legitimize a person's right to know and make decisions about their treatment regimens [101]. Beginning in the 1960s, clinical ethicists began to explicitly target the use of placebos "for the good of the patient", framing their administration as examples of unnecessary, unjustifiable, and covert medical paternalism. Thus, by the end of the 1970s, the word placebo was now synonymous with fake, deceptive, or fraudulent interventions, viewed by physicians and lay persons alike as unethical outside the boundaries of an RCT [101]. Remnants of this view of placebos are still prevalent today, with many potential participants refusing to partake in clinical trials because of the chance they might receive a placebo, and many physicians still viewing the clinical use of placebos as violations to their ethical code of conduct. Thus up until recently, the physiological effects associated with the placebo response were first believed not to exist, and studying this phenomenon later came to be seen as unethical and of no benefit to the patient.

Fortunately, due to technological advances in neuroscience and converging results across a series of laboratory-based and clinical trials, we are slowly pushing past the methodological and philosophical obstacles that have hindered a real appreciation and utilization of the placebo effect. We now understand that the placebo effect is a powerful psychobiological occurrence rooted in underlying and identifiable neurobiology while simultaneously being effected by psychosocial contexts [102]. Moreover, we also now know that the placebo response is a phenomenon that has been observed across

a variety of diseases, biological systems, and treatments [103] and that it is psychologically durable, with responses lasting well beyond what would be expected of feigned relief (in some cases up to or exceeding 12 weeks post treatment) [104].

3.1.1. The placebo effect in pain – placebo analgesia

Importantly, placebo effect has shown some of the largest and most consistent effects in the field of pain, which isn't surprising since pain perception is subject to influences by cognition and emotion [105]. Analgesia in the placebo arm of clinical trials is observed often, especially for chronic pain drug studies [106]. In back pain specifically, one study found that repeated sham (placebo) treatment resulted in greater pain relief than that caused by the tested conventional therapy [107]. It has been shown that placebo effects in both acute and chronic pain tend to be larger when participants present with high levels of pain [108]. Thus, the placebo's potential for substantial analgesia in clinical settings cannot be understated.

A body of literature has already established that the neurophysiology of placebo analgesia is complex and depends on multiple neural circuits. Placebo analgesia is known to recruit endogenous pain pathways acting upon the opioid system to regulate descending inhibition via the PAG [109, 110], a mechanism that can be reversed by the administration of the opioid antagonist, Naloxone [11, 111]. Additionally, in some cases, individuals who report having the highest levels of analgesia resulting from active drug effects also show the largest placebo effects, indicating that some analgesic medications and placebos may share common mechanisms or pathways [99]. Besides these anti-nociceptive circuits, which reduce spinal responses to pain [112], the placebo effect is also dependent upon limbic circuitry and higher-order frontal mechanisms involved with context generation and de-coding, expectations of treatment outcome, emotional appraisal of events, and reward learning [99, 113-115].

3.1.2 The psychology of the placebo effect

In addition to biological mechanisms of the placebo effect, psychosocial mediators of placebo response have also been heavily investigated, albeit to little or no avail. Since the 1950's, individual personality differences have also been hypothesized to contribute to placebo response magnitude or the

presence or absence of a response. Historically, most of the research aiming to capture a "placebo responder personality" has used psychological profiling [116], with a prominent focus on affective traits (positive: extraversion, optimism, and openness to experience [117-120]; negative: neuroticism and pessimism [121, 122]), as well as research on the role of social learning and empathy [115, 123], reward and motivation [124], suggestibility [125], and expectations toward treatment [120, 126-129]. For example, individuals who score high in optimism, openness, and suggestibility, low in neuroticism and pessimism, and have positive expectations toward a treatment are often considered to be good potential placebo responders. However, there is still no conclusive evidence of a set of patient-reported outcomes (PROs) or self-report measures that predict placebo response, are consistent between studies, and correlate well with brain pathways related to response [130]. Additionally, many of the personality traits that are associated with large placebo responses in healthy controls have been shown to be easily influenced by context; in particular, both dispositional optimism and extraversion are only associated with larger placebo responses in situations that include warm emphatic interactions with caregivers and researchers [97, 131, 132].

3.2. Roadblocks to studying placebo response in chronic pain

Given the power of the placebo effect and the lack of efficacious medicines for chronic pain patients, it makes intuitive sense to study the placebo response within the context of chronic pain treatment. However, the relative contribution of each of the cognitive, emotional, neurological, and anti-nociceptive systems (outlined above) to the clinical placebo response in chronic pain patients remains unknown. So, why hasn't there been a substantial effort in investigating the clinical efficacy of placebo analgesia? And of the research that exists, why are there so many inconsistent findings between studies investigating placebo analgesia? Below I outline 3 reasons ("problems") that I feel contribute to the gaps in our knowledge regarding the intersections of chronic pain relief and the placebo effect, although by no means is this list exhaustive. These issues must be addressed if we are to seriously move forward in either predicting clinical placebo utility and/or manipulating placebo response in chronic pain patients.

3.2.1. Problem 1: Inconsistent framing

One of the reasons for to inconsistent results between studies may be that the field is still divided about the fundamental "nature" of a response. It is still actively disputed whether the placebo response represents a relatively stable *trait* (e.g., a person's neuroanatomy, in combination with their genetics and long-standing habits and behaviors, predetermines his/her ability to respond positively to an inert treatment), or instead, if it is a dynamic *state* subject to interference from a variety of psychological, biological, and environmental factors (e.g., a person who responds to a placebo treatment in one form is no more likely than the next person to respond to a placebo treatment in an alternative form due to potential differences in context and mood). This philosophical debate is important because it suggests that if the placebo response is a state, it may be able to be predicted but only in certain and very specific instances; it also indicates that the placebo effect would be more easily influenced or manipulated. In contrast, if the placebo response is a trait or a component of a trait, it should be able to be predicted across time and reduced to a unique model. This discussion is also significant because of its potential impact on study designs - framing the placebo response according to only one of these schools of thought could lead to divergences in interpretations of any psychological differences found, or, more likely, to differences in the questionnaires chosen to measure the phenomenon in the first place, resulting in a systematic bias in many placebo studies.

The literature provides evidence for both state and trait possibilities. Supporting the state framework are research studies showing that the placebo effect can be affected by all sorts of external factors, including but not limited to: frequency of treatment, dose, appearance and branding, route of administration, treatment ritual, pre-treatment conditioning, interpersonal interactions, cultural constraints, feelings of empowerment, previous exposure to therapeutic relationships, and timing of expectation [104, 133-136]. For example, studies have demonstrated that taking two placebo pills twice a day was not only more effective than no treatment but also more effective than taking one placebo pill twice a day, implicating a perceived "dose-dependent" response [101]. As another example, the color of a treatment has been shown to affect the placebo effect, with a red pill being more likely to cause a response than a blue pill [101], and there is recent evidence that the extent of response is correlated with the length of a clinical trial (such that the longer and more expensive a trial is, the more placebo response seen [137]). A

nice summary of the potential ways that different contexts might impact placebo response can be found below in **Figure 6**.

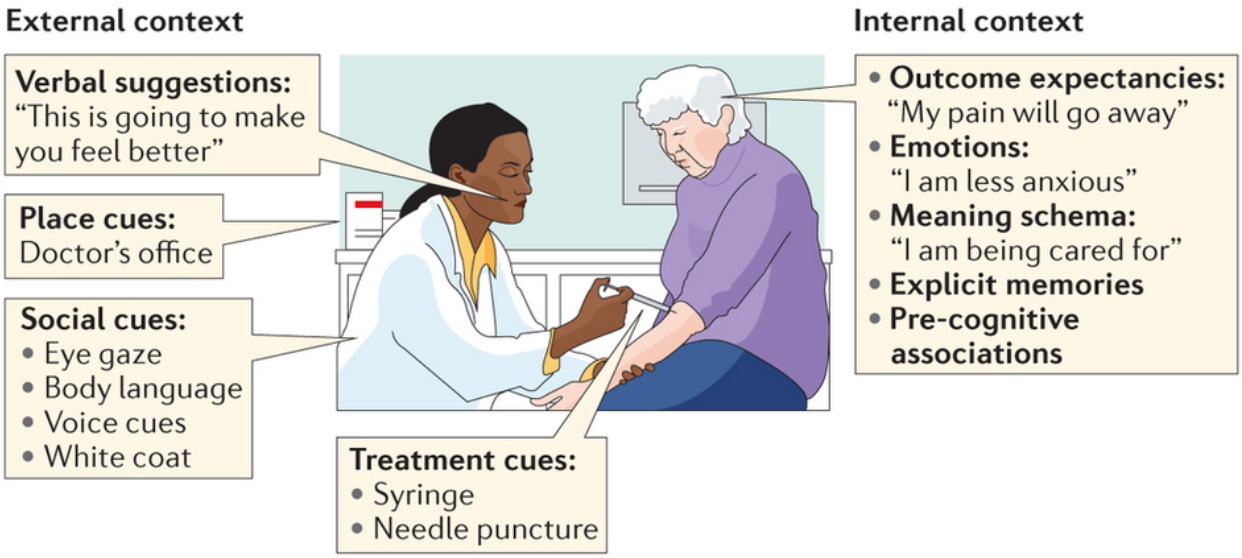


Figure 6: Context can affect placebo response. This figure was taken from Wager and Atlas's 2015 review regarding the neuroscience of placebo effects [99]. It gives examples of how various internal and external stimuli can influence the psychology of the placebo effect. The authors state, "Clinical settings that surround treatment include multiple types of context information that are perceived and interpreted by the patient's brain. The external context includes treatment, place and social cues, along with verbal suggestions. The internal context consists of memories, emotions, expectancies, and appraisals of the meaning of the context for future survival and well-being. These features combine to make up the treatment context and are the 'active ingredients' of placebo effects."

Placebo response can also change depending on the ailment – one person might consistently respond to placebo in the context of acute analgesia but not in the context of reduced anxiety or improved mood, for instance [101, 134, 135]. Even within the same condition, there are within-subject inconsistencies. In a study measuring placebo response to three different analgesic contexts, there were no significant associations between placebo pills, sham acupuncture treatment, or cue conditioning effects, indicating that individuals may respond to unique healing rituals in different ways, properties that are more indicative of a state [114, 138]. Likewise, the pharmacological evidence suggests that there exist many kinds of placebo responses as opposed to one single placebo effect [139, 140] – responses may recruit different mechanisms and circuitries across the body, including the opioid and cannabinoid systems, dopaminergic pathways, the seritonergic system, and the endocrine system [101]. What all of

these findings point out is that there is inherent variability and inconsistency within the placebo effect [141] between and within subjects, moving away from theories suggesting that the placebo response comes from personality traits, specific biological pathways, or particular brain region activation.

However, there is also evidence that the placebo response (although variable) can still be conceptualized as personal, consistent, and trait-like, and this involves considering an individual's genetic predisposition, neurobiology, and other environmental factors [133, 134]. For instance, studies using repeated placebo administration have shown that individuals responding to placebo in one set of environmental cues had a high likelihood of responding again in the exact same setting, arguing for a stable and reliable placebo response [116]. Furthermore, other researchers have shown that the placebo effect can still persist even when participants know they are receiving inactive treatment, illustrating how reinforcing treatment cues with positive outcomes can create scenarios where analgesia is independent of state-based expectations of pain relief [142]. The neurobiological evidence of trait-like placebo response is also convincing. Structurally speaking, studies have found that white matter tract integrity in descending pain pathways are correlated with inter-individual differences in placebo analgesia [143]. Functionally speaking, a recent neuroimaging study investigating depression medication found that placebo responders had distinct neural patterns that changed from baseline measurements over 6 weeks of chronic treatment, indicating a process of neural adaptation in specific brain regions overtime. Importantly, these patterns were nearly the exact inverse of those of non-responders, whose neural signature during the 1st week of treatment remained stable throughout the study [102, 144]. The finding that these functional connections, which are prone to fluctuations and perturbations, were distinct between groups at baseline suggests some level of trait-like predisposition to respond.

Regarding pain specifically, despite evidence that the placebo effect can be malleable and dependent upon context in pain paradigms [145], a few neuroimaging studies provide evidence for predictability of the placebo analgesia based on brain functional properties. For example, the functional connectivity between mPFC and bilateral insula strongly and reliably predicts response in CBP patients given a placebo patch instead of a lidocaine patch [146, 147], and the efficiency of resting state network properties correlates to the level of placebo analgesia reported in OA patients [148]. Additionally, midfrontal gyrus resting state connectivity has been shown to differentiate placebo responders from non-

responders with OA pain, as well as correlate with standardized pain outcome measures [149]. Importantly, these brain regions of interest remain relatively consistent across different types of placebo rituals examined in either CBP or chronic osteoarthritis (OA) knee pain [147-149]. These findings provide us with a solid base from which to identify brain characteristics of the placebo response in chronic pain patients (those that are stable and those which may fluctuate) and identify reliable biomarkers from these and other measurements. However, none of these studies specifically looked at psychological or personality-based characteristics that might contribute to or interact with resting state properties to give rise to placebo analgesia. Moreover, none of them had a proper control arm (discussed below).

Additional research looking at the genetics underlying treatment response provide further evidence of predetermined, trait-like qualities in the placebo effect. Both serotonergic-related and catechol-O-methyltransferase (COMT) gene polymorphisms have been consistently linked to individual placebo response in social anxiety and major depressive disorder [150]. Such pre-existing genetic biomarkers of response, which are relatively invariant to time, also provide strong evidence for the placebo effect as a predictable, stable phenomenon.

Returning briefly to the question of why a state versus trait framework might influence the robustness or reproducibility of placebo findings, one could imagine that if a researcher viewed the placebo effect as a state, s/he may design a study void of measuring personality characteristics or stable neurological properties, potentially missing key elements of a response or response propensity. In contrast, if a researcher thought that the placebo effect was a trait, s/he might only include self-report measurements at one or two time points pre-treatment, failing to account for dynamic changes in mood or expectations throughout the study. Both viewpoints are problematic because they disregard the notion that a physical, mental, or emotional state for one person may be more trait-like for another person, and vice-versa. Importantly, both perspectives also ignore the potential complexity of the placebo effect. Instead, a placebo response could reflect a *combination of* or *interaction between* state and trait characteristics, such that certain personality dimensions, brain properties, genetics, and learned behaviors (i.e., traits) make it more likely for a person to exhibit the placebo effect but, due to previous personal experiences, expectations, and/or the current context (i.e., states), does not guarantee that this effect will always occur. In reality, there may be a system or set of systems involved in a placebo

response (trait) that may respond differently depending on the context that the placebo is given in or the current mood of the participant (state). One could imagine that the placebo response reflects not only the specific hardwiring of a distinct universal neurological system and its corresponding genetic and psychological components (traits) but also different thresholds and patterns of activity (internal states) within this system that are activated only in given contexts (external states), which may be unique to a person and possibly even relatively stable for that person (traits). Viewing a response in this way allows for more nuanced inquiries investigating the intersections of biology, personality, context, and culture. Given that long-term pain is a complex phenomenon consisting of both traits (genetics, biology, and personality), as well as states (intensity of pain, current activities, current medications, and current psychological well-being), reframing placebo analgesia to potentially be a combination of both state and trait properties fits better with the nature of chronic pain. There are very few, if any, current projects (outside of the one presented in this dissertation) with research designs that allow for the study of both predetermined and context-specific attributes of the placebo effect, and more of these are needed to further scientific progress in this area.

3.2.2: Problem 2: Overuse of healthy participants

The second reason why little is known about placebo effects in chronic pain is largely because the neurobiological and psychological mechanisms underlying placebo analgesia have been almost exclusively studied in healthy pain-free participants who are given an acute painful stimulus (such as an injection or series of stimulations) under paradigms involving conditioning or manipulation of expectations [151]. In these subjects, as mentioned above, placebo analgesia appears to reflect altered transmission in pain pathways, including: (1) reduced activity in many brain regions involved in acute pain perception and emotional processing, (2) activation of areas important for modulation of pain-related regions, including the engagement of descending pain circuitry, and (3) activation of endogenous opioid and dopamine systems [102, 105]. Moreover, the placebo effect in healthy people seems to be primarily a state rather than a trait characteristic (or at least framed as such), with placebo responses showing a lack of consistency across different routes of administration [138, 152]; such results have been used to explain the difficulty of identifying placebo responders in the general population [130].

This reliance on healthy individuals for studying the placebo response is problematic for three key reasons. First, while it is possible and highly likely that placebo analgesia shares common physiological pathways and neurological mechanisms among all individuals regardless of health status, brain anatomy and physiology is distinct in chronic pain patients, both due to predisposing factors and plastic reorganization with pain chronification [42, 48, 52, 153] and also perhaps as an aftermath of having had long and repeated exposures to myriad medical rituals which may render patients more consistently or predictably placebo responders/non-responders. Importantly, the regions implicated in the predisposition or continued presence of long-term pain are also associated with placebo response. There is the frontal-parietal system that is important in decision making (being in a study and taking the medication consistently), emotional control and reappraisal (being aware of what you are feeling while you take the medication), and maintaining cognitive contexts (knowing what possibilities of treatment you are getting now and who is doing the research). There is also a more distributed system that is involved in the reward of pain relief, evaluation of tasks and stimuli, processing of ascending and descending pain signals, and forming memories and learned expectations based off of all of these things; this system includes the mPFC, OFC, NAc, amygdala, hippocampus, thalamus, and periaqueductal grey (PAG). Both systems are important and integral to pain perception and placebo response – together, they connect a diverse set of brain regions responsible for sensation, perception, and emotion, and by virtue of their specific functions and distributed connections with other regions, are likely responsible for linking situational context and sensation to meaning, evaluation, and behavior [101]. Thus, it would make sense that some of the same regions that serve as biomarkers involved in pain chronicity might also be involved, albeit differently, in identifying placebo response. However, because of this overlap, those networks associated with placebo response in healthy pain-free participants may already be perturbed by the presence of long-term pain or by having to cope with this pain for many years prior to placebo treatment, resulting in compensatory mechanisms that may impact the characteristics or magnitude of placebo responses in patient populations. This has largely been ignored in the placebo literature.

Likewise, the majority of psychological variables associated with response (reported in **3.1.2**) were studied in healthy populations and, here again, their overlap with clinical cohorts is unknown as certain personality traits may be more or less prevalent in chronic pain populations. Since chronic pain

patients may present with co-morbid depression, anxiety, and/or risk-taking behaviors, one cannot assume that the same personality traits found in healthy placebo responders will also be attributable to pain patients. Nor can we presume that a person's current disposition occurs in a vacuum outside of their long-term health conditions (that is, perhaps a patient's likelihood to respond to a placebo changes depending upon the kind of chronic condition that they have and how it impacts their quality of life). The only "trait" consistently found thus far among placebo responders in chronic pain is a high baseline pain severity [116]; to date, sensory and affective pain characteristics, total duration of pain, and demographics (including gender and race) have not been found to affect placebo response in clinical cohorts. The personality traits and psychological characteristics of clinical placebo analgesia in the context of chronic pain have remained relatively unexplored.

The third reason why relying on healthy controls to study placebo is problematic is because the settings of a laboratory often differ vastly from those in a clinical trial, which may impact the placebo responses in various ways. Clinical trials usually involve repeated visits, longer follow-up time, and different researcher-to-participant interactions (e.g., blood work, physical exams, probing for adverse events, and/or exit interviews), all of which might influence the context under which a patient might respond to placebo (e.g., being in a medical setting, receiving increased attention, talking often about their symptoms, etc). Moreover, individuals with chronic pain are already fundamentally dissimilar to their healthy counterparts when it comes to expectations of treatment efficacy and/or already learned responses to treatments or clinicians. A lifetime of interacting with various medications, procedures, physicians, and the medical system changes the baseline state at which many chronic pain patients start a clinical trial, a confound that cannot and should not be ignored when investigating the placebo effect. And unlike laboratory investigations, clinical trials do not usually manipulate or condition response to placebo – the participant comes in with his/her own expectations of randomization and potential treatment efficacy. Therefore, the findings regarding the placebo effect in healthy participants may not easily translate to chronic pain populations.

3.2.3. Problem 3: Lack of a proper control group

The final reason why there is relatively little known about placebo response in chronic pain is that the majority of studies do not include an additional control group beyond those participants receiving placebo, a flaw present in most RCT designs. In a sort of research paradox, the assumptions and aims that underlie most RCTS make it difficult to actually investigate placebo responses and the factors contributing to them [154]. This is because the goal of an RCT is not to understand the placebo response but rather attempt to *control* for it, and because of this, clinical trials essentially work to define what a placebo is without going any further - that is, they define a placebo as a control for active effects, but ironically do not investigate the effects of the placebo treatment itself. Due to the assumption that a placebo group is a proper and "good enough" control for an active treatment group, the large majority of studies do not include a proper no-treatment arm. This is problematic because a no-treatment arm takes into account the natural history and progression of the disease, regression to the mean, natural fluctuations in symptoms, physical co-interventions, and purely psychological effects [101]. Without a control like this, one cannot definitively conclude that changes seen in the placebo arm (or the active treatment arm, for that matter) are not due to random fluctuation about the mean. Additionally, without a no-treatment arm, researchers cannot measure the magnitude of the placebo response specifically.

3.3 Significance of studying the placebo in the context of chronic pain

In addition to its potential to clarify mechanisms of resilience, adaptation, and analgesia at a basic research level, studying the placebo effect within the specific context of an RCT for chronic pain could significantly impact the way medicine and research is practiced. The placebo effect will remain a nuisance to the medical industry as long as its properties are not understood to the extent with which it can be predicted, manipulated, and in turn administered to the proper subjects under optimized conditions [155]. Therefore, gaining a better understanding of the neurobiology of clinical placebo and identifying individual characteristics of placebo propensity is necessary for harnessing placebo as an aid in diminishing chronic pain disabilities. Such an effort would also lead to increased efficiency and accuracy in performing clinical trials, especially if placebo responders could be removed from active treatment groups prior to commencement or if the placebo effect in RCTs could be manipulated to have a null effect size. The importance of this latter effort cannot be understated. A large majority of clinical trials of novel

medications designed to treat chronic conditions fail. In recent years, one of the most interesting findings is that the placebo response in many of these trials is increasing while the drug responses remain stable, resulting in a smaller separation between the drug and placebo arms. This trend has been seen in both analgesics [156] and anti-depressants [157]. If we can understand why this is happening and prevent it, manipulate it, or exploit it, we might not only be able to improve medicinal standards for RCTS but also revolutionize the way clinical trials are performed.

Section 4: Language as a quantitative tool

There is evidence that semantic properties of chronic pain patients' speech may provide important and currently missing information about the nature of placebo analgesia. Language is often seen as a "window into the mind", since it is simultaneously a mediator and shaper of mental concepts [158]. Because of its intimate role in the subjective construction and explanation of phenomena like thought, emotions, and experience (as well as being one of the natural consequences of these things), an interest in using language to investigate human behavior has gained momentum. In recent years, researchers have used various language properties to capture many elements of the human experience using properties mined from interviews, blog entries, and social media posts/tweets. From these investigations, we now know that we can use language to measure personality and mood characteristics, relationship stability, and political associations [159-163], calculate approximate age of speakers [164], predict psychological disorders such as schizophrenia [165], identify classes of drugs ingested [166], and quantify the extent of various neurological ailments (such as aphasias [167]). The results of these research initiatives have been promising and their findings provide support for utilizing language to not only better understand complex conditions and pathologies, but also develop tools to predict future behaviors and changes in health status. Language in particular is an ideal candidate for these kinds of analyses due to its unique position as both the cause and consequence of thought, and in turn it has the ability to pull many biological and psychological elements together in one measurement. Given all of these findings, language analyses have the potential to elucidate neurological and psychosocial mechanisms regarding pain, including recovery from subacute episodes, coping with chronicity, or in the case of this thesis, responding to a placebo treatment.

Interest in the language used by individuals in pain is not new. Numerous ethnographies of chronic pain patients have emerged over the last 50 years [168-171], and more recently, papers regarding pain physician's thoughts and narratives about their patients have also been published [172, 173]. However, all of these inquiries have been limited to the social sciences, and in turn, the large majority of the analyses have been qualitative, using coding and grounded theory to find patterns and themes in people's stories. While these methods are important, they do not translate well to clinical measurements or to reportable, quantifiable outcomes. Those efforts that have been made to quantitatively study the language of pain have only been done relatively recently, and so far have only been attempted either in too broad of a context, like the entire English [174] or Greek [175] lexicons, or in healthy participants talking about acute pain [176]. The latter is particularly problematic, since the experience of pain in individuals with chronic conditions is vastly different than people experiencing acute pain or remembering previous pain episodes, and therefore it does not follow that the language of pain would be the same between chronic pain sufferers and healthy individuals. This is primarily because of the many neurophysiological and psychological differences between healthy and chronic pain patients due to both changes caused by prolonged suffering and predetermining factors that contribute to the chronification process; such differences would be expected to result in differences in language used. Additionally, chronic pain patients bring with them myriad medical, personal, and social encounters and expectations that many people without these conditions do not have; such previous experiences will inherently change how these individuals conceptualize and talk about their pain. Put simply, the starting point or baseline of pain patients and healthy individuals with acute pain is not the same and therefore neither is their linguistic repertoire.

There are currently no efforts being made to study the language of chronic pain patients in general or in the context of placebo, even if the results of such investigations could be incredibly valuable. To our knowledge, we are the first to systematically look at language in chronic pain specifically, with the goal of searching for quantifiable semantic differences in patients' narratives related to their analgesic response to a placebo pill in a clinical trial.

4.1 Overview of qualitative interviewing

This thesis does not use qualitative methods to analyze language data, and therefore these methods are not discussed here (although they are mentioned in the **General Discussion** as we might use them in future analyses). However, the interview utilized in this dissertation was designed with a qualitative approach in mind, which is briefly explained below.

Qualitative interviewing aims to understand how individuals make meaning of their lives, how they organize their social relationships and identity, and how they subjectively perceive information [177]. Such a technique can offer insight into larger and broader topics such as culture, politics, and other social forces, or it can provide information about one person or a group of people's unique experiences and rituals in the world. Although interviewing techniques have rarely been brought together with neuroscientific approaches, the combination may prove to be quite powerful, especially in the context of placebo and pain. Through a combination of neuroimaging and language analyses, we may be able to afford information about individuals responses, thoughts, experiences, and natural behavior that can then be correlated with the same individual's neurobiological functioning, providing "a link between the acculturated mind and brain" [177]. Applying this to CBP and placebo specifically, we might now be able to ask and answer questions such as:

> (1) How does the chronic pain state influence one's thoughts and in turn the words that are used to store, express, and change these thoughts (about one's pain, about one's current pain treatment, and about one's life in general)?
> (2) How does someone's placebo propensity also affect his/her thoughts and word choices? Do these effects overlap with those caused by chronic pain?
> (3) Can these influences be captured in an interview, quantitatively measured, and used to predict various behaviors?
> (4) Can we link complex experiences and memories from people's verbal disclosures about pain and placebo response directly to the complex mechanisms and circuitry of each individual's brain and the state(s) it is in during the study?

Interdisciplinary explorations and questions such as these are rare and they are incredibly difficult to do well. However, very recent work combining anthropological methods (like interviewing) and neuroscience methods is beginning to show that this type of methodology can be fruitful. As an example, researchers recently investigated cultural differences in the ability to emotionally empathize with people in physical and mental distress; they found that differences in an individual's use of affective language during a one-

on-one interview predicted individual differences in activation of particular cortices in response to these distress situations [178]. Moreover, the researchers saw this predictive capacity even if these individuals did not self-report feeling more strongly or empathetically than those participants who used more cognitive-centered language, and the findings were also dependent upon culture (i.e., Americans showed this relationship but Chinese participants did not) [178]. This suggests that language can pick up more-subtle relationships between emotions, personality, decision-making, culture, and neurophysiology than other approaches or tasks, even in the absence of participant awareness (possibly resulting in a less biased measurement of the subject of interest, such as placebo response).

The interview created in this thesis was semi-structured as opposed to structured. The idea behind this kind of guidance format is to foster more of a conversation with participants about their lives and experiences and make them feel more comfortable, all the while following a set of questions that will be common between people and thus relatively consistent between subjects. Here, the interviewer directs the participants' answers from one broad topic to the next, but does so in a way that remains in the context of the ongoing conversation, meaning that questions may be differently ordered between participants [141]. Additionally, the interview was made to be open-ended, another qualitative format that gives participants "the freedom to discuss in greater detail or at greater length aspects of their experiences they feel are the most important or noteworthy" [141].

Finally, the interview attempted to maximize the likelihood that we would identify words or themes uniquely linked to the placebo response and its neural signature in CBP patients. This was done by including questions that captured the distinctive characteristics of chronic back pain spanning the disease continuum, as well as the physical and emotional toll long-term suffering takes. It also probed for previous experience with treatments and the medical system, expectations in the study, and assessment of placebo efficacy. Importantly, some questions also significantly overlapped with those asked in the battery of questionnaires dispensed to the participants – this was done so that we would not only possibly be able to tie language with a placebo response and with neuroimaging, but also we could link language directly with questionnaires measuring personality traits and psychological states. Such an approach is novel and also critically important for development of a tool (interview or otherwise) for placebo propensity. If such a tool was inspired by and created from a combination of brain mechanisms

(neuroimaging), personality (questionnaires) and lived experience (language), and if these elements can be shown to correlate to one another, then any predictive measure given by the tool for placebo propensity will not only be psychologically valid but also physiologically valid.

4.2 Overview of different quantitative methods

The last 60 years has also brought about a number of different methods to quantitatively study language, an effort largely encompassed by the term “Natural Language Processing” or (NLP), which describes interactions between computers and human languages. Broadly speaking, there are five main ways researchers have aimed to quantify language, although these are by no means exhaustive or totally independent from one another: (1) word counts and general tokenization, (2) syntax tagging, (3) semantic tagging, (4) latent semantic analysis, and (5) a combination of two or more of the above. Each one is briefly described here, including their utilization in research and known pitfalls.

4.2.1. Counting words

First and most basic are word counts, which are used to capture at the very least three key measurements: verbosity (how much does someone speak/how many words do they say), vocabulary (how many unique words does someone use in a given segment of speech), and lexical diversity (how many of the total words spoken were unique words – of everything they said, what words had unique or specific forms?). The latter term has been referred to as “a poor man’s version of lexical entropy”, and it’s initially calculated as a ratio of vocabulary/verbosity (also called the types-to-tokens ratio, TTR). As a ratio, its max is 1. A TTR = 1 (1/1) is only possible if a person never repeats a term, as in counting numbers, and a TTR = 1/tokens is only possible if a person says the same thing over and over again (during a tic or stuttering episode for instance). Obviously, most people fall somewhere in between. However, there is an inherent problem with TTR calculations in many languages, especially English, due to the repetitive nature of language in general. TTR has an inevitable downward trajectory because English requires the speaker to use a lot of what’s called “functional vocabulary” like articles and prepositions. Because of this, the verbosity calculation is often biased in the ratio and thus as TTR falls (denominator - verbosity – increases) because it can’t distinguish between speech of different lengths.

For example, a 100 word children's story might have the same TTR as a Shakespeare play because the constant repetition of functional words overwhelms the greater diversity in absolute terms of the latter (making TTR fail with changes in sample size) [179]. Therefore, it is necessary to correct for the length of the text through some factor in the denominator (called a corrected type-to-token ratio, CTTR). There are many different types of corrections, but the easiest and most common is taking the square root of the verbosity measure, or some sort of variation of this. Word counts are prevalent in linguistic research because any basic NLP software can easily and quickly calculate these measures and sometimes they are very useful. For example, they have been used to identify recovery from a stroke or accident (such as a traumatic brain injury) that maybe have inhibited or damaged part of their language network [180]. Word counts have also been used to measure how proficient people are when learning a new language [161, 181]. For the purposes of this dissertation, word count analysis was used as a quality check to make sure that our patients were displaying normal language properties and thus none of the other language results could be explained by some basic disordered speech characteristics. Additionally, all the other NLP methods involve word counts to some extent, in that they must first tokenize the text into different words and then count kinds of words or number of occurrences, albeit more specifically. The major limitation of word count analyses is that they really don't provide a lot of information on their own about the content or meaning of texts. Therefore, they rely on at least another level of inquiry (such as syntax or semantics) to be useful in most research questions.

4.2.2. Tagging words part 1 - Grammar

A second common method of language analysis is Linguistic Inquiry and Word Count (LIWC) which primarily looks at basic syntax of sentences and speech. LIWC involves tagging parts of speech (e.g., adverbs, adjectives, or pronouns) and either counting the number of times each of these types are used or trying to find a pattern in the order they are used. Again, it is a relatively simple measure that is relatively easy to calculate, and it has been used for a variety of research areas[160-162, 181, 182]. For example, the kinds of pronouns someone uses (e.g., "I" versus "we" or "us") can be used to calculate the probability that they are a Democrat or Republican or what position of power they hold within a social ranking or work hierarchy [182]; additionally, a unique combination of the number of verbs, prepositions,

quantifiers, and conjunctions have been shown to dissociate honesty from deception in certain language tasks [182]. While useful, language is more than just grammar and the identity of the words do not place any weight in this type of analysis (and thus valence and definition do not get considered in these instances). Therefore, LIWC syntax analyses cannot quantify the meaning of language outside of correlations with other behaviors.

4.2.3. Tagging words part 2 – Valence and meaning

Third but related, LIWC has also been used to quantify basic semantics via tagging words according to pre-specified connotations that are either dictionary/program-driven or subjectively determined. For example, one could tag the word “happy” as positive valenced and “sad” as negative valenced, “think” as cognitive and “feel” as emotional, “perplexed” as complicated versus “confused” as simple, or “evil” as intense versus “mean” as mild language. After tagging, one can group words according to tags and count the number of words in each kind of tag, calculating differences in time or between two groups based on these parameters. For instance, Toivonen and colleagues [163] found the 50 most frequently-used words that described emotional experiences and through hierarchical clustering based on subjective ratings of similarity and valence of each concept, was able to detect differences between men and women in the kinds of words they used and how they viewed those words. As another example, Schwartz et al [164] analyzed 700 million words, phrases, and topics collected from Facebook messages and was able to correlate semantically tagged words with various measure of personality and self-reports of age and gender. While this type of analysis gets us closer to capturing meaning, it has a major limitation in that it is often biased because it ignores context. A great example of this is the word “pretty” – without context, we first don’t know if this is an adjective or an adverb, which allows us to further narrow down its meaning. In the case of the first possibility, it’s likely positively valenced (“the pretty flower” or “the pretty woman”). In the second case, however, it’s harder to tell the valence because it is likely an intensifier of an adjective or verb and here context is extremely important (“he was pretty happy” or “the pain was pretty bad”). Furthermore, even with context, tagging based just on valence or intensity might still likely fall short, as in the case of “time flies pretty fast around here”, where “pretty” is still an intensifier but appears to be neutral because we lack information about whether time passing quickly is

good in this context, bad, or neither. Additionally, tagging of these categories can be dependent upon the software used and the dictionary databases they may use as a references for tagging.

4.2.4. Latent semantic analysis

This brings us to the fourth kind of quantitative language method used, one which has been growing in popularity since the 1990s [183] and is the primary method chosen for this dissertation. This is Latent Semantic Analysis (LSA), also sometimes referred to as LSI (for indexing). LSA analyzes a large set of documents to find underlying hidden (i.e., *latent*) meaning or concepts in a smaller amount of words. In some ways, LSA represents the meaning of a word as a sort of average of all the meanings of all the passages in which it appears, and in turn, it frames the meaning of a given passage or speech as a kind of average of the meaning of all the words it contains [184]. It initially arose when trying to find relevant documents from search words entered into search engines; in order to accomplish this task, it is not enough to just compare words between what is typed and what is being looked for. Instead, the real comparison is between the meaning or concepts behind the words of interest. LSA tries to do this by mapping both words and documents to a shared conceptual space and then comparing them within this space. To accomplish this, there are a couple of short-cuts and simplifications that LSA takes: First, documents (which are any sources of text, including interviews, books, blog posts, etc) are represented as “bags of words”. This means that all words are jumbled together and the order of the words in the document are not important; instead only how many times the word appears in the document and other documents (indicative of how much it is used overall as part of the lexicon) are important. Second, concepts are represented as patterns of words that usually appear together in a set of documents (for example, the words “catnip”, “brush”, and “litter” might all appear in a piece of text about owning a cat). Finally, in most cases, words are assumed to have only one meaning (an idea which I will return to).

LSA first requires construction of a “co-occurrence matrix”. This matrix contains a list of words in one dimension (m) and a list of texts or documents in the other dimension (n). The cells in the m x n matrix contain counts of how many times a given word (m_i) is used in the given document (n_i), with the idea being that words used together (i.e, those that co-occur) under different documents might share some sort of context or meaning. With a small number of words or a small number of texts, a pattern may

be visible to the naked eye. However, when faced with 10s of 1000s of words or documents, there needs to be a way to remove noise, simplify the problem, and extract meaning from larger patterns. Therefore, the second step in LSA is using singular value decomposition (SVD), which is an algorithm that breaks up the m x n matrix (referred to as A) into the product of 3 different matrices (A = U*S*V) that each explain a different part of the data. U is called the "term by feature matrix" and describes the relationships between the words of interest (rows) and features (columns). S is the "singular value matrix" and consists of a diagonal of numbers (singular values) in descending order. The magnitude of these values describes the relative strengths of each of the features (that is, how much variance a feature explains in the dataset, similar to factors in a PCA). Finally, matrix V is the "document by feature" matrix (usually shown transposed as V' or VT to be a feature by document matrix). In instances of LSA that have a relatively small # of documents (<100), V describes the relationship between the features and the documents; in instances of a large corpora of documents, this interpretation is slightly different and will be discussed shortly.

SVD thus finds a reduced dimensional representation of our initial matrix (A) that keeps as little of the semantic structure as necessary while empathizing the strongest relationships. The key is figuring out the optimal number of dimensions to use, as the resulting matrices are truncated to "k" dimensions as part of the third step in LSA. Too few dimensions (small k) can result in important patterns and meaning being left out, but too many dimensions (large k) can cause noise to be kept in or reintroduced via random word choices or inappropriate word usage. Previous research on LSA dimension size indicates that the number of dimensions typically selected is between 100 and 500, and Landauer and Dumais's seminal paper showed that they obtained the best performance when truncating at around 300 dimensions, which has become a sort of field standard in this approach [183, 184]. Thus returning to matrix size and interpretation, one might now imagine an m X n matrix that has just undergone SVD and has been shortened to k dimensions (n = k). If k < number of documents entered from the original matrix (which often is, since corpora of 1000s of texts are commonly used), then matrix V cannot be interpreted simply as a feature by document matrix. Moreover, if n (also k) = m, matrix V will be square; however if m > n (and therefore greater than k), then matrix V will be truncated at k (length n) for both dimensions (also a square matrix) because there cannot exist more documents than features (since features were

determined by documents). Therefore, V in these instances will still describe documents by features but the interpretation will be in terms of features (i.e., a feature by feature matrix).

The resulting matrices can be multiplied together in different ways in order to test the data. For example, the product of U and S allows you to see which words are best associated with which factor and how strongly associated they are. Although feature 1 is always the strongest (like factor 1 is in a PCA), it is usually so because it captures the mean of the data or the most common word used, and thus it is often ignored (taking features 2 and 3 instead as having new information). Additionally, using the features from LSA, one can calculate other linguistic measures, including semantic distance of whole bodies of text or speech to words of interest. This is done by taking the dot product of all k features of the concept (word) of interest with all the k features of the words used in a text or by a participant, averaging across all words to come up with a mean semantic proximity score for that text. This calculation essentially represents how semantically similar a concept, word, or theme is to a collection of words (how closely does this word stand for these other words, on average).

LSA has many advantages. First, it allows for mapping words used in speech or text within a larger semantic context and in the same semantic space for comparisons. Second, it provides a method for data reduction, where fewer semantic dimensions will result in less noise and more robust information. Finally, because it is inherently global (as opposed to looking at words within a local context, such as the current sentence), it can find trends and patterns in words usage that might not be as readily apparent in other modes of inquiry. Importantly, LSA is an approach that does not depend on explicit semantic representation, nor does it need word-for-word correspondences between terms, so people with different vocabularies or different ways of explaining things can still be compared to one another and their speech can be represented side-by-side [185].

However, it also has some caveats. First, LSA assumes that a word's usage in a passage of text has a Gaussian distribution with a Frobenious norm [179]. This is not always the case, as many words have been shown to have a Poisson distribution and thus LSA might not always provide the best fit in its algorithm [179]. Second, and as eluded to previously, words are assumed to have only one meaning, which we already know is not always the case. Polysemy describes the phenomenon by which a term or a symbol has multiple meanings (e.g., board, top, left, bank, table, crane, etc). This is problematic

because it biases semantic connections or calculations for some words and not others. In the example of the word “bank” which might be on the side of a river or might be a place to store money, LSA would put the term close in between “money” and “river” as it should, but it might also in turn calculate the semantic distance between money and river as being closer than what it actually is because of the attributed singular definition of “bank” Some attempts have been made by researchers to control for this by using an online dictionary, Wordnet (which has these multiple definitions built into it) as their corpora of interest [158], but this is not done often enough in the field and isn’t yet standard practice.

Similarly, although LSA often throws out the 1st feature, it can still be influenced by over-used words or words that are part of common phrases or colloquialisms and thus not weighting words is potentially problematic because it can result in “noisy” features. Some instances of LSA do weight words by calculating a text’s TF-IDF (Term Frequency-Inverse Document Frequency). This corrects for the number of documents, the number of times a word appears in all the documents and the current document, and the total number of words in the current document, essentially making rare words (those that hold privileged or specific meaning and importance and which occur in only 5% of a set of texts) weighted more heavily than common words (which might occur in 90% of texts in a set of documents) [186]. This calculation is done from the beginning on the raw matrix co-occurrence counts (modifying SVD inputs). First, the word frequency is converted into its log form and the entropy of each word is computed as p*log(p) over all entries in its row (with each cell entry then divided by the row entropy value). Thus the word occurrence is simultaneously directly weighted by an estimate of its importance in the passage and inversely weighted with the degree to which knowing that word provides information about which passage it appeared in (so words that are used a lot get downsized because they don’t provide important classifying information) [184]. An additional weighting-like option involves calculating different orders of co-occurences through a term by term matrix (U*U’) and the semantic proximity measurement based on vector distance mentioned previously, both of which provide information about similarity of terms. For each pair of words in this term matrix, the order-of co-occurence is computed by tracing “co-occurrence paths” where each word is considered a node in a graph, and semantic paths can be drawn between nodes/words that are in the same document [187]. Order is assigned generally with n + 1, where n = number of hops needed to connect the nodes in the graph. Nodes that are connected are considered

first-order pairs with no hops in between, whereas nodes that have one intermediate word between them are second order, with two intermediate words as third order, and so on. Calculation of co-occurrence order can also be derived by first binarizing the term-by-term matrix and then multiplying it with itself for n-successive times. Thus if A is the original matrix, the values in (i, j) entry of A^n will represent the number of paths of length n between term i and j. The analyses can be done for each of the different ordered representations to determine which is best (that is, which has the shortest characteristic path length between terms). Again, however, weighting practices like these are still not widely used. Finally, LSA also doesn't make use of word order, and in turn does not utilize information from syntax, logic, or morphology. While it can still extract correct reflections of passage and word meanings accurately despite missing this information, it's likely that some incompleteness or error could and does result on some occasions because of this [184].

4.2.5. Combining methods

The final method of quantifying language involves combining some or all of the above analyses together to try to make up for areas where one is lacking. As one example, one study combined syntax- tagging with semantic similarity of words to identify important phonetic properties related to the intensity and emotional content of action verbs [188]. As an additional example, researchers have combined semantic context with syntax to look at directed speech network properties in schizophrenia and bipolar patients and in participants who ingested various one of 3 substances [165, 166]. Network analyses like these indicate that graph theory (the study of how parts of a system or graph are connected) can also be employed to syntax or semantics (or their combination) to further explore language. Language in particular is especially suited for this kind of analysis as (1) all of its units are interdependent and need the presence of each other to convey meaning [189], and (2) it has been shown that, like the brain and other complex networks, language is also scale-free and has small world properties [158].

The method of choice often depends on the time commitment, level of expertise, and general hypothesis/question tested, in addition to the amount and kind of linguistic data obtained. However, once set up and tested, it can be made relatively automated through programmed pipelines and scripts.

4.3 Importance of studying language in the context of chronic pain and placebo analgesia

While neuroimaging represents an extremely important methodology due not only to its predictive capacity but also especially to its ability to clarify underlying mechanisms of chronic pain and placebo response, performing neuroimaging scans for future analyses is not necessarily clinically ideal for a couple of reasons. First are logistical concerns– fMRI scans are expensive and functional images are difficult to interpret without the proper software tools, server space, processing time, and sophisticated statistics. Therefore, relying primarily on neuroimaging methods to predict placebo responders is simply not a feasible solution for clinical trials or pain clinicians due to time, money, and lack of expertise. Second, while self-report measures are cheaper and simpler in terms of administration and calculation, as mentioned in **Chapter 1,** questionnaires looking at placebo response and personality in healthy participants have produced very few reproducible results across trials, placebo rituals, or clinical cohorts (so it is unknown how consistent these measures will be in chronic pain patients). Therefore, there still is no tool that exists that can capture and predict placebo propensity while remaining easy to implement and relatively cost-effective, and there is no method that has been shown to be strongly related to existing psychosocial and neurobiological measures of placebo response in chronic pain patients. Language may prove to be the ideal candidate for such an endeavor, as it theoretically should be able to capture aspects of both questionnaires and underlying neurobiology.

Systematically and quantitatively studying what chronic pain patients say and how they say it has enormous potential utility for the field of pain and many clinical applications. Linguistic analyses could help researchers and physicians better understand the chronic pain experience and in turn why some patients recover or respond to various medications in time. Additionally if language can serve as a simultaneous surrogate for both psychosocial measures and brain response or anatomy, interviews may be able to be used as an alternative tool for prediction that not only meets the main requirements of cost-effectiveness and easy implementation, but also is both physiologically and psychologically valid.

Section 5: The memory of pain

Everyday existence is a mixture of experiences and memories, the confluence and interaction of which guide internal thoughts and future actions. This relationship between momentary experience and

memory remains one of the most fundamental topics of inquisition for both neuroscience and philosophy, having major implications for human behaviour and decision-making, as well as for more complex phenomena, such as identity and selfhood. As such, unraveling mechanisms regarding the interaction between lived experience and recalled experience is a cornerstone from which neuroscience can inform and advance psychology.

5.1 Mismatch between experienced and recalled pain

However, the biological and psychosocial bases of such interactions remain essentially unknown. This is unfortunate since a better understanding of how memory and experience influence each other would be of great clinical benefit, especially in the realm of pain. One of the main challenges physicians face is determining the amount of pain experienced by their patients. In addition to the intrinsic difficulty in describing and quantifying pain, it has been repeatedly shown that memories for painful events are often inaccurate - when asked to recall a past painful event, people tend to overestimate their pain, with the intensity reported to be more severe than that which was actually experienced [190]. The magnitude and direction of the discrepancy between remembered pain and actual pain are dependent upon many factors, including the emotional context under which the pain was experienced and later recalled [190-196], the individuals' personality traits and mood [197, 198], and the participants' pain history and previous experience with pain [190, 199, 200].

The psychophysical properties of acute experimental pain also account for a large proportion of the error in remembering pain. An influential study by Redelmeier and Kahneman [201] demonstrated that patients' memories of the amount of discomfort reported after a minimally invasive procedure was determined primarily by the intensity of pain at both the procedure's worst and most recent episodes, a phenomenon now known as the "peak-end rule" (**Figure 7**). The authors reported that the initial amount of pain, the overall total amount of pain experienced in time, and the duration of the procedure all had little effect on the patients' retrospective ratings. These observations have since been replicated in other studies, experimental designs, and participant cohorts [202-204].

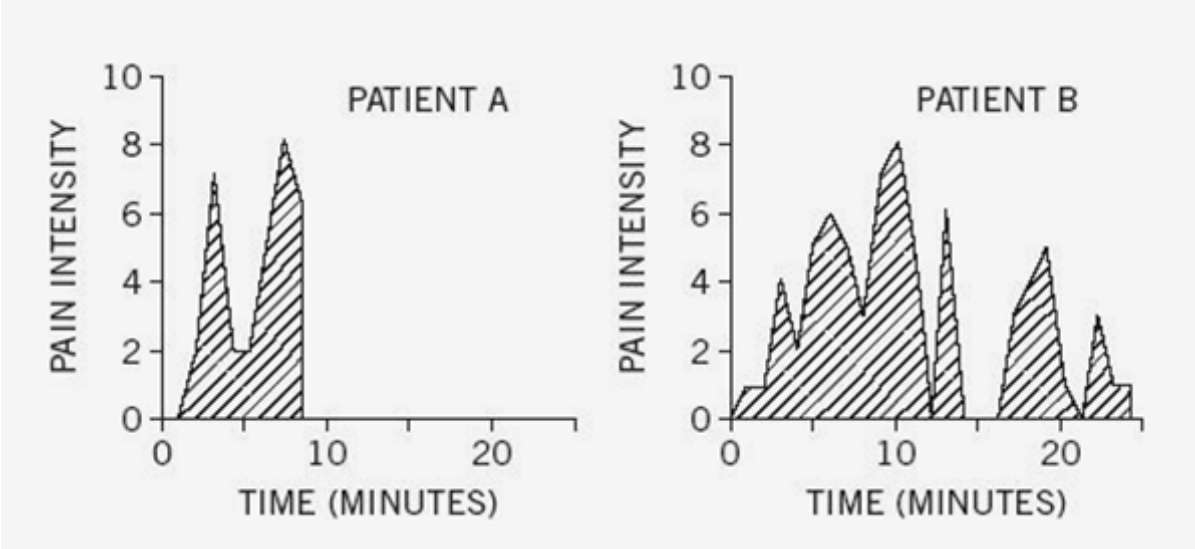


Figure 7: An example of Kahneman's peak-end rule. In 1996, Daniel Kahneman and Donald Redelmeieir studied 154 patients undergoing a colonoscopy. They asked these participants to rate their pain on a VAS scale (0-10) every 60 seconds during the procedure. The procedures varied in length, with the shortest being 4 minutes and the longest 69 minutes in duration. Shown here are plots of pain intensity reported by two of these patients during the colonoscopy (taken from Kahneman's "Thinking Fast and Slow" [203]). At the end of the procedure, participants were asked to rate the total amount of pain they had experienced during the procedure and report this number to the researchers. Now consider this question: Assuming that the two patients used the VAS scale of pain similarly, which patient suffered more? And why? Most people say B because he spent at least as much time as patient A at any level of pain and the AUC is clearly larger for B as well. Of course, B's procedure also lasted a lot longer than A's did, too. However, the patients did not display this type of result; instead A actually reported more pain than B, illustrating that the real-time ratings and restrospective assessments were systematically different from each other. Kahneman later demonstrated that this recall bias was due to patients' memories of the level of pain reported at the worst moment of the experience and the level of pain at its end, which he later called the "peak-end rule". Importantly, duration of the procedure had no effect on its recall. Applying this to the patients, we can see that the worst rating (8/10) was the same for A and B – they had the same peak – but the end was different (7 for A and 1 for B). This means that the peak-end average was therefore 7.5 for patient A and 4.5 for patient B, explaining why A retained and rated a much worse memory of the procedure than B.

The memory biases described above have also been documented in chronic pain patients, with evidence that long-term pain is remembered less accurately than acute pain [190, 200] and that people with persistent pain report intensity of previous pain less accurately than healthy people [205]. These inaccuracies in the recall of spontaneous episodes of chronic pain can become worse over time and even impact memories of treatment efficacy [199]. Chronic pain populations also have higher rates of psychological co-morbidities and mood disturbances, which can in turn influence memory. Studies have demonstrated that increased depression, elevated levels of emotional distress, and/or sustained

presence of negative mood can all result in the overestimation of recalled pain in patients with various kinds of chronic pain conditions [206-209].

5.2 The importance of accounting for memory bias in RCTs

The extent to which the combined effects of psychophysical pain and mood properties explain memory bias in chronic pain remains unknown. Importantly, a neurological substrate responsible for this bias has yet to be identified in either healthy or patient populations. From a placebo standpoint, it's important to understand which regions are involved in this bias since the anatomy and neural circuitry of memory at least partially overlap with that of chronic pain perception and placebo response (and thus may influence one another). Additionally, since most clinical trials rely on retrospective pain assessments or reports of treatment efficacy, investigating the extent of bias in chronic pain populations would have utility in clinical decision-making (e.g., when to start treatment, what dose to give, or whether to even include someone in a trial). Furthermore, placebo analgesia in particular is likely influenced by this kind of bias, in that subjects often rely on the memory of their pain prior to treatment in order to estimate their subsequent improvement (or lack thereof) [210]. Studies have demonstrated that when retrospective evaluations are used, the magnitude of the placebo effect can be 3-5 times greater than the effect calculated from real-time ratings [211] , and that there is often a greater memory bias at baseline in future placebo responders than in those who do not respond [212].

Section 6: Model building and machine learning

All studies in this dissertation build some kind of model for explaining variance in the data, and some also make use of machine learning techniques. The scope of these topics - models and machine learning - is incredibly vast and could make up hundreds of dissertations; as such, they are not covered in much detail. However, a few of the key elements in model development are described below, as are brief explanations of each of the techniques used in the thesis, as they are important in understanding the intermediate and final model analyses in each of the studies presented.

6.1 Definitions and basic knowledge

For the purpose of this dissertation, a model can be defined as a description of a system (e.g., the neurobiology and psychology of chronic pain patients) or behavior (e.g., placebo response) using mathematical concepts, statistics, and language. In the following studies, the models will be "predictive", meaning that we will use statistics on a set of independent variables (predictors) to predict a dependent variable (outcome) that we expect is related to these predictors. Depending upon how one defines it, predictive modeling overlaps significantly with the field of machine learning. Machine learning is part of the computational sciences and is a term used to describe the study and construction of algorithms that can learn from and predict data. One of the main goals in machine learning is to build and train a model from sample inputs, output this model as some kind of equation or formula, and test it on a new set of input data (although this goal may change depending upon the question asked and the starting knowledge of the data). Importantly, machine learning makes use of the computational power of modern day computers, with the idea being that a machine can sift through extremely large amounts of data at speeds exponentially faster than a human could ever dream of achieving, and in doing so, can produce more complicated and accurate models of the world and uncover insights that might not have otherwise been discovered by human eyes/actions alone.

For many predictive applications, the outcome of interest is binary (e.g., given some criteria, did a person respond to a placebo or not?), and the information from the model can then be expressed as a probability [213]; this is known as classification. For other applications, the dependent variable is a continuous variable (e.g., pain ratings or memory scores), and in these cases, the model provides information about the amount of variance that the set of independent variables explain in the outcome. Although aimed at predicting the future, many predictive models are applied to any type of unknown problem regardless of when it occurred. In **Study 1** of this thesis, we analyze placebo response propensity in chronic pain patients using data collected prior to randomization and therefore *before* any response occurred. **Study 2** inputs language parameters into a predictive model explaining placebo response, even though the language data was collected after responses had taken place – in instances like these, the models aren't always predictive in the traditional sense, but they can still be used to assess how well parameters can distinguish between two different groups. In **Study 3**, we input rating data and

neuroimaging data into a model explaining variance in the memory of pain, and we then use this model in a separate group of participants to predict their reported memories before looking at their actual recalled values.

6.2 Multivariable linear regression

For continuous scalar variables, some form of linear regression is often used as a predictive model. In this case, the dependent variable (response) is thought to be a linear function (combination) of all of the independent variables (regressors) inputted into the model. Expressed as an equation:

$$y = xB + E,$$

Where y is the actual observed response (dependent variable), B is the vector of unknown parameters (coefficients), x is the set of regressors (explanatory independent variables), xB is the predicted response, and E is the error accounting for the different between y and xB. Additionally, there is always a constant term included in the regression that indicates where the fitted regression line crosses the y-axis (intercept, B_0).

Interpretation of linear regressions usually involves looking at the R^2 value of the final equation and its overall significance level. The R^2 value (also known as the coefficient of determination) indicates how close the data are to the fitted regression line and is a measure of the percentage of variation in the response variable that is explained by the model (i.e., explained variation/total variation). R^2 values are between 0 and 1, with 0 being that the independent parameters explained none of the variance in outcome (i.e., they did not affect the response) and 1 being that all the variance in response is explained by the regressors (and the model is completely sufficient). In general, the higher the R^2, the better the model fits the data (although this is not always the case and caution must be taken to ensure that these values are not biased). The p-value of the regression tests the hypothesis that the slope of the fitted line is equal to zero (meaning that there is no relationship between the response and the regressors); a low p-value ($p<0.05$) in the final model indicates that this null hypothesis can be rejected and suggests that the predictor variables are associated with changes in the response variable in a way that is not by chance alone. Additionally, investigating the B values for each of the regressors provides important information about the model. Specifically, these values indicate the direction and magnitude of expected change in y

for a one-unit change in x_1 when the other covariates ($x_2 \ldots x_n$) are held fixed. These B's, along with their corresponding statistics (p-values), can give clues as to the relative importance or power each variable has in the model (and, in turn, in the outcome variable).

Linear models assume that the data is normally distributed (Gaussian), denoised (without significant outliers), and independent (not correlated with one another).

6.3 Multivariable logistic regression

Logistic regression is another common technique used in computational neuroscience and psychology for predicting outcomes. It is named because of the function that it uses – the logistic or sigmoid function –that was developed in the 19th century by statisticians studying population growth [214]. They noticed that in the early stages of a developing society, population grew quickly and steadily but over time eventually reached a point of near plateau (a maxing out of capacity). To fit this data, they utilized an s-shaped curve that was able to take on any real-valued number and map it onto a value between 0 and 1. Like linear regressions, logistic regressions also use an equation in their representation, with x, y, and B values and a constant term. However, a key difference between these equations is that y in logistic regressions is assumed to be a binary outcome (0 or 1) as opposed to a meaningful number. Expressed as an equation:

$$y = e^{(B0 + BX)} / (1 + e^{(B0 + BX)}),$$

Here, the B values (constant term and coefficients) and x values are interpreted in the same way as they would be in a linear regression, with the only exception that they are exponents of Euler's number (e), which is the base of the natural logarithm and is equal to approximately 2.718; this arrangement transforms the linear relationship of the variables into a logistic equation. In this case, y represents the probability (p) of the combination of regressors leading to one of two classes. If we let f(x) stand for the linear combination of regressors, we can also think of the equation in this way:

$$(p) = (e^{f(x)})/(1+e^{f(x)})$$

A criteria for p is usually set to define which class a given combination of regressors would best predict; sometimes this threshold is arbitrarily determined and other times it is apriori determined by previous or

current data. For example, if we expect 50% of our patients to respond to a placebo, we might set the criteria to p = 0.50; in this case, if p >= 0.5, patients would be classified as a responder (class 1) and if p<0.5, they would be classified as a non-responder (class 2).

Outside the obvious use of logistic regressions to obtain classification probabilities, interpretation of logistic regressions is similar to that of linear regressions in that R^2 values can also be used to assess the amount of variance explained, and coefficients and p-values can also inform about the strength of the model and the relative contribution of each of the explanatory variables. In the case of B values, the strongest coefficients are those that result in the model predicting a value very close to 1 (class 1) *and* a value very close to 0 (class 2). To get these, a search procedure seeks values that minimize the error in the probabilities predicted by the model to those that are actually in the data (i.e., it tries to find values that result in the best match between predicted and observed). Like linear regressions, logistic models also assume that the data is de-noised, Gaussian, and independent.

6.4 Naïve Bayes

Another common machine learning technique that is complementary to logistic regression is a Naïve Bayesian (NB) approach, which is named due to its reliance on Bayes Theorem, proposed by Reverend Thomas Bayes in the 18th century [215]. Baye's critical insight was that in order to determine how probable a certain outcome was, one would need to also determine how probable that same event was if different scenarios were true. This type of probability is now called a likelihood, and it is the foundation of Bayesian principles of data analysis. Bayes Theorem thus describes the likelihood (probability) of a certain outcome provided a set of input parameters (features) and in turn allows us to predict a class given a set of features (which are described in terms of their probabilities as well). In an NB classifier (or any classifier that uses Bayesian inferences), these features are assumed to be unrelated to each other, that is, each parameter is thought to contribute independently to the probability of a certain outcome occurring. For example, the features "furry", "striped", and "around 9 lbs in weight" would be viewed as independent contributors to the probability that an animal is a cat, regardless of any correlations between the features. Such independence is often not true of features, which is why the

algorithm is labeled “naïve”, but despite this shortcoming, NB classifiers are still able to outperform other more complicated machine learning algorithms.

A key component of the theorem is calculating posterior probabilities from the data. In its most simplest form, this can be done with the equation below:

$$P(c|x) = (P(x|c) * P(c)) / P(x),$$

where P(c|x) is the posterior probability of a class (c) given the feature (x); P(x|c) is the probability of the feature x given the class c (this is likelihood measure, which is opposite of the posterior probability); P(c) is the class prior probability (how many of the class c is expected in the total sample), and P(x) is the predictor probability (how many of the feature x is expected in the total sample). Importantly P(x) and P(c) are calculated without regards for each other. If there is more than one feature, the equation can be re-written to multiply the probabilities of each feature together. In an example with 3 features (x1-x3), this might look like:

$$P(c|x1,x2,x3) = (P(x1|c) * P(x2|c) * P(x3|c) * P(c))/ (P(x1) * P(x2) * P(x3))$$

Like the above models, in addition to assuming independence, NB also assumes Gaussian distributions of the features.

6.5 Support Vectors

The last type of modeling method this thesis addresses is that of support vector machine (SVM) learning. An SVM is a kind of supervised learning model that takes in specified inputs (training data) and specified desired outputs (binary) and asks the machine to come up with a rule or equation that maps the inputs to the outputs as best it can. An optimal algorithm would allow the model to generalize its outcomes to unseen data (referred to as testing data, which is either purposefully left out or not yet collected). To accomplish this, SVM makes use of both linear and non-probabilistic strategies. Given a set of training data whose classifications are known and specified, SVM first plots each piece of data as a point in space, finds a line or hyperplane that best separates the points according to their class, and then assigns new points to one class or the other according to the line. Importantly, the line or plane (called a classifier) is fit in such a way that the distances from the closest point in each of the two classes will be the farthest apart (in other words, it attempts to find the biggest gap between defined groups). The two

points closest to the line are referred to as the support vectors, and the gap is referred to as the margin and is typically twice the distance to the support vectors (i.e., it's exactly in between the closest data points). The side of the line that the new testing data lands on determines the class that it is assigned. Often, the training data is so complicated that a simple line is not sufficient to divide the classes easily. For these cases, some level of non-linear optimization is needed, including transforming the data into a higher-dimensional space, to aid in finding the hyperplane that best fits the classes [216].

6.6 Model validation

After model development, a quantification of its performance is needed so that one can judge whether the model is adequate for its purpose and/or better than an existing model [217]. One of the main things validation analyses inform us about the models we test is the extent to which they are over-fit. Overfitting refers to the phenomena that happens when models are too complex, resulting in a model that produces noise or error instead of the underlying relationship. Intuitively, one might expect that the more parameters inputted into a model, the better it will perform since it has more information to go off of; thus it might seem like a 20-factor model might be better than a 2- or 3-factor model. However, things are not usually so simple. While it is true that the fit of the model should improve with additional parameters and with this, more variance in the outcome will be explained, a better fit does not guarantee better prediction [215]. This is because the model might be biased by the data in its training set, in turn becoming too sensitive to the particular data points that happened to be observed in the current sample (which might not necessarily be observed in a different sample). If models have too many features relative to the number of their observations, it can become so finely tuned to that specific data set that the solutions it produces in new data sets are highly variable and noisy. The more factors in a model, the more points that a novel data set must fit; the less factors, the more likely that slight variations from different inputs would still be able to fit the essential pattern [215]. A good model falls somewhere in between – not so simple as to fail to capture an underlying relationship in the data but not too complex so as to introduce error. Model testing can combat this potential problem by first detecting the presence of overfitting with methods like cross-validation (explained below) or by penalizing more complex models, as in regularization and Lasso methods (not explained here) [215].

A primary aim in model validation is to have at least internal prediction validity (meaning that the model is strong and reliable for subjects within the same sample or underlying population). However, if a model is to be considered truly valid, this predictive capacity would preferably generalize to different samples (individuals with the same condition that the model did not train on) and/or to related but different populations (e.g., a model trained on CBP would also have some predictive accuracy in OA). For these reasons, validation of a model on a fully independent and/or external dataset is the best way to assess its performance. Unfortunately, there are many cases where it simply is not possible to have a validation or test data set. In these cases, the validity of the model must be tested on its own data in various ways to essentially guess at how well it will fit data that it hasn't seen. One of the most common ways to do this is through cross-validation (which is used in **Study 1**) [218]. Cross-validation investigates how well a model fits the data that it's given as well as how well it generalizes to data it hasn't seen. To do this, cross validation methods "hold back" or "leave out" parts of the data temporarily, building the model on the remainder of the data and later testing it on the data it kept out. Often, this procedure is repeated many times, sometimes over all possible iterations of the data. The number of data points to leave out depend on computational power/time and sample size. A common but computationally intensive option is exhaustively testing all the possible ways to divide the data into discovery and validation sets; these are referred to as Leave-p-out-cross-validation (LPOCV), where p is the number of observations used as the validation set. With larger sample sizes, LPOCV can become difficult to calculate. Leave-one-out-cross-validation (LOOCV) sets p = 1 and has the same advantages as LPOCV but with faster processing time; for each iteration of the data, you leave one observation out and test it on the model later. Accuracy in determining the left out data point is determined at each iteration and averaged over all iterations to obtain an approximate percent accuracy for the model. LOOCV is used in **Study 1**. Other types of cross validation include k-fold CV (dividing the sample into k equal-sized subsamples, training on one and testing on the other) and repeated random subsampling (randomil splitting the data into training and testing data with no dependence on number of folds or observations) [219].

Section 7: Thesis aims and hypotheses

The overall aim of this thesis is to better understand the biological, psychological, and linguistic mediators of clinical placebo analgesia in chronic pain and develop a model that can be used as a tool to detect placebo propensity in future clinical trials and cohorts.

The aim of **Study 1** is to identify and characterize neurological biomarkers and personality traits that predict placebo response in chronic low back pain patients. We conducted an 8-week-long blinded randomized control trial (RCT) investigating mechanisms of clinical placebo analgesia. This trial had daily pain and mood ratings, an extensive battery of questionnaires, and multiple functional and anatomical neuroimaging scans collected longitudinally to capture placebo responder properties. We hypothesized (1) that brain regions involved in emotional and reward processing, memory, pain, and cognitive control should show differences, anatomically and/or functionally, between responders and non-responders at baseline prior to treatment, and (2) that psychological traits and states important for emotional and pain regulation would also differentiate future placebo responders from non-responders.

Study 2 had two primary aims: the first aim was to investigate the extent to which semantic language properties could capture differences in placebo responder and non-responders with chronic pain. To answer this, we conducted a semi-structured, open-ended exit interview as part of the same RCT in **Study 1**. We hypothesized that placebo responder narratives would differ significantly from non-responder narratives in their semantic relationships with concepts including emotional awareness, coping behaviors, pain experience, and expectations, and that a unique combination of these semantic concepts could predict response. The second aim was to demonstrate that the semantic differences found in Aim 1 corresponded to neurological and psychosocial predictors of placebo response in these patients. We hypothesized that if there were robust semantic differences seen between groups, then these language parameters should be significantly correlated to underlying and pre-existing brain organization and function, as well as to personality traits and emotional states related to placebo response. We also hypothesized that if language could truly be conceptualized as a intermediary between the brain and the mind (or as a variable that captured aspects of both neurological function *and* psychological processes), then the semantic differences seen between responders and non-responders should be able to identify

the same functional connections seen previously, as well as additional brain signatures that might also be related to or mediating of placebo response.

Study 3 was unanticipated and resulted from exploratory analyses of the data collected in Study 1. The aim of this study was to identify and characterize neurological, psychological, and psychophysical mediators of pain recall bias in chronic pain patients. Utilizing data from the same RCT as **Studies 1 and 2**, we combined daily measures of pain and mood collected using the smartphone app, psychological scores from the questionnaires, and morphometry of the hippocampus to explain pain memory bias in CBP. We hypothesized that CBP patients would show a discrepancy where their recalled pain at the end of the rating period would be significantly higher than the actual pain intensity they experienced while rating. Given the importance of the hippocampus in memory encoding and retrieval, and its role in the development of chronic pain [45, 46, 50, 91], we further hypothesized that memory biases seen in participants would be associated with differences in the anatomy of the hippocampus as well as personality characteristics.

CHAPTER 2: METHODS

Participants

129 participants with chronic low back pain (CBP) were initially recruited from the general population and clinical referrals via hospital databases and advertising in the community. To meet inclusion criteria, individuals must have been 18 years or older with a history of lower back pain for at least 6 months; this pain must have been neuropathic (radiculopathy confirmed by physical examination was required) with no evidence of additional co-morbid chronic pain, neurological, or psychiatric conditions. Individuals must have agreed to stop any concomitant pain medications and must have indicated ability to use a smartphone or computer to monitor pain twice a day. Additionally, the enrolled patients must have reported a pain level of at least 5/10 during the screening interview, and their averaged pain level from the smartphone app must have been higher than 4/10 during the baseline rating period (explained below) before being randomized into a treatment group. Finally, for safety precautions, clinical measurements taken at Visit 1 must have been within a pre-specified healthy range and all participants must have passed the MRI safety screening requirements at each scanning visit.

Figure 8 illustrates the flow of patients through the clinical trial. From the initial 129 chronic back pain (CBP) patients recruited in the study, 4 individuals were assessed for eligibility but met exclusion criteria before consenting. Of the enrolled 125 patients, 43 failed screening due to meeting exclusion criteria at Visit 1 or during the 2-week baseline period between Visits 1 and 2. The remaining 82 patients were randomized into one of three groups – no treatment (n=25), active treatment (n=5), or placebo treatment (n=57). Of the no treatment group, n=5 were either discontinued from the study or lost to follow up; of the placebo treatment group, n=11 were either discontinued or lost to follow-up, with an additional 2 participants being excluded from final analysis due to having an average pain rating values during baseline below 4/10. Note that the inclusion of active treatment group was used only to ensure that the double blind for placebo treatment was maintained for the duration of the study. Therefore, the 5 participants randomized in the active treatment group were not analyzed.

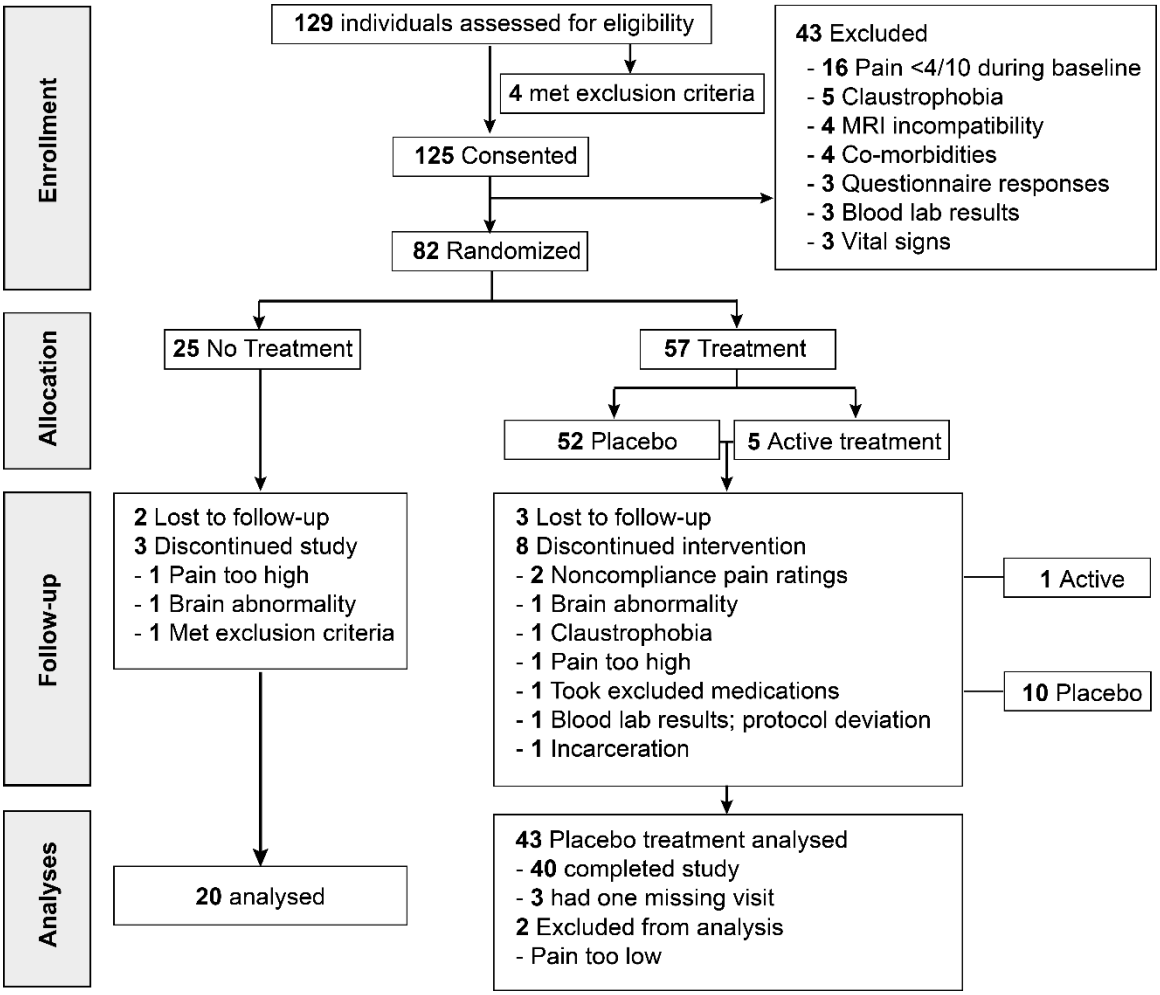


Figure 8: CONSORT diagram. Disposition of all study participants from study entry to study completion, including those who screen failed, those who discontinued, and those who successfully completed all study visits. Of the 129 people who were screened, 4 met exclusion criteria prior to consent; 125 individuals were consented and entered into the study. Of these, 43 screen-failed between visit 1 and visit 2 due to the reasons listed; the remaining 82 people were randomized into a no-treatment (n=25) or treatment (n=57) group according to a block design. Of the 25 individuals in the no treatment group, n = 5 were either lost to follow-up or were discontinued post-randomization due to reasons listed, leaving a final n=20 that successfully completed all 6 visits of the study and were subsequently analyzed. Of the 57 treated individuals, 52 were allocated to placebo treatment and 5 were allocated to active treatment (Naproxen + Esomeprazole) in a double-blind fashion. After randomization, 10 placebo-treated individuals and 1 active-treated individual were either lost to follow-up or discontinued for reasons specified in the diagram. The remaining n=4 active treatment participants were not analyzed as their only purpose was to aid in maintaining the double blind. The remaining n = 42 placebo participants successfully finished all study visits; however, 2 individuals were found to have baseline pain ratings not meeting eligibility criteria and were subsequently excluded, leaving n=40 placebo treated individuals to complete the study. In addition to these participants, we added 3 people who had all scans but never made it to the final visit (i.e., they were only missing visit 6's interview and questionnaires); thus we ended up analyzing n=43 placebo participants.

The final sample size for **Study 1** included 20 CBP patients randomized to the no treatment group and 43 CBP patients randomized to the placebo treatment group; demographics for these individuals can be found in **Chapter 3**. Participants were compensated $50 for each visit completed, and they were reimbursed up to $20 for travel and parking expenses if applicable.

For **Study 2**, the numbers remain relatively the same with the exception of the 1 person in the placebo treatment group who failed to complete the exit interview, resulting in n = 20 no treatment and n = 42 randomized to placebo treatment**.** A table of demographics can be found in **Chapter 4**. For **Study 3**, 72 of the 129 initially recruited CBP patients were investigated – in addition to meeting the inclusion criteria specified above, they must have completed at least the first 2 visits of the trial (randomization into a treatment or no-treatment group was not necessary, as we were interested in memory bias prior to placebo treatment). 69 of these 72 participants had usable data across all of these modalities. 2 patients were excluded due to having no variability in their pain or mood ratings (i.e., they consistently rated their pain or mood at the same number), and another participant was excluded due to missing one of the retrospective questions (explained below). 48 patients were used as a discovery group in the analyses, and an additional 21 participants were used as a validation group. Demographics can be found in **Chapter 5.**

Additionally, for **Study 3**, a second dataset was taken from a completed longitudinal study identifying neural substrates of pain persistence, portions of which have been used in previous publications[45, 48, 79, 220]. Data from 22 healthy individuals who served as control (CON) participants and 21 individuals with subacute back pain that persisted to become chronic (SBPp) were used in the present analyses. Each group had multiple scans collected throughout their participation; for this thesis, data from the first baseline scan (scan 1) and the fifth scan (scan 5, second scan here) that occurred approximately 1 year later were used. All participants used in these additional datasets were also recruited from general and clinical populations via community flyers and ads, as well as from physician referrals and hospital databases when applicable. Eligibility criteria were comparable across all studies and patient populations. To be recruited and eligible, all patients with SBPp had to report an initial duration of pain between 4-16 weeks. Additionally, SBPp participants were diagnosed with back pain by a clinician and reported pain intensity of >40/100 on a visual analogue scale. Their persistence in pain (as

opposed to recovery) was defined by the observation that their pain levels taken at each visit did not decrease by at least 20% by the end of the study. Healthy controls must have had no current pain or history of sustained pain in the last year. As with the CBP patients in the primary analysis, both SBPp and CON participants must have had no reports or presence of co-morbid systemic, chronic pain, psychiatric, or neurological disorders (including history of head injuries and high levels of depression).

The Northwestern University Institutional Review Board approved the study (and all studies from which additional data was taken), and all participants gave written informed consent prior to commencement of any research activities. A waiver of documentation of consent was provided for the follow-up analyses in **Study 3** since they were not initially planned; those individuals who participated in the follow-up phone call (explained below) to assess memory of the study provided verbal consent prior to answering any questions.

This study was funded by the National Center for Complementary and Integrative Health (NCCIH) and was registered as a clinical trial at ClinicalTrials.gov (study identifier: NCT02013427). All study procedures were completed in accordance with federal and state guidelines.

Study design and procedures

All studies presented here were taken from one larger study, which was conducted in the setting of a clinical RCT specifically designed for assessing the placebo response. The total duration of the study lasted approximately 15 months: the first patient was seen on 11/06/14 and the last patient was seen on 02/04/16. The study consisted of 6 visits spread over approximately 8 weeks, including a baseline monitoring/screening period and two treatment periods, each followed by a washout period. The design was set up to track placebo response in time and to test the likelihood of response to multiple administrations of placebo treatment in order to optimize accuracy in the identification of responders and non-responders. Although the overall protocol included four scanning sessions, only the brain imaging data from the first scanning session performed at Visit 2 and the questionnaires administered at Visit 1 were analyzed for **Study 1** because our initial goal was to determine the psychological and neurobiological mechanisms predisposing an individual to respond a placebo treatment. Data from other

scanning sessions performed after the first treatment period will be analyzed in a separate manuscript addressing different research questions.

Studies 2 and **3** made use of different aspects of the data collected in this RCT; therefore, their study procedures are depicted in separate figures for clarity purposes only (all diagrams are displayed the **Results** section for each respective study). Briefly, **Study 2** used the interviews collected at visit 6, the questionnaires collected at visits 1,2, and 6, and the brain imaging data collected at visit 2. **Study 3** used the same data as **Study 1**, but with a different research question and additional data collection after the trial.

Randomization

The randomization scheme was performed using 2 kinds of blocks, each with 8 patients; the first block assigned 5 patients to placebo and 3 to no treatment, and the second block assigned 5 patients to placebo, 2 to no treatment, and 1 to active treatment. Each patient ID was randomly attached to a randomization code. The initial randomization included codes for the first 80 patients and was followed by a second randomization of 50 additional codes about 6 months later. For those assigned to either of the treatment groups, the allocation was performed in a double-blinded fashion: a biostatistician performed the randomization; drugs were ordered and re-encapsulated by the Northwestern research pharmacy and bottled by designated lab members; a member of the Northwestern University Clinical and Translational Sciences (NUCATS) institute matched the appropriate treatment drug with patients' randomization code; and study coordinators picked up the blinded agent from NUCATS for storage and dispensing. All drugs were stored at room temperature in a locked cabinet within the lab. The double blind for treatment groups was maintained by the identical encapsulation of the study agent - blue pills were either Naproxen (500mg) or placebo (lactose) and bi-colored pills were either Esomeprazole (20mg) or placebo, and each person assigned to treatment received a mixture of blue and bi-colored pills. This way, neither the participants nor the researchers knew which treatment the participant had received. For those assigned to the no-treatment group, no blind was maintained, as both study staff and participants knew that they were not receiving the study agent. Once approximately 50% of all participants had been entered into the study, a preliminary analysis of the electronic pain rating data was completed in order to confirm that

there were participants who were experiencing a diminishment in pain (no action was taken). Randomization details are important for **Studies 1** and **2** only.

Visit schedule

Visit 1 (**all studies**)

Visit 1 was the screening visit - at this time, participants were screened for eligibility and consented. Following informed consent, a blood sample was drawn (for a comprehensive chemistry panel, a complete blood count, and a pregnancy test if applicable), vital signs were taken (blood pressure, heart rate, respiration rate, height, and weight), and a medical professional completed a physical examination and took a comprehensive pain history. Participants were then asked to complete a battery of 29 questionnaires regarding basic demographics, pain, mood, and personality (**Table 1**).

Category	Questionnaire	Visits	Description	References
General	MQS	1	*Medicine Quantification Scale*: used to score how much and what kinds of medications participants were using before or at the start of the study; if participants were taking concomitant pain medications, they were asked to stop these for the duration of the study	[221]
	PHI	1	*Personal Health Information*: collected a general medical history, a history pertaining specifically to back pain - including causes, previous treatments, or surgeries – and information about education, income, sleep, smoking, and alcohol usage.	n/a
	Demographics	1	*NIH Demographics Form:* Collected gender, race, and ethnicity	[222]
Pain outcomes	NRS	all	*Numeric Rating Scale*: provides an additional pain rating from 0 (no pain) to 100 (worst pain imaginable); standard method of collecting perceived pain level in a clinical setting	n/a
	MPQ-sf	all	*McGill Pain Questionnaire (short form)*: measures location, duration, intensity, and quality of pain	[223]

	NPS	all	*Neuropathic Pain Scale*: measures the neuropathic components of participant's pain	[224]
	pDT	all	*painDETECT*: provides information about location, duration, intensity, and quality of pain at different time scales	[225]
Pain & emotions	CPAQ	1	*Chronic Pain Acceptance Questionnaire*: measures the effort participants put into either actively controlling their pain (activity engagement subscore) or passively accepting their pain (pain willingness subscore)	[226, 227]
	CPCI-42*	all	*Chronic Pain Coping Inventory*: measures 3 kinds of maladapative coping strategies (guarding, resting, and asking for assistance) and 6 kinds of adaptive coping strategies (exercising or stretching, relaxation, task persistence, purposeful self-statements, pacing activity, and seeking social support) in response to chronic pain; the idea is that certain illness-focused, maladaptive behaviors are associated with more disability, where as some well-ness focused behaviors with less disability or recovery.	[228, 229]
	PCS	1	*Pain Catastrophizing Scale*: assesses how much people worry about their pain and its possible causes (yields a total score and subscales of rumination, magnification, and helplessness)	[230]
	PASS-20	all	*Pain Anxiety Symptoms Scale*: measures pain-related fear, avoidance, and anxiety	[231]
	PSQ	1	*Pain Sensitivity Questionnaire:* assesses participants' sensitivity to imagined painful and non-painful stimuli	[232]
Study agent efficacy	GIC	all	*Global Impression of Change*: to score the perceived change of pain from one visit to the next on a 5 point scale, including much better, better, no change, worse, and much worse.	n/a

	TSS	3,5	*Treatment Satisfaction Survey*: lab-developed, un-validated survey asking individuals to rate on an 11-point scale (-5 to +5) how satisfied or dissatisfied they are with the study treatment and to explain why	n/a
Expectations related to medications and health	SETS*	2,4	*Stanford Expectations of Treatment Scale*: assesses participants' positive and negative expectations to the upcoming treatment, and their overall level of understand what the treatments' purpose is	[233]
	PSM	all	*Perceived Sensitivity to Medication*: measures how sensitive people think they are to medication in general, which might affect if they respond to a placebo	[234]
	HCAMQ	6	*Holistic Complementary and Alternative Healthy Questionnaire*: assesses via two subscores participant beliefs about whether alternative and complementary medicinal techniques work and should be used; we were interested in seeing if these beliefs were correlated to placebo response	[136]
	MHLC – C	6	*Multidimensional Health Locus of Control (form C)*: Assesses what factors participants believe are responsible for and in control of their health (themselves, the medical system, luck, or other people to various degrees)	[235]
Mindfulness and Emotional Control	MAIA	1	*Multidimensional Assessment of Interoceptive Awareness*: measures the extent to which someone is aware of his/her body and emotions and how well they can either focus or distract themselves from these sensations (8 subscales reflect various aspects of this awareness); additionally, scores on the MAIA have been shown to be lower in individuals with back pain	[236, 237]
	ERQ	1	*Emotional Regulation Questionnaire*: measures two kinds of strategies people use to control their positive and negative emotions (including a re-appraisal and a suppression subscore)	[238]

	ACS	1	*Attentional Control Scale*: assesses the voluntary control of attention during a variety of situations	[239]
	eACS	1	*Emotional Attentional Control Scale*: assesses the voluntary control of attention during emotionally demanding situations, which could include pain	[240]
	FFMQ	1	*Five Facets of Mindfulness Questionnaire*: a combination of many well-known and validated questionnaires, this measured the five main components of mindfulness as a skill set, which may correspond to placebo propensity (includes	[241]
Suggestibility	MISS-sf	6	*Multidimensional Iowa Suggestibility (short form)*: measures the extent to which participants can be influenced by a variety of other external and internal factors	[242]
Personality	NEO-FFI	1	*NEO Five* Factor Inventory: measures participants' scores on personality dimensions (extraversion, agreeableness, conscientiousness, neuroticism, and openness); previous research has shown that personality plays a role in placebo response in healthy controls	[243, 244]
	LOT-R	1	*Life Orientation Test (Revised)*: measurement of dispositional optimism, which has been shown to influence placebo propensity in healthy individuals	[245]
	LAQ	1	*Loss Aversion Questionnaire*: measures how sensitive participants are to a wide variety of potential "losses" in their lives; since we have already published results showing that chronic back pain patients are more gain sensitive (less loss averse), we were interested in investigating whether this trait also affected propensity to respond to placebo	[83, 246]
Affective State	BDI-Ia	1,3,5	*Beck Depression Inventory, Version 1a:* measures the extent to which a participant may be clinically depressed; a score of >=19 was an exclusion criteria	[247]

	PANAS	all	*Positive and Negative Affect Schedule:* assesses the extent to which participants are feeling a list of positive and negative emotions on the day of the visit to try to quantify the current affective state	[248]

Table 1: List of 29 questionnaires administered to patients. 29 self-report measures (47 items total if divided into respective subscales) were completed at designated visits in the study. The name and abbreviation are provided for each of the questionnaires, along with the rationale for why each measure was included in our battery. Also included are the visit(s) at which each measure was administered. An asterisk marks a questionnaire that wasn't analysed due to poor participant understanding of the questions (SETS) or poor questionnaire compliance by either selection of extremes or skipping items (CPQI).

These self-report measures were collected online via REDCap (Research Electronic Data Capture version 6.5.16, © Vanderbilt University) through a survey link sent to the participant's email address (or a back-up study email if they did not have an email account); once submitted, questionnaire answers were finalized in the database and un-editable by both participants and study staff. To best avoid questionnaire fatigue due to the number of questionnaires administered, participants were allowed to take breaks and walk around the testing room, although they were required to complete all questionnaires at the designated visit. Any remaining information, including clinical data collected at the visit, were entered manually into the database by study staff; verification of information was done via double-data entry by different staff members at a later time. At the end of Visit 1, participants were asked to stop all medication they were taking for controlling their pain. Rescue medication in the form of acetaminophen tablets (500 mg each) was provided as a controlled replacement to be used at any time in the study if their pain became too intense. At this time participants were also trained on how to use our electronic pain rating application on either the phone or the computer (explained below); if participants did not have access to either, they were provided with a smartphone and data plan for the duration of the study. The baseline rating period started at the end of this visit and lasted until they came back for their second visit approximately two weeks later.

Visit 2 (**all studies**)

If patients' pain ratings and blood lab results met inclusion criteria, they returned for Visit 2 where they completed a 35-minute brain imaging session that collected a T1-weighted image, 2 resting state scans, and 2 diffusion tensor imaging (DTI) scans (DTI not used here due to error in scanning parameters). Following the imaging protocol, the patients completed another battery of questionnaires, a subset of which were repeated from the first visit to track longitudinal changes in pain, and they were queried as to whether they experienced any change in health status since the last visit and how often they needed to take the rescue medication. Additionally, they were asked to verbally recall their average pain levels over the previous 2 weeks, over the last week, and over the last 48 hours. For example, they were asked "What was your pain level on average, from 0-10, over the last week?" Any number reported – whole, decimal, or fraction- was recorded; for individuals who reported a range of numbers, we took the mean of that range (for example, if 6-8 was recorded, 7 was used as the final answer). This self-reported recalled pain was referred to as "pain memory" and was used as an alternative outcome measure of pain levels in **Study 1**. It was also used as the primary dataset for **Study 3**.

At the end of this visit, participants were randomized into one of three groups: no-treatment, placebo treatment (lactose) or active treatment (the standard of care, which was combination of Naproxen, 500 mg bid, and Esomeprazole, 20 mg bid). Participants in the treatment groups were instructed to take a blue pill with a bi-colored pill in the morning and again at night with plenty of water, and they were asked to record this in their electronic rating app. Note that study staff never informed participants about the odds for receiving active versus placebo treatment - this is important because the current study aimed to not introduce an added layer of bias to potential responders; the goal was to have participant's own baseline expectations influence whether or not they responded to the placebo treatment. Both treatment and no treatment groups continued to receive rescue medication to use if needed, and all participants were asked to continue rating their pain and mood twice a day until Visit 3. The duration of this first treatment period was ~2 weeks long.

Visit 3 (**Study 1**)

Patients returned at Visit 3 and were queried about their memory of their pain, any changes in health since the last visit, and rescue medication usage. If on treatment, patients were asked to report

any side effects experienced and bring back any unused medication so that study staff could calculate their treatment compliance. Participants underwent another scanning session that was identical to the one completed at Visit 2 and completed another set of questionnaires with some repeated from the previous visit. At the end of Visit 3, individuals assigned to the treatment group were told that the study agent would be temporarily discontinued until their next visit so that the effects of the agent could "wash out" of their system. Again, all participants were given rescue medication to use if needed and were asked to continue using their app twice a day until the next visit. This first washout period was ~1 week long.

Visit 4 (**Study 1**)

Patients returned at Visit 4, where all measurements and procedures from Visit 2 were repeated identically, including the scanning session and questionnaires. Again, they were queried about their pain memory, rescue medication usage, and changes in health. The study agent was reintroduced to those individuals allocated to one of the treatment groups according to the same regimen described above (treatment assignment was kept the same within subjects, as this was not a cross-over study design). All participants were given rescue medication and asked to rate their pain and mood twice a day as with previous visits. Like the first treatment period, the second treatment period was also ~2 weeks in length.

Visit 5 (**Study 1**)

Following this period, participants returned for Visit 5, where all measurements and procedures from Visit 3 were repeated identically. Briefly, patients were queried about any side effects, their rescue medication usage, and their pain memory, and they were asked to return unused pills to assess compliance if on study treatment. Patients underwent the same scanning procedures as on visits 2-4. Finally, patients filled out a series of questionnaire about their pain, some of which were repeated from the last visits. As before, those participants allocated to a study agent had their treatment discontinued for a second washout period, which was also approximately 1-week long. Participants continued to use their electronic app twice daily and were given rescue medication if needed.

Visit 6 (**Studies 1** and **2**)

Patients returned for the last visit during which they were again queried about their pain memory, changes to health, and rescue medication usage. During this visit, the patients completed a semi-structured, open-ended exit interview with a designated staff member that asked them more detailed questions about their pain and medical history, quality of life, overall mood, and time in the study (details found below). Participants finished with a final battery of questionnaires and were asked to return study smart phones, if applicable. There were no scanning procedures on this visit. Any ratings submitted for the duration of the study were totalled, and in addition to their visit compensation, participants received their additional compensation for the electronic app at this time.

Exit interview implementation and design **(Study 3)**

All interviews were completed at the beginning of visit 6 before any self-report measures were administered and prior to final compensation. While this may at first seem counter-intuitive, there was a reason the interview was timed to be at the end of the trial. Talking about chronic pain can be an emotionally-charged experience due to descriptions that may arise, such as memories of suffering, lost productivity, changes to health, and medical expenses to name a few. Therefore, it was possible that some chronic pain patients could have found an interview about their pain to be cathartic or even therapeutic, or in contrast, could have viewed it as distressing or anxiety-provoking [249]. Because the placebo effect is known to be dependent upon context and manipulated by affective cues and unconscious emotional associations, we did not want the interview or any emotions (positive or negative) that it may have provoked to influence potential future response in our participants. Therefore, we purposefully chose to conduct the interview at study completion to prevent or circumvent such interference. This unfortunately means that we lost the ability to use language as a parameter with which to predict placebo outcome. However, we were still able to use the interviews to investigate semantic differences between responders and non-responders and in turn see how well these findings correlated to neuroimaging and questionnaire data that are predictive of placebo response pre-treatment.

To control for potential environmental cues that may affect placebo response or the memory of it, participants were asked to return to the same place where they had completed all previous visits; here, all

interviews were conducted in a quiet, private room with the door closed. For consistency between participants, only one researcher (SEB) led all interview questions for all participants; however, as a safety precaution and to control for potential gender effects, a male colleague with whom the participant was familiar was also present in the room as an observer. Participants were instructed to only interact with the main interviewer until the interview was complete, and they were also asked to refrain from asking any questions about the study until after the interview was over.

The researcher explained that the general purpose of the interview was to better understand the participant's pain and health history, including how chronic back pain had impacted their overall mood, quality of life, and interactions with others. Participants were also told that they would be asked about their previous experiences with medical professionals due to their pain, as well as probed about their opinions and knowledge about various pain therapies. The researcher informed patients that the exit interview would be used to improve upon future studies and thus they would be asked about their time in the study, including their experience with the study medication if in a treatment arm. Finally, participants were told that their information would remain confidential, de-identified by their participant id, and not shared with anyone outside the study. Even though they agreed to these conditions when signing the consent form at visit 1, all participants provided verbal consent to have their voice recorded before the interview began. All interviews were audio-taped with an electronic hand-held device and upon recording commencement, were verbally stamped with the date, time, and PID number prior to beginning with questions.

To verify that the questions asked would be appropriate for chronic pain patients and would result in descriptive and meaningful answers, as well as test whether any of the questions' wording needed to be edited, the script was first piloted on 10 participants with knee osteoarthritis (OA) who had previously completed another clinical trial in the lab [250]. This cohort's answers indicated that that interview was able to provide language data of sufficient length and detail; very few changes were made to the initial wording. **Table 2** shows the final interview script. The interview included "warm up" section that asked questions purposefully not related to pain or the study, and a main "substantive" section that probed participants about their current and previous pain, mood, medical experiences, opinions on traditional and alternative medicines, and time in the study. The warm up section was used primarily to get participants

comfortable with talking to the research and tried to help them forget that the recording device was nearby. However, it was also used to collect linguistic information that would theoretically be less directly biased by concepts like pain or mood that may contribute to added sources of noise and variability in the data. Importantly, two of these questions (one describing a dream and the other a recent event) were re-interpreted from recent papers investigating predictive language properties of future psychotic episodes [165] and drug usage classification [166], respectively. In general, the interview was designed to be semi-structured but open-ended. This means that while there was a general set of questions that every participant was asked (11 primary questions), questions were purposefully designed to be broad in order to allow patients the full range of possible responses and interpretations. Moreover, the order of questions was not necessarily consistent between participants and was instead based on their most recent responses in order to keep the flow more natural and conversational. Thus the interview provided flexibility for both the researcher and the participant, with the goal being to get as much information as possible that was consistent across participants with as influence from the researcher as possible. In cases where participants were less verbose, or in the instances where their stories lacked information that we wanted to capture for all patients, the interviewer asked them additional follow-up questions or probes for clarification, being careful to minimize their impact on the participant's speech content. Excerpts from selected interviews have been provided in **Appendix I** as examples of some of the content of patients' narratives.

Interview Questions	Potential Follow Up Questions	Potential Probes	Reason for Inclusion
1. Tell me about yourself.	n/a	job, hobbies, interests, goals, family	interested in what comes to mind first outside of pain
2. What top 4 words would you use to describe yourself and why?	n/a	n/a	interested in how they see themselves in general, despite pain
3. Tell me a dream that you remember that you can describe vividly.	n/a	could be recent; could be from childhood; could be reoccurring	used to predict future episodes of psychosis so wanted to collect it for future comparison
4. Describe a recent (within the last year) event that you	n/a	could be a positive, a negative, or a neutral event;	because it involved friends or family members, it was

had with family or friends.			similar to previous studies that asked about people of importance, so wanted to collect it for future comparison
5. Tell me about your experience with your back pain?	1. How has your pain been recently? 2. How long have your had your pain? 3. What do you think may have caused it? 4. How has your pain changed over time from when you first noticed it to now? 5. What does your pain feel like? Describe it. 6. How has your pain changed over the course of the study? 7. Do you see yourself differently because of your pain? 8. What are some things that influence your pain? What makes it better or worse? How do you manage your pain? 9. What do you understand about your back? Do you think back pain is different for different people? 10. Do you have a family history of chronic pain? Please explain. 11. Do you know anyone who has/had chronic pain? Tell me more about them. How do/did they deal with their pain? Did this affect how you deal with your pain? 12. How do you think your pain will be a year from now? Why? 13. Have you ever woken up and thought that the pain was gone? Do you believe your pain might disappear one day? 14. How does your pain affect your mood? Does it ever make you feel angry/sad/anxious/frustrated? 15. How does mood affect your pain? 16. Does pain affect your concentration/decision-making/or memory? Tell me ore about that. 17. Tell me how pain affects your relationships with other people. Do you feel like people understand what you are going through and are supportive?	(8) increased or decreased activity; specific physical activity; sleep quality; body position; medications; faith/spirituality; weather or temperature; bad or good day; other distractions including: smoking or drinking, social activities, family/friends, or support systems (9) explain what you think is happening in your back to cause this pain (14-15) did good emotions alleviate pain or make it easier to ignore? Did negative emotions make it worse? (17) affects how you interact with others or how they interact with/treat you?	Wanted to get an extensive pain history to understand how pain had impacted their quality of life thus far
6. How many experiences have you had with doctors and/or the medical system because of your pain?	1. What have those experiences been like for you? Can you tell me about them or give an example? 2. How do doctors and medical staff react to your pain? 3. Have you had any difficulty getting health care for your pain? 4. Do you think there are medications that would help your pain but you haven't tried them or don't know about them or can't afford them?	(1) in- versus out-patient procedures, insurance, family, surgeries, injections, physical therapy, referrals, pharmacy/medication issues, payment; afraid of any kinds of treatments (pills, shots, operations, etc) (2) do they feel trusted/believed by doctors and do they themselves trust doctors/medicine? Have they searched for second opinions or different medications; do they feel people are attentive to their needs (3) did financial situation affect kind of help you have or haven't gotten? Insurance problems? Did your race/gender/sexual orientation affect kind of treatment they received?	Wanted to get an extensive medical history related to their pain; wanted to know how satisfied they were with the medical process/previous treatment (to understand potential expectations) and wanted to know how open they were to different kinds of treatments (and how much they thought they might work)
7. Has the treatment in the study helped your pain? Please explain.	1. Do you think that you received the active medication or the placebo as treatment? For both periods or just one? 2. What are your reasons for thinking this? 3. Would you recommend this treatment to someone else?	n/a	Wanted to compare what they thought happened with their ratings, their memory reports,

	4. How did this treatment compare to other treatments you've tried in the past? 5. What do you know about placebos? Do you think they can work for others?		and our definition of response – do they correspond?
8. Have you ever tried any unique treatments other than prescribed or over-the counter medication to treat your pain? Explain.	1. If so, what may you try these things? What did you try? What were those experiences like and would you try them again? 2. If not, why not? Are you interested in trying something else? 3. Have you ever tried to mentally control your pain? Did it work? Do you think this can work for other people?	(1) herbal therapies, essential oils or candles, acupuncture, yoga, massage, biofeedback, hypnosis, prayer or faith-based healing, support group meetings, cupping…	Wanted to assess whether mind-body therapies and alternative therapies work for them or they believe in them
9. Is there anything that you think I should know about your pain that I haven't already asked you?	n/a	n/a	Since interviewer did not have chronic pain, she could have missed a key question about the pain experience
10. Is there anything you think I should know about your experience with the study that we haven't discussed so far? How was your time in the study? Any improvements we could make?	n/a	n/a	Since context is important in placebo response, wanted to see if they had a good or bad experience with the study; could also gain information on expectations before/during the study
11. Is there anything else that you think I should know about your life –, any recent and/or important life events – that may have affected your time in the study, how much pain you have, or how you answered any of the questions?	n/a	n/a	Wanted to assess possible confounds in the data (placebo response, pain levels, and interviews themselves)

Table 2: Overview of exit interview script. The interview was conducted at the final visit of the clinical trial and was designed to be semi-structured and open-ended in all 3 sections (first column). This means that every person was asked the same broad interview questions (second column) and allowed to talk freely and openly about these questions for as long as they chose. In order to get approximately the same quality of data between participants with similar information over all, the interviewer was also allowed to follow-up with specific questions (third column) in case a person's narrative did not cover all of the topics. Sometimes, participants needed clarification or additional focus if the question was intentionally broad; in these instances, the interviewed provided specific probes to aid them in getting started (fourth column). Reasons for why these questions were asked are provided in the fifth column.

Interview Preprocessing **(Study 2)**

Transforming the interviews into data appropriate for quantitative analyses involved a multi-step procedure. After successful recording of the interview, the mp3 file was uploaded to a secure server, renamed with the PID and date of interview, and immediately deleted from the recording device so that no one else would have access to the information. Copies of these mp3 files were sent in batches of 10 to an outside company for medical transcription services (Lee Perfect Transcription, Inc., Chicago, IL). Interviews were transcribed verbatim, with all "um's", "ah's" and other fillers included; slang terms, colloquialisms, and improper syntax were also kept. Electronic transcripts in the form of word documents were sent back within 48 hours of their upload date. As part of the quality control process and to ensure transcription was done adequately, researchers manually checked every 5 transcripts for overall accuracy by tracking the document while listening to the corresponding mp3 file, correcting any mistakes found along the way. Although there were a few scattered spelling errors and a number of words that were inaudible due to the sound quality in the audio file, there were no systemic errors that warranted checking all 62 interviews. In the 12 transcripts that were reviewed, less than 5% of the data needed to be edited.

Next, the interviews went through a systematic cleaning process whereby words that carried little to no meaningful information were removed. Word documents were converted to plain text files so that they could be easily manipulated with text processors and toolboxes. First, all interviewer versus participant indicators (including PIDs and other identifiers) were deleted, along with timing brackets that specified long pauses or inaudible words. Next, all instances where the researcher spoke were deleted from the text; this included questions, clarifications, or any conversational material not associated with the main interview. To prevent confusion during parsing of the data, contractions were converted to the two corresponding base words from which they were formed (e.g, can't = can not). After this step, all punctuation was removed so that their characters would not be included in future token analyses. At this stage, the interview files consisted of only of words that the participant spoke, and a basic word count analysis was performed (explained in **Data Analyses**).

Despite substantial preprocessing at this point, the interviews were not yet ready for more complex analyses because they still contained substantial noise in terms of retaining a large amount of empty, non-meaningful words. To clean the text further, we used Python's NLTK (Natural Language Toolbox) to identify and subsequently remove all stop words (e.g., words like "the", "and", "a", and "it").

Following this, a script was used to tokenize the data (that is, parse the remainder of the interview into distinct word units known as tokens). The tokenized text was then lemmatized. Lemmatizing refers to reducing each word to its common base form, known as its lemma. For instance, the words "am", "are", and "is" would all become "be", and the words "cat", "cats", and "cat's" would all become "cat". Lemmatization is a crucial normalizing and de-noising step in most language analyses because it (a) helps reduce the number of inflectional forms a word can take, (b) makes all related forms of a word as similar as possible, and (c) essentially forces all words into a common template so that they can be compared with one another. At this late stage, the interview text is now unreadable but still retains its basic (and theoretically most important) meaning. Finally, after all of these steps were completed, all interviews were converted from text to number, with the number representing the alphabetized location (index) of the word in the dictionary. For instance, if a participant said the word "pain", it was converted to the number 50010 as it is the 50,010th entry in the dictionary. Thus, each interview was represented by a string of N numerical tokens (w_i) – (w_1,w_2,…,w_N) that were used as the primary inputs for the semantic programs, which were run in Matlab (MathWorks, version R2016a).

Follow-up phone call **(Study 3)**

After all neuroimaging analyses were completed, an additional analysis was conducted for participants whose data were used in the third study. A follow-up phone call interview was created to assess participants' episodic recall of study events during their time in the entire clinical trial, as well as test their general short-term memory (STM). **Table 3** provides the questions asked to participants during this phone call; in general, the interview was designed to be around 5 minutes in length (average: 4.5 ± 1.6 minutes) so that the STM question could be assessed appropriately. In addition to assessing episodic memory, different questions had different purposes. We asked participants about their memories of pain and mood during the first 2 weeks of the study so that we could compare long-term memory biases with previous short-term memory biases, as well as compare memories of pain and mood (which we did not capture during the study). We also asked participants to recall the number of days they were in the study, the number of visits they came to the lab, and the number of visits that involved scans – this was done to provide us with a relatively neutrally-valenced set of numbers that would be similar across participants. In

contrast, we also posed two questions about money – we asked participants to remember the total amount of compensation they received in the study and the worth of one rating from the pain application. Because pain is negative and associated with punishment, these questions allowed us to compare negatively-valenced memories with positively-valenced memories of reward. Finally, an STM prompt was used to capture general ability to recall items, and we used this to test whether memory biases seen in our participants could be dependent upon impairment or difficulty in short term memory. We copied the STM prompt from the Montreal Cognitive Assessment's (MoCA, [251]) delayed memory task. Briefly, as the first item in the phone interview, participants were told that they would be asked to recall five random words; these words were then listed and the participants were required to verbally repeat them. At the end of the phone call, after all other questions were answered, participants were asked to recite any of the words they could remember, with one point given for each word they recalled without a cue from the researcher. For the remainder of the follow-up questions described above, participants were asked to report what they remembered as honestly and accurately as possible and, in cases where they were unsure, to make a guess if they could. We did not pressure participants to provide an answer if they did not want to, resulting in missing data for a few individuals which we note in the **Results** section. Additionally, data from question 4 about number of days in the study was excluded because there was too much variability in responses (with large extreme values at both ends).

Questions asked:	
1	I'm going to ask you to remember 5 words. At the end of this phone call, I will ask you to recall as many of them as you can. The five words are: "face" "velvet" "church" "daisy" "red". Please repeat them now.
2	From 0 to 10, with 0 being no pain and 10 being the worst imaginable, what was your average pain during the first 2 weeks of the study?
3	From -10 to +10, with -10 being the worst mood imaginable and +10 being the best mood imaginable, what was your average mood during the first 2 weeks of the study?
4	How many days were you in the study?
5	How many visits did you have as a part of the study?
6	Of the visits you remember, how many of them included an MRI scan?
7	How much total compensation did you receive over the course of the study?
8	How much money was one phone/computer rating worth?
9	What is your current pain intensity right now form 0 to 10?
10	What is your current mood right now from -10 to +10?
11	What were the five words I asked you to remember? Please tell me as many as you can remember.

Table 3: **Follow-up phone call script**. List of questions asked to 33 CBP participants as part of a follow-up phone call. In addition to assessing short term memory with a delayed recall task, participant's memory of the study was assessed by comparing their responses on the phone to the actual values from the study and calculating a discrepancy (recalled-actual) for each item.

Monitoring pain intensity with phone app **(all studies)**

The pain of each patient was monitored electronically using an application (app) designed specifically for the study (**Figure 9**). This app was used to track the patients' pain and mood over time and to query them on their medication usage; it could be accessed using either a smartphone or a website link on a computer. The app had two VAS scales with sliding bars: the first asked participants to rate their current pain level from 0 (no pain) to 10 (worst imaginable), and the second asked them to rate their current mood level from -10 (saddest imaginable) to +10 (happiest imaginable) with 0 being neutral. The app also included fields to indicate the participant's assigned ID number, query if participants had taken any rescue medication at that time, and ask if they had taken the study medication, and there was a comments section that they could use to describe their pain, mood, or medication usage if they chose. Participants were instructed to use the app twice a day, once in the morning and once at night. To encourage compliance, participants were compensated $0.25 for each rating they submitted, up to $0.50/day. This additional payment was given to them on the last visit of the trial. Submitted ratings were immediately sent to a secure server and both date- and time-stamped. Rating compliance was assessed by a separate program monitoring whether the list of currently enrolled patients had provided the necessary ratings during the previous day. In the case that a patient omitted a rating, staff were alerted via an email. If patients missed more than 2 consecutive ratings (~24 hours-worth), a member of the study team contacted them to remind them to use the app. Two patients were discontinued from the study because they did not comply with the daily rating requirements despite repeated contact from the study team.

Figure 9: Electronic application. Participants were asked to rate their pain twice a day for the duration of the study using a smart phone application (app) or computer. Pictured here is a screenshot of the app; participants entered their assigned ID and then rated their pain on a scale from 0 to 10, with 0 being no pain and 10 being worst pain. All ratings were sent to a secure server and stamped with the date and time completed. The app also had a scale to rate mood from -10 to +10, questions inquiring about rescue medication usage and treatment compliance (if applicable), and a comments section (not shown). Participants were paid $0.25 for each rating completed up to 2/day.

To verify that pain levels remained within the inclusion criteria specified above, all participants' ratings were closely monitored for the first two weeks of the study as part of a run-in/baseline pain period. Individuals not meeting this level were deemed ineligible and did not continue in the study (n=16 screen failures). It was later noticed that 3 additional participants in **Study 1** had met this exclusion criteria but accidentally continued in the study. One person was assigned to no-treatment and was discontinued as a protocol deviation before study completion; the other two individuals finished the study in the placebo treatment group and were later excluded from the analysis to make sure the results were not dependent upon their outlying pain levels at study commencement.

Stratification of patients into placebo Responders and NonResponders **(Studies 1** and **2)**

Rating data from all participants were downloaded from the server as text files and preprocessed as follows. Although participants were asked to rate two times a day (and only compensated for this

amount), many participants exceeded this number of app ratings in 24 hours due to over-compliance, reassessment of their pain or mood level, and/or cellular service problems. If pain ratings were entered within 30 minutes of each other, only the last rating was kept and taken as indicative of the participant's final assessment of their pain and mood levels at that time. Any additional ratings outside of this 30 minutes window were not considered duplicates and were kept as valid entries. Beside this cleaning process, no other changes were made to the ratings. In the instances where participants missed ratings, no attempts were made to interpolate or re-sample the data so that the temporal aspects of the ratings were left intact.

This smartphone technology permitted us to track fluctuation in pain levels throughout the study. To best make use of the daily rating data, we developed a new classification scheme of responders versus non-responders that is currently not used in the literature. It was important that this method accounted for the within-subject variability of pain levels in their day-to-day experience (something that neither a percent change in average pain levels nor a change in pain calculated between 2 scores at different visits provides). Each patient was classified based on a permutation test between the pain ratings acquired during the baseline rating period (Visit 1 to Visit 2) and the pain ratings acquired during the treatment periods (either baseline versus treatment 1 or baseline versus treatment 2). The null hypothesis was generated by randomly re-sampling 10,000 times the distribution of pain ratings, which provides all possible t-values obtained from the rearrangement of the pain ratings. The real t-value obtained between baseline and treatment was used to determine if the null hypothesis could be rejected ($p<0.05$) for each of the treatment periods. In the cases where the null hypothesis could not be rejected for either of the treatment periods, the patient would be stratified as a "Non-Responder". Alternatively, the patient would be stratified as a "Responder" if the treatment induced significant diminution in the pain ratings. The main advantages of using a permutation test is that it (a) takes into consideration the variability across pain ratings during the baseline and treatment periods and (b) represents a statistically-defined cut off point for response and thus is more likely to be reproducible in another study (unlike cut off points arbitrarily defined by a percentage change in pain).

Comparison of phone app with other pain outcomes **(Study 1)**

The pain outcome measured with the phone app was compared with six additional outcome measures of pain level. The numeric rating scale (NRS) and the memory of pain were the two other pain outcomes also relying on numerical scales. The NRS represents the traditional standard pain measurement usually used in clinical trials assessing pain levels of participants for both placebo-controlled trials (compared against an active medication) and placebo-only trials (where the placebo effect is being manipulated) [127]. The memory of pain represents the one of the standard pain assessments used by physicians in clinical practice and has been shown to correlate well with daily pain diaries in previous studies [252]. Other pain outcomes were collected using the McGill pain questionnaire (MPQ) affective and sensory scales, the pain detect, and the neuropathic pain scale (NPS), which have been widely used in both randomized clinical trials and research labs, although their utilization in placebo-only trials remains minimal [253].

Using the phone app to calculate psychometric properties of interest **(Study 3)**

After ratings were cleaned as described above, metrics for pain and mood data – including average, standard error, peak, minimum, end, and area under the curve – were all computed in Matlab (version R2016a). These metrics were chosen based on previous literature [201, 203, 204]. While the authors of these papers also used “initial pain” as a parameter, due to our participants having chronic pain for months to years, we did not include this metric as there was no way to capture the onset of their pain (although we did investigate the effects of pain duration on these measurements as part of our analysis).

Blinding of the analyses **(Studies 1** and **2)**

Given the recent issues regarding lack of reproducibility in scientific findings [254] and the importance of transparency in data analysis, we followed recommendations by MacCoun and Pearlmutter [255] and employed cell scrambling to further blind our data and minimize bias. For all endpoints, a lab member not involved in analyses was selected to organize data files and spreadsheets for processing and statistical analyses of the data. This person first renamed all the data files to ensure that analysts were blinded to each participant’s unique id to minimize bias due to previous interactions with patients

during data collection. Next, all analyses were performed with 3 randomized codes (which we refer to here as “classifiers”) for each condition, with only one of them being the proper classification of responders, non-responders, and no treatment. We refer to this as “triple blinding” because analyzers were blind to participant id, participant treatment, and correct participant group classification. The selected lab member did this blinding prior to any analyses, with the exception of the pain ratings from the app, which was used to stratify patients first. This resulted in each analysis being done three different times in an unbiased manner. Importantly, the three lab members that contributed to the analyses were not informed that they were provided different classifiers to make sure they could not collaborate to figure out which one was the real code. The results were presented in a public lab meeting where the lab member un-blinded the analyzers to the data to confirm which results were true. Although we refer to these 3 classifiers throughout the paper, we only present the outcomes and data from the correctly classified group in each instance. Results from the 2 false classifiers will be presented where applicable in the forthcoming manuscript supplementary materials for the purpose of comparison (outside of GMD, they are not shown in this thesis). This triple blinding procedure aimed at decreasing uncontrolled bias during data analyses and enhancing the reproducibility of results.

fMRI and anatomical scanning protocol **(all studies)**

Participants were scanned on a Siemens Magnetom Prisma 3.0 T whole-body system at Northwestern University’s Center for Translational Imaging (CTI) with capacities matching the human connectome project parameters. All scans were acquired with a 64-channel head coil. The procedure consisted of high-resolution T1-weighted brain images, two consecutive 10 minutes resting state functional scans, and a multi-shell diffusion imaging scan (data not shown). The entire procedure was completed in about 35 minutes, but an extra 25 minutes was allocated to install the patients in a comfortable position to keep their back pain at a minimum and to re-acquire images if the data was contaminated by head motion.

High-resolution T1-weighted brain images were collected using integrated parallel imaging techniques (PAT; GRAPPA) representing receiver coil-based data acceleration methods. The acquisition parameters were: isometric voxel size = 1 X 1 X 1 mm, TR = 2300 ms, TE =2.40 ms, flip angle = 9º,

acceleration factor of 2, base resolution 256, slices = 176, and field of view (FoV)= 256 mm. The encoding directions were from anterior to posterior, and the time of acquisition was 5 min 21sec.

Blood oxygen level-dependent (BOLD) contrast-sensitive T2*-weighted multiband accelerated echo-planar-images were acquired for resting state fMRI scans. Multiband slice acceleration imaging acquires multiple slices simultaneously permitting denser temporal sampling of fluctuations that improves the sensitivity of detection of signal fluctuation by up to 60% [256]. The acquisition parameters were: TR = 555 ms, TE = 22.00 ms, flip angle = 47º, base resolution = 104, 64 slices with a multiband acceleration factor of 8 (8 X 8 simultaneously-acquired slices) with interleaved ordering. High spatial resolution was obtained using isomorphic voxels of 2 X 2 X 2 mm, and signal to noise ratio was optimized by setting the field of view (FoV) to 208mm. Phase encoding direction was from posterior to anterior. The time of acquisition lasted 10 min 24 sec, during which 1110 volumes were collected. Patients were instructed to keep their eyes open (to reduce alpha-band interference and lessen likelihood of sleeping) and to remain as still as possible during acquisition. The procedure was repeated two times in the event that the first resting scan could not be used because of artifacts such as excessive head motion or patients falling asleep. For **Studies 1** and **2**, only the first resting state scan was used for all patients. For **Study 3,** only the T1 image was used.

Additionally, **Study 3** relied on another dataset collected from an earlier study. SBPp and CON participants were also scanned on a Siemens 3T Prisma whole-body system with slightly different parameters. High-resolution T1-weighted brain images with a 12-channel head coil using the following acquisition parameters: isometric voxel size = 1 X 1 X 1 mm, TR = 2.5 ms, TE = 3.36 ms, flip angle = 9º, in-plane matrix resolution – 256 X 256; slices = 160, and FoV = 256 mm.

fMRI data preprocessing (**Studies 1** and **2**)

Several steps were necessary for de-noising the original time series of fMRI volumes. The pre-processing was performed using FMRIB Software Library (FSL) and in-house software. The first 120 volumes of each functional dataset were removed for allowing magnetic field stabilization leaving a total of 990 volumes used for functional connectivity analyses. The effect of intermediate to large motion was initially removed using fsl_motion_outliers. Time series of BOLD signal were filtered with a Butterworth

band-pass filter (0.008Hz<f<0.1Hz) and a non-linear spatial filter (using SUSAN tool from FSL; FWHM=5mm). Following this, we regressed the six parameters obtained by rigid body correction of head motion, global signal averaged over all voxels of the brain, white matter signal averaged over all voxels of eroded white matter region, and ventricular signal averaged over all voxels of eroded ventricle region. These nine vectors were filtered with the Butterworth low band-pass filter before being regressed from the time series. Finally, noise reduction was completed with Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC tool in FSL) that identified components in the time series that were most likely not representing neuronal activity. This ICA regression process was kept very conservative so that only components obviously related to motion or noise were removed. Finally, all analyses were performed with and without volume censoring to ensure that head motion had no influence on our functional results. Censoring (also known as scrubbing) was performed in 3 different ways according to Power et al. guidelines [257].

The functional image registration was optimized according to a two-step procedure. All volumes of the functional images were averaged within each patient to generate a contrast image representative of the 990 volumes. This image was then linearly registered to the MNI template and averaged across patients to generate a common template specific to our CBP patients. Finally, all pre-processed functional images were non-linearly registered to this common template using FNIRT tool from FSL. The registered brains were visually inspected to ensure optimal registration.

Data analyses

The analyses are presented in order of study, although there is some overlap between studies (which is noted where applicable). Unless otherwise specified, all MRI data were analyzed using FMRIB's Software Library (FSL) version 5.0.8.

***Brain networks constructed from resting state fMRI* (Study 1)**

The brain was divided into 264 spherical ROIs (5-mm radius) located at coordinates showing reliable activity across a set of tasks and from the center of gravity of cortical patches constructed from resting state functional connectivity [258]. These ROIs were merged to generate a maximally-spanning

collection of ROIs according to Power et al ([258]). Because limbic structures are believed to play a role in placebo response, a 5-mm radius was manually added in bilateral amygdala, anterior hippocampus, posterior hippocampus, and NAc. Linear Pearson correlations were performed on time courses extracted and averaged within each brain parcel. Given a collection of 272 parcels, time courses are extracted to calculate a 272x272 correlation matrix. These matrices allowed for the construction of weighted brain networks, where nodes represent brain regions and links represent weighted connections from Pearson correlations between any given set of these regions.

Community detection analyses **(Study 1**)

We used the Louvain algorithm integrated in the Brain Connectivity Toolbox (BCT; https://sites.google.com/a/brain-connectivity-toolbox.net/bct/[259]) to determine consistent community structures across a large number of network partitions [260]. For each subject, the individual community structure was initially constructed from 100 repetitions of the same network. The group community was then constructed from 100 x 63 patients generating a total of 6300 networks. The final community structure was created by thresholding the averaged within-module connectivity likelihood matrix at 0.5, meaning that if the likelihood for two nodes belonging to the same module was above 50% they were considered in the same module. This permitted us to identify six separate communities including the four communities of interest that we ended up studying (see below).

Identifying communities of interest **(Study 1**)

We used a localizer from an independent data set consisting of osteoarthritis patients studied during a clinical trial determining the impact of Duloxetine and placebo on the brain[250]. We used resting state functional connectivity to identify four regions predicting patients in the placebo arm that responded to treatment: the right mid-frontal gyrus connectivity (x=28, y=52, z=9), the anterior cingulate cortex (x=-3, y=40, z=2), the posterior cingulate cortex (x=-1, y=-45, z=15), and the right somatosensory cortex (x=60, y=-7, z=21). We next entered these coordinates as seeds in the Neurosynth analytic tool (http://neurosynth.org) and extracted three networks sharing strong connectivity with these seeds: the DMN, the frontoparietal network, and the sensorimotor network. We identified communities corresponding

to these networks based on spatial overlap, by multiplying the networks of interest with the nodes pertaining to each community. Communities overlapping with networks of interest were the DMN, the lateral frontal, the parietal, and the sensorimotor communities. A total of 113 nodes were affiliated with these communities. The 151 nodes affiliated with the visual and saliency communities and those nodes without affiliation to any community were excluded from the analyses. The limbic nodes and a node located in the PAG from the Power parcellation scheme (which was not affiliated with any community) were added for a total of 122 nodes of interest.

Network statistics **(Study 1**)

Group comparisons (responders, non-responders, and no treatment) were initially performed on the average r-values between nodes within a same community or nodes between different communities using one-way ANOVA. Group differences between placebo responders and non-responders were further examined using a permutation test (5,000 permutations) on the 14,762 connections of the weighted network (122 * 121 nodes) controlling for false discovery rate (FDR $p < 0.05$) using the Network Based Statistics toolbox (NBS; [261]). The r-values of the significant connections were extracted and entered in one-way ANOVA (between group: responders, non-responders, no treatment) with Bonferroni corrected post-hoc tests to determine differences with respect to the no-treatment group.

Grey matter density **(Study 1**)

Grey matter density was examined using voxel-based morphometry from FSLVBM . All T1-weighted images were first brain extracted and then segmented into grey matter, white matter, or cerebrospinal fluid. A common grey matter template was generated for CBP by registering and averaging all grey matter images. The grey matter image of each participant was then registered to the common template using non-linear transformation. A voxel-wise permutation test was used to test the significance of group differences between placebo responders and non-responders to a distribution generated from 5000 permutations of the data for each voxel of the template using a sigma filter of 3 mm for a smoothing. The initial analysis established significance level using the Threshold-Free Cluster Enhancement (TFCE) method (FWE $p < 0.05$) and a follow-up analyses compared the average grey matter density in nodes of

the communities showing group differences in functional connectivity (between group: responders, non-responders, no treatment; Bonferroni corrected for 50 nodes; $p < 0.001$).

Cortical thickness **(Study 1)**

Cortical thickness was examined using Freesurfer software library (http://surfer.nmr.mgh.harvard.edu/). In brief, the structural processing includes skull stripping, intensity normalization, Taliarch registration, segmentation of subcortex, reconstruction of cortical surface, and tessellation of the gray/white matter boundary and pial surface. Following reconstruction of the cortical surface, brains were inflated, averaged across participants to produce a study-specific brain, and then smoothed using a 15 mm full-width at half maximum Gaussian kernel. A direct measure of cortical thickness was calculated using the shortest distance (mm) between the pial surface and gray-white matter boundary at each point or vertex. Cortical thickness analysis for each hemisphere was conducted using FreeSurfer's Query, Design, Estimate, Contrast (QDEC) graphical interface. The initial vertex-wise comparison was performed between placebo responders and nonresponders for each hemisphere. Correction for multiple comparisons was performed using random-field-theory-based significant clusters at $p < 0.05$. Values of cortical thickness were extracted in the significant cluster surviving multiple comparison and compared between responders, non responders, and no treatment groups in SPSS.

Subcortical volumes **(Study 1)**

Volumetric analyses of T1-weighted images were performed through automated processes using both FSL (version 5.0.8) and FreeSurfer (version 6) softwares for **Study 1**. For this first study, we investigated volume differences in 3 subcortical nuclei selected apriori; the NAc, the amygdala, and the hippocampus. After using FSL's brain extraction tool (BET) to remove the skull from all images, FSL's integrate registration and segmentation tool (FIRST) was utilized to segment these specific subcortical regions and extract their volume measurements [262]. Unilateral volume measurements for each region were initially compared between responders and non-responders. Given the recent evidence from the ENIGMA consortium showing that subcortical volume asymmetry can provide a brain signature for psychopathologies [263], we also investigated the possibility that asymmetry differences may provide a

biomarker for placebo propensity in our data. All subcortical regions' volumes were summed for the right and the left hemisphere separately; for each patient, the ratio between the two (right/left) was created, where a result =1 would be indicative of perfect subcortical symmetry, whereas numbers >1 or <1 would indicate asymmetry biased toward the right or left hemispheres, respectively. Volumes and subcortical assymetry were compared between responders, non-responders, and no treatment groups in SPSS. Volumes of the right and left hippocampus were also analyzed separately as part of **Study 3** looking at memory bias of pain; these volumes were simply correlated to the participants' discrepancy scores.

Analysis of questionnaire data **(Study 1*,*** although all studies used overlapping questionnaire data)

Over the course of the 6 visits, participants filled out 29 unique questionnaires. These specific self-report measures were chosen for one of 4 reasons: (1) to gather basic information about participants, including demographics and pain/medical history, (2) to track any changes in the quality and/or intensity of pain characteristics as measures of treatment efficacy, (3) to monitor any changes in emotional affect which may have influenced someone's time in the study or their treatment response, and (4) to capture trait-based qualities, general habits and beliefs, or state-related expectations of individuals that may predispose them to respond to a placebo. Questionnaires used to track pain and mood changes overtime were repeated across all study visits. Questionnaires that targeted expectations towards treatment and satisfaction after treatment were conducted at two visits each– either before treatment sessions (visits 2 and 4) or after treatment periods (visits 3 and 5), respectively. In contrast, measures that aimed to identify more stable traits of participants were completed at visit 1, which allowed us to use them as possible predictors of response. Finally, a subset of questionnaires regarding beliefs toward alternative medicines and suggestibility were administered at the final visit after the exit interview. A full list of all questionnaires used, along with descriptions and references, can be found in **Table 1** above. The data analyzed for **Study 1**, with the exception of the pain questionnaires collected at every visits to determine treatment outcome, come from those questionnaires collected at visit 1 only, as we were interested in looking at predictors of placebo response. **Study 2** relied primarily on questionnaires collected at visit 1 and visit 6 (time of interview), and **Study 3** relied on those collected at both visit 1 and visit 2.

Data from these self-report measures were downloaded directly from REDCap as a CSV file and scored in Excel according to their references. Because all questionnaires were converted to an electronic format in order to be used in REDCap, an option to "skip" a question was provided if the participant did not feel comfortable answering a certain item. If more than 30% of the data from a given questionnaire (or questionnaire subscale, if applicable) was missing, the person's data for the questionnaire was not scored; for all other missing data, the mean was used to fill in missing items (if the questionnaire had sub-scoring, the mean was calculated from the remaining items in the sub-dimension as opposed to the entire questionnaire); this approach is one of the most commonly used methods in data analysis [264] and was utilized in order to conserve statistical power given our relatively small sample size. Of all the self-report data analyzed, less than 3% was totally missing and thus unable to be filled in as described above.

Two methods were used to analyze questionnaire outcomes: (1) a data-driven approach with factor analysis using all questionnaire data as inputs, and (2) a hypothesis-driven approach using only measures selected apriori as inputs. For the first approach, a hierarchical factor analysis was utilized to identify a minimum set of self-report outcomes as predictors of placebo response as defined above. In addition to the individuals used in the neuroimagining analysis (n=65 before the exclusion of 2 individuals having their pain too low), we also included 46 additional participants who had completed all questionnaires at Visit 1 but later screen-failed or discontinued (total n = 111). These additional participants were used to create more robust and better-differentiated factors.

Briefly, all visit 1 questionnaires were entered into a principle component analysis (PCA) using STATA (Student Edition, version 11.0). The resulting factors were visualized using a scree plot to see how many components had eigenvalues >1; this initial plot suggested as many as 5 factors above an elbow (a point of clear separation in the fraction of variance explained [265]. To statistically choose how many of these potential 5 factors to keep in the model, a parallel analysis was run in Mplus using 10,000 random permutations at the 99th-percentile – this provided 4 statistically relevant factors, so the 5th potential factor was eliminated. The PCA was then constrained to a 4 factor matrix, which was orthogonally rotated using a VARIMAX function to make the components more independent. Factor loadings were thresholded, discarding any questionnaires that did not have loading values of ± 0.4. For a simple approach, instead of using the raw loading values, each of the elements in the four factors were

un-weighted by multiplying the questionnaire items by 0, 1 or -1. For example, if an item had a loading value of 0.72, it was assigned a "1", and if it had a value of -0.55, it was assigned a "-1"; those items not passing the 0.4 cut-off threshold were assigned a "0". These unweighted factors were then used as covariates in a logistic regression predicting the probability of response. Coding for the binary independent variable was made so that 0 = nonresponder and 1 = responder.

In the second approach, 15 questionnaires of interest were grouped into 4 broad themes that were hypothesized to be important to the placebo response: attitudes and emotions toward pain, positive affect, negative affect, and emotional/attentional regulation and awareness. The themes and their assigned questionnaires can be found in **Table 4.** Importantly, the selection of these items and their organization into groups were chosen prior to triple blinding and subsequent analysis. Themes were treated as if they were factors, running all questionnaire items within each theme into 4 separate backwards-stepwise logistic regressions predicting probability of response. Here again, the independent binary variable was coded so that 1 = responder and 0 = non-responder. Only surviving items from each theme's regression were kept ($p<0.05$). After all regressions were run, all surviving items were used in the final full model for the questionnaire data.

Theme 1: Attitudes and Emotions Toward Pain
CPAQ (2 subscales) ○ *Activity engagement* ○ *Pain willingness* **PCS (3 subscales)** ○ *Helplessness* ○ *Rumination* ○ *Magnification* **PASS (4 subscales)** ○ *Avoidance behavior* ○ *Cognitive anxiety* ○ *Physiological anxiety* ○ *Fear* **MPQ** (*affective subscale*)
Theme 2: Positive Affect/Personality
PANAS (*positive subscale*) **Extraversion** **LOT-R**
Theme 3: Negative Affect/Personality
PANAS (*negative subscale*) **Neuroticism** **BDI** **LAQ**

Theme 4: Emotional and Attentional Regulation and Awareness
ACS
eACS
ERQ (2 subscales)
○ *Reappraisal*
○ *Suppression*
MAIA (8 subscales)
○ *Noticing*
○ *Not-Distracting*
○ *Not-Worrying*
○ *Attention Regulation*
○ *Emotional Awareness*
○ *Self-Regulation*
○ *Body Listening*
○ *Trusting*

Table 4: Hypothesis-driven analysis of questionnaires. The self-report measures were secondly organized into 4 themes that were hypothesized to be important to placebo response in CBP. The questionnaires or subscales were initially chosen from the literature according to these themes, which served as general guidelines of personality traits of interest. All themes and groupings were done prior to analysis and triple blinding. We ran items from each theme in 4 separate logistic regressions (for each of the 3 classifiers) and surviving variables ($p<0.05$) from any of the regressions were kept for our final questionnaire model.

Analysis of final placebo propensity model* (Study 1*)

This final step in the analysis aimed at addressing two questions: can we map the personality traits associated with placebo response with brain functions associated with this response, and if not, do the brain properties explain variance in placebo response that is unique, providing information beyond what is captured by personality traits? After all neuroimaging and questionnaire analyses were completed, the correct classification scheme was revealed. The final model was therefore tested without remaining blind. We initially correlated all significant brain parameters with all significant questionnaire parameters to get a sense of their co-dependence. Next, we tested the combination of neuroimaging markers with questionnaire variables in predicting response. To avoid over-fitting our data due to the number of variables and our relatively small sample size, all significant imaging findings were first entered into 4 separate stepwise, backward logistic regressions with each of the 5 surviving questionnaire items; those brain parameters that survived each regression and remained significant ($p<0.05$) in the model were kept for the final multivariate model. As an additional step to further reduce our potential variables of

interest, only those self-report measures that survived in all of these partial regressions (i.e., only those that were consistent in all 4 models) were then entered into our final multivariate model.

Finally, we trained 3 different types of classifiers to discriminate between responders and nonresponders using the surviving variables from our final multivariate logistic regression using a leave one subject out cross-validation procedure (LOOCV); these included a logistic regression based on Mahalanobis distance, support vector machine (SVM), and a Naïve Bayesian classifier. For each classifier, average accuracy in group classification was calculated over 1000 iterations of the data.

Word Count Analysis **(Study 2)**

After cleaning the interviews (but prior to full lemmatization, stemming, and stop-word removal), we calculated 3 linguistic metrics for every participant: verbosity (number of words spoken), vocabulary (number of unique words spoken, and lexical diversity (the ratio of vocabulary-over-verbosity, used as the simplest version of a lexical entropy measurement) using a Python script developed with NLTK methods. The latter measure was adjusted using a corrected type-to-token ratio (CTTR) that takes the square root of twice the total verbosity in order to account for the length of the text (this is also known as Carol's TTR). Without this correction factor, lexical diversity would not be able to discriminate between interviews of different lengths and would have a skewed negative trajectory due to language's inherently repetitive nature and excessive use of articles and prepositions [266]. These calculations were used as control measures to verify that they did not differ between groups and thus could not account for any differences seen in the semantic results.

Latent Semantic Analysis **(Study 2)**

As described in [166], the meaning of a text can only be understood with regard to the "mutual dependencies of words within a language", and because of this, simple linguistic analyses such as word counts or parts of speech tagging were not sufficient in capturing semantic content. Instead, we employed Latent Semantic Analysis (LSA, [267]) to calculate the semantic distance between the concepts of interest and the words used in the interviews. LSA is a high-dimensional associative model that is used to find hidden meanings or concepts by quantifying the associations between patterns of words

found in large set of documents (such as dictionaries or encyclopedias). Therefore, LSA operates under the assumption that semantically-related words will be present across texts with coherent topics or themes, and it generates a multidimensional linear representation of the semantic meaning of words based on their co-occurrence with other words in a large and diverse text corpus. The frequency of word co-occurrence across different documents in this collection represents how well given words are semantically related to each other[183].

LSA starts with a word-by-document (N X M) occurrence matrix X, with each row corresponding to a unique word in the corpus (N total words) and each column corresponding to a unique document (M total documents). Following recommendations by Bedi and colleagues [166], here we use the Touchstone Applied Science Associates (TASA) collection of documents as our corpus to generate a co-occurrence matrix. TASA consists of thousands of text documents spanning novels and newspapers that represent a common knowledge base across the US educational system through the end of high school. Our matrix X was then 77,998 distinct words from TASA (N) by 37,651 text materials in TASA (M), where any given (Ni, Mi) combination represented the number of times that Word (Ni) occurred in Document (Mi).

Next, we performed a dimensionality reduction to condense X into a matrix that had less columns while still retaining as much of the similarity structure between rows (words) as possible. This step allows for easier interpretation of the data by projecting each work into a “semantic space” where meaning is represented by its corresponding vector within that space. This reduction was accomplished through singular value decomposition (SVD), where our co-occurrence matrix was mathematically broken down into the product of 3 matrices (U,S, and V) and then cropped to k dimensions. The value of k is important because it dictates the accuracy of the results – too few dimensions means that important patterns will likely be left out, but too many dimensions can reintroduce noise from things like random word choice or over-usage. Previous research on LSA dimension size indicates that the number of dimensions typically selected is between 100 and 500, and (Landauer and Dumais’s seminal paper [183] showed that they obtained the best performance when truncating at around 300 dimensions, which has become a sort of field standard in this approach). Therefore, we used k =300 as our cutoff for the number of columns to retain post SVD. Thus X was converted to the combination of U (77,998 words X 300 semantic features), S (square matrix of 300 X 300 semantic features with diagonal elements representing the singular values

of each of the 300 factors), and V^T (transposed matrix of 300 semantic features X 300 projections to these features based off the reduced TASA corpus data). For our purposes of looking at semantic relationships between words, we only utilized the matrix U for our analyses, which we refer to here as our "dictionary" and which represents a general English lexicon weighted by the frequency of words. For each word spoken in each participant's interview, the corresponding dictionary word was found and its vector of 300 semantic features was copied into a new matrix for analysis, resulting in an N words X 300 feature matrix for each subject representing the semantic space of each interview. Individual subject matrices were appended together according to participant ID such that the first 800 words belonged to PL001's interview, the next 700 to PL002's interview, the next 1000 to PL003's, and so forth through all 76,929 total words used by all participants; this data reorganization allowed for more efficient calculations while still maintaining the order of words within each interview. We performed SVD again on this participant-driven dictionary by feature matrix (76,929 X 300, referred to as Xi) to obtain semantic relationships that not only encompassed the dictionary lexicon but also were specific to our population. We refer to the resulting matrices as Ui, Si, and Vi (with i representing interview). Because singular values are organized in descending order, the first feature is the strongest in the dataset, followed by the second and third. We therefore took the first 9 out of 300 features from Ui representing the strongest linguistic elements in the dataset and compared them between groups; the choice of 9 was arbitrary (and could have easily have been higher or lower).

Semantic proximity analysis **(Study 2)**

Finally, we calculated the semantic proximity for all words in each interview to a selection of 60 that served as concepts of interest based on 8 themes. These themes were hypothesis-driven and were chosen because of their potential to capture elements of the chronic pain experience and qualities of the placebo effect; however, the themes were used for organizational purposes only, as they were arbitrarily defined and not determined via linguistic analyses. The 60 words within the themes were selected from a variety of sources spanning ethnographies, questionnaires, and peer-reviewed journals; positive, neutral, and negatively valenced words were included so as to encompass a larger range of potential semantic space and meaning. Semantic proximity is a measurement of how similar the meaning is between two

words, and it's calculated by taking the cosine between the two corresponding word vectors. In our dataset, the semantic proximity between two words is computed as the dot product of all 300 of the normalized semantic features for Word A and Word B obtained through SVD; the semantic vectors were normalized to constrain the possible outcomes to be between -1 and 1. The proximity to each of the 60 concepts was measured for all words in each participant's interview; from these, a mean was calculated for each participant across all words used for each of the concepts of interest to obtain an average semantic similarity value per person for the 60 concepts. This semantic similarity value essentially demonstrated how close each concept was to a person's unique language in the interview. The semantic proximity values for each concept were averaged within group (responders, non-responders, and no treatment) and used for future statistical analyses.

Creation of semantic language factors **(Study 2)**

The 60 concepts of interest were not expected to be completely independent of one another. Therefore, it did not make sense for us to search for 60 different language parameters in our questionnaire and neuroimaging data. Instead, we supposed that a combination of these concepts would be more important in explaining the variance in placebo response than would any one singularly. However, we chose not to test all possible combinations of 60 semantic proximity values and alternatively implemented a data-reduction approach that combined hypothesis-driven methods with data-driven analyses to further narrow down the number of language parameters used in our analyses and to see which combinations of word similarities were important.

The 60 words of interest were entered into a principle component analysis (PCA) using STATA (Student Edition, version 11.0). The resulting factors were visualized with a scree plot to see how many components had eigenvalues >1; this suggested as many as 6 components above an elbow (a point of clear separation in the fraction of variance explained [265]. The PCA was then constrained to this 6-component matrix, which was rotated using a VARIMAX function to make the underlying component structures more evident. The choice of an orthogonal rotation instead of an oblique rotation was made because even though the concept words and similarity scores were probably somewhat interdependent (as language in general is), the themes of interest were not necessarily related to one another, and the 60

words would not all be expected to occur in the same context, thus orthogonalizing them would allow for further separation from one another. Component loadings were thresholded, discarding any concepts that did not have loading values of ± 0.4 in each factor.

Those word similarities that survived this threshold were used as covariates in 6 backward stepwise logistic regressions (one regression per component) predicting the probability of being a being a placebo responder. Coding for the binary dependent variable was made so that 0 = non-responder and 1 = responder. At each step, the word with the highest p-value was removed from the model until it reached a total significance level of $p<0.05$; there was no inclusion/exclusion criteria for individual regressors. Resulting logistic models indicated the combination of semantic concepts that explained a statistically significant amount of variance in characterizing placebo responders. Those components (which we refer to as factors) that produced significant models were named according to the words in them, and their equations were used to create combination scores of all associated words in a given factor. These factor scores were used as regressors in future neuroimaging analyses.

Additionally, a 6^{th} and final factor was created post-hoc as a way to increase the semantic signal for future neuroimaging analyses. All 60 concepts were compared between responders and non-responders via un-paired t-tests to detect words that significantly differed in their semantic proximity between groups. These significant words were entered into two logistic regressions predicting placebo propensity. The first model was like the others and did not have a statistical threshold; it was used to capture the most variance in the data (i.e., like the first factors, the total model was significant but some of the individual words were not); the second model was a backward stepwise logistic regression of all significant words (with inclusion level of $p<0.05$) predicting binary placebo response, resulting not only in a significant model but also in individual significant predictors. The reason why we chose to set a statistical cut-point for only the 6^{th} factor is because the other 5 factors would have provided redundant information, being driven largely by the significant words in their models. We used this final model to test proof of concept – that is, the predictive accuracy of semantic language parameters in explaining placebo response.

Note: to avoid confusion when talking about the results of SVD and PCA (which are very similar methods), we have made the language consistent throughout the paper. Whenever we are referring to

SVD, we use the term *features*; for PCA, we use the term *components*, and for semantic similarity calculations that relied on PCA initially but have gone through additional data reduction methods (e.g., logistic regressions or statistical cut-off), we use the term *factors*..

Data-driven linguistic analyses **(Study 2)**

Pre-selecting words and concepts may result in unintentionally biased results. To avoid this, as an alternative method we also performed a data-driven analysis of semantic similarity (similar to the supplementary analysis completed in [166]), which we viewed as a "black-box" approach in that we did not pre-specify any words apriori and did not apply additional data reduction steps after SVD. Instead, we calculated the semantic similarity across all words in each interview to all the words in the dictionary (U). Thus each interview was represented by a 77,998-dimensional vector whose components were the average semantic similarity over an interview to each word in the lexicon, weighted by the frequency of occurrence of the word in the TASA database. This resulted in a matrix M for 42 subjects and 2 conditions (responder and non-responder) so that the dimensions of M = 77,998 X 42. SVD was performed on M to achieve the following matrices: Um = 77,998 X 77,998, Sm = 77998 X 42, and Vm = 42 X 42; for this analysis, Um can be seen as a set of multidimensional features and Vm as projections on those features. Because the data was not centered, the first feature essentially captures the mean, and therefore we only compared the second feature (the next strongest) between the two groups. To assess differences between responders and non-responders, we entered the loading of each interview on the second feature into an unpaired t-test.

Correlating language to the brain **(Study 2)**

In order to see if language was related to placebo propensity, the language factors were first correlated to variables that we had previously shown to be significantly predictive of placebo propensity in CBP (**Study 1**): these included 2 questionnaire measure (the Multidimensional Assessment of Interoceptive Awareness, MAIA, emotional subscale; and the Emotional Regulation Questionnaire, ERQ, suppression subscale), 2 functional links (lateral frontal to sensorimotor connectivity and lateral frontal to periaqueductal gray, PAG, connectivity), and 3 anatomical markers (limbic volume asymmetry, post-

central gyrus gray matter density, and superior frontal gyrus cortical thickness). We also correlated language factors to additional questionnaires and subscales of interest based on additional models in **Study 1** and on the idea that language should be able to capture aspects of emotion and personality that might not be as readily seen in imaging data.

Next, the language factors were used to investigate if there were any additional pre-treatment brain and personality measurements of placebo response that they were related to (outside the previously confined networks of interest). As part of a whole brain functional connectivity analysis, each of the 6 language factors were correlated to every one of the possible functional connections in the connectivity matrix, thresholded to identify the most significant correlations (i.e., those connections that were the most significantly related to language), corrected for multiple comparisons, and compared between groups. Details about this network analysis and associated statistics can be found below.

Machine learning classifier for language analyses **(Study 2)**

When developing models aimed at classifying different groups, it's important to ensure that the model is not too closely fit to the specific dataset that it is trained on such that it cannot generalize to novel datasets (a phenomenon known as over-fitting). The best way to verify that a model is not biased in this way is to test it on an independent validation dataset. Because we did not have a large enough sample size to feel comfortable leaving out a portion of the data as a validation set, we instead chose to use a cross-validation approach to test different patterns within the pooled training data. Here we used a Support Vector Machine (SVM) classifier. We reduced the problem of binary (responder vs. non-responder) classification to information provided by the first 9 features from the SVD analysis and the semantic similarity to each of the words in the strongest language factor. We implemented leave-one-out-cross-validation (LOOCV) consisting of the 42 participants assigned to treatment and 2 conditions (responder or non-responder). Discriminative models were computed by learning each of the parameters (features or word similarities, depending on the data provided) on N-1 subjects and testing the remaining (left-out) participant on all of the possible classifications. Accuracy of the models (from 0%, meaning failed at discrimination, to 100%, meaning perfect discrimination) was measured at each test and repeated for a total of 1000 iterations; average accuracy across all iterations was used as the final

accuracy assessment for each measure. For the data-driven approach, a single vector (the projections on the second feature) were entered into an SVM with LOOCV to test the result.

Whole brain functional connectivity and language network analysis **(Study 2)**

In **Study 1**, we used a "localizer" from a different dataset to narrow our focus on four networks of interest (the DMN, lateral frontal, parietal, and sensorimotor communities), resulting in 122 out of 272 ROIs being included in that analysis. Although language parameters were correlated to resulting functional links from those previous results, here we look at networks constructed from all 272 ROIs described in the first study (Power's parcellation plus 8 limbic regions of interest). We chose to perform a whole-brain functional connectivity analysis instead due a recent publication by Huth and colleagues [268] showing that language meaning is not confined to any one brain region but rather stored in an organized, redundant, and widespread set of functional networks distributed across the entire brain.

The 6 language factors were linearly correlated to each of the 73,712 functional connections of the weighted network (272 * 271 nodes). To identify the most important links, correlations between each language factor and the brain were thresholded at $p<0.001$. Only those connections that were significantly related to language according to this criterion were analyzed. Within- and between- network connectivity of these links was determined by calculating the percentage of functional connections out of the total possible that were between nodes of the same community (e.g., DMN to DMN) versus between nodes of different community designations (e.g., DMN to limbic). After controlling for false discovery rate (FDR, $p <0.05$), group comparisons between responders, non-responders, and no-treatment were performed on the remaining functional connections identified in each language factor. These significant r-values were extracted for all participants, averaged within each group, and entered in a one-way ANOVA with Bonferroni corrected post-hoc tests to determine differences with respect to the no-treatment control arm.

Correlating language with questionnaires **(Study 2)**

Over the course of the trial, participants filled out 29 unique questionnaires that aimed to capture trait and state characteristics of the placebo effect. A full list of all questionnaires used, along with their descriptions and information about scoring and modeling, can be found in **Study 1**. The questionnaire

data analyzed for **Study 2**, with the exception of the pain and mood questionnaires collected at every visit to determine treatment outcome, come primarily from self-report measures obtained at visit 1 pre-treatment, as we were interested in investigating the extent to which language correlated to predictors of placebo response. Questionnaires collected at visit 6 were also analyzed, as this time point corresponded to the date of the exit interview. Unlike previous reports where we looked for differences between placebo responders and non-responders based on their questionnaire scores, here we were interested in whether language corresponds to previous personality metrics which captured these differences, as well as what additional self-report measures were related to language differences seen between responders and non-responders. For this, we used simple Pearson correlations to capture relationships between interview data (language factors) and questionnaire data (subscores).

Calculating pain memory bias **(Study 3)**

From pain memory reports taken at Visit 2, we calculated a "discrepancy score" for each participant, which was the number obtained after subtracting their average pain over the last week (ratings on the phone app) from their recalled average pain over that same time period. This discrepancy score was the main parameter used in **Study 3**.

Subcortical shape analysis **(Study 3**)

Unlike volume measurements which provide a general whole-structure summary of a region, shape can reveal more subtle and nuanced changes in structure at subregional levels and is thought to indicate alterations or innate differences in the underlying neurocircuitry of a region. In order to capture locations of shape differences and the direction of these differences, points (vertices) are projected along the surface of a region according to predefined anatomical locations, and statistics are calculated on a vertex-by-vertex basis. Following structural segmentation and volume calculations, a vertex-based mesh surface for the right and left hippocampus was constructed based on the average surface of all participants; this was an automated procedure that was also part of FIRST (first_utils). To determine whether memory discrepancy was correlated with hippocampal shape, we created a design matrix where the first column contained the values of the variable of interest (discrepancy) in the same order as the

participants' files; these values were automatically demeaned as part of FIRST. We first generated 3D image files of the surface of each hippocampus to provide a general visualization of the results. These colored surfaces represented the uncorrected multivariate F statistics associated with pain discrepancy; the color indicated the strength of the correlation between discrepancy and shape at that location. Zooming in on the surface allowed visualization of arrows of various lengths – these indicated the direction of change of the shape in relation to discrepancy (referred to here as "vertex displacement"), with arrows pointing toward or below the surface representing a thinning or shrinking of the structure, while arrows pointing outward representing a thickening or expansion in that region.

Due to limitations in correction for multiple comparisons and the absence of additional statistics or quantification measures for the 3D surface mesh files, we re-ran our design to instead obtain a mask of the average surface shape (an outline around the hippocampus) and a 4D file with the displacement values at each of the 732 vertices of the hippocampus for all participants (one image per person). Positive values indicated outward displacements from the mean surface (expansion) whereas negative values indicated inward displacements from the mean surface (shrinkage). FSL's randomise option was then used for nonparametric statistics; data were permuted 5000 times and threshold-free cluster enhancement (TFCE) was used to correct for multiple comparisons.

The primary outputs from this analysis included an uncorrected t-statistic map, a map of corresponding uncorrected p-values (1-p), and a map of p-values post TFCE correction. To better visualize the results, we used Matlab (version R2015a) to extract the signed t-value at every vertex coordinate for all participants in the 4D file. T-values were thresholded at ± 2.0 to confine our statistics to the most significant areas, and for those vertices that met or exceeded this cutoff, a k-means algorithm was used to arrange and label these vertices into distinct clusters. The number of input clusters chosen was based on how many areas showed potential significance in the initial surface F-stat map. Because k-means provides arbitrary assignments to the data, the boundaries of partitions (in this case, the number of coordinates in each cluster) could change slightly from one repetition to the next. To ensure a stable cluster assignment, we ran the k-means algorithm for 50 iterations to obtain 50 patterns of different cluster assignments; from these, we found which set of clusters occurred most often and chose as the final cluster arrangement the grouping that had the highest similarity to the most common pattern. After

identifying this set of clusters, we extracted the signed shape displacement values from all of the vertices in each cluster for every person. These displacement values were then averaged within each cluster and correlated to the memory discrepancy scores. To account for possible memory or shape differences due to age or gender, we regressed age and gender values from vertex displacements to verify that the results did not significantly change (all statistics and figures are calculated with these covariates regressed). For internal consistency, the same methods were applied to the scanning data from the validation group.

For the secondary longitudinal analysis with CON and SBPp hippocampal data, we wanted to investigate whether areas identified in the CBP group changed as a property of time and/or pain. Therefore, we constructed a paired t-test design file for each group to compare the mean shape between the first scan (time 1) and the fifth scan (time 2). Vertex displacement values for both SBPps and CONs were extracted only from coordinates within hippocampal clusters that survived correction for multiple comparisons in the CBP discovery group. The average displacement values in these clusters and the difference in displacement between scans were compared between groups. Additionally, to test time effects, the correlation between the number of days between scans and the difference in average cluster displacement was calculated for both groups. Within-subject analyses were performed by first thresholding the t-values in each contrast at ± 2.0 (as was done in the CBP groups), and then testing whether those vertices that exceeded this cutoff had p-values <0.05 in their contrast map between scans (both uncorrected and corrected for multiple comparisons).

Modeling memory bias **(Study 3**)

We first examined the relationships between the retrospective ratings (which we call “recalled pain”), the real-time pain and mood measurements taken, and same-day pain and mood questionnaire scores through Pearson correlations. To assess adequacy of these variables in explaining pain memory and best compare our results with others’ cited in the literature, we tested four multivariate models: a pain only-model, a pain+mood model, a comprehensive behavioral model that combined previous pain and mood rating elements with the questionnaire scores, and a final neuroimaging + behavior model that combined the psychophysical and questionnaire variables with the neuroimaging results. Each multiple

regression analysis was used to quantify the influence of the given set of independent variables on the dependent variable of memory (recalled pain). Each model was tested using backward stepwise regression; criteria for inclusion was $p<0.05$ and criteria for exclusion was $p>0.10$. Additionally, a hierarchical approach to the regressions was used in order to conserve degrees of freedom. Thus, only those variables that remained in a prior model (i.e., significant or borderline based on our criteria) were then entered into the next model; if they did not meet criteria, they were discarded from any future models.

Validation analyses **(Study 3**)

To test the validity of our memory bias model, the hippocampal surfaces from the validation group of CBP patients were correlated to discrepancy using the same design as before (with discrepancy being the covariate of interest). Shape displacement was extracted within cluster boundaries using the vertex coordinates for each of the 3 clusters from the discovery group and averaged within each cluster. In addition to these measurements, all parameters from the phone app were also calculated. Using the regression equation from the discovery analysis, only those variables that remained within the model were entered from the validation group to predict their reported memory of pain (i.e., the intensity of pain that they recalled). This predicted memory score was then correlated to their actual memory reported to determine the accuracy of the predictive model.

Follow-up analyses **(Study 3**)

As part of the post-hoc phone call, we were interested in answering 3 main questions. First, we wanted to examine if any memory differences seen in CBP participants were representative of a general disturbance in short-term memory, since research suggests that chronic pain impacts attention and can cause both short-term and working memory deficits [74, 269-271]. Therefore, participant's overall scores on the 5-word STM task (questions 1 & 11 from **Table 3**) were correlated with their initial memory discrepancy while in the study ("previous pain discrepancy") and with the memory discrepancy of their study pain reported during the phone call ("current pain discrepancy"). Second, we wanted to investigate the specificity of our results: were any discrepancies found in participants' memory during the study *pain*-

specific or were they related to a person's baseline biases (i.e., their tendency to over- or under-exaggerate memories regardless of the valence, intensity, or context). We calculated the differences between participants' reported memories of different aspects of the study (questions 3-8 from **Table 3**) from their actual values while in the study to obtain discrepancies for each item (recalled – actual). These discrepancy values were then correlated to current and previous pain discrepancies. We also investigated whether personality played a role in these biases by correlating the discrepancy scores with 4 questionnaire measures relating to pain sensitivity, pain avoidance/anxiety, pain catastrophizing, and loss-aversion, some of which have been shown to influence memory in previous studies [272-277]. Third and finally, we wanted to know how generalizable our anatomical results were – did they also explain other behavioral or memory-related data in the study outside of pain memory. To investigate this, we correlated any surviving hippocampal clusters from our final multivariate linear model with the phone call discrepancy values and the scores from the selected personality questionnaires. Finally, we ran a mediation analysis using Mplus (7.0) to further explore the relationships between discrepancy, loss aversion personality, and hippocampal shape; the indirect effects of this mediation were tested by bootstrapping the data over 1000 iterations.

CHAPTER 3: RESULTS FOR STUDY 1

The neurobiological and psychosocial mechanisms predicting placebo response in chronic pain

Demographics

From the 82 CBP patients initially randomized, data from 63 patients that completed the study were analyzed: 43 of these patients had been allocated to the placebo treatment group and 20 patients had been allocated to the no-treatment group. 5 additional patients had been allocated to the active treatment group, but were not analyzed, as we were only interested in placebo response. The 43 patients in the placebo treatment group were exposed to multiple treatment periods and stratified into responders and non-responders based on their pain ratings using a smartphone app. Because of the temporal dimension of the data, we opted for a stratification strategy accounting for the amplitude of placebo analgesia after considering the variability across pain ratings. We used a permutation test randomizing pain ratings collected during baseline and those collected during each treatment period. Patients were dichotomized as responders (n = 24) if they responded to at least one of the two treatment periods and as non-responders (n = 19) if otherwise (demographics are presented in **Table 5)**.

Group	N	Age (years)	Female (%)	Duration of Pain (months)	Education (years)
Non-Responders	19	45.0 ± 13.4	N=5 (26.3)	62.0 ± 108.5	11.9 ± 3.2
Responders	24	46.9 ± 11.1	N=9 (37.5)	45.4 ± 48.9	12.7 ± 3.75
No Treatment	20	46.2 ± 13.2	N=10 (50.0)	57.6 ± 61.2	13.8 ± 3.9

Table 5: Demographics for Study 1. There were no differences between groups in age ($F(2,60)=0.12$, $p = 0.88$; one-way ANOVA), gender (Pearson $chi^2(2) = 0.90$, $p = 0.64$), duration of pain reported ($F(2,60) = 0.29$, $p =0.75$; one-way ANOVA), or years of education ($F(2,60) = 1.34$, $p = 0.27$; one-way ANOVA). Table shows mean ± STD; all variables here were reported at visit 1.

Importantly, external factors such as pain intensity during baseline, phone rating compliance, overall treatment compliance, treatment duration, and previous medication usage were not related to placebo

response (**Table 6**). Only self-reported number of hours of sleep per night dissociated placebo responders from non-responders, where responders reported sleeping about an hour less each night (**Figure 10**).

2a: Starting Pain Qualities			
Questionnaires at Visit 1	**Groups**	**Mean ± SEM**	**one-way ANOVA: F(2,60) = f-stat (p-value)**
NRS	Non-Responders Responders No Treatment	63.50 ± 4.13 56.28 ± 5.75 55.37 ± 3.78	1.33 (0.27) n.s.
MPQ_sensory	Non-Responders Responders No Treatment	13.21 ± 1.17 12.84 ± 1.02 14.20 ± 1.57	0.32 (0.73) n.s.
MPQ_affective	Non-Responders Responders No Treatment	3.12 ± 0.60 2.80 ± 0.58 3.35 ± 0.72	0.20 (0.82) n.s.
PainDetect	Non-Responders Responders No Treatment	9.38 ± 1.39 10.11 ± 1.02 13.66 ± 1.57	2.89 (0.06) &
NPS	Non-Responders Responders No Treatment	41.42 ± 2.59 37.83 ± 2.58 33.4 ± 2.45	2.29 (0.11) n.s.
2b: Other Potential Confounds			
Variable	**Groups**	**Mean ± SEM**	**one-way ANOVA: F(2,60) = f-stat (p-value)**
Pain app VAS Over Baseline	Non-Responders Responders No Treatment	5.85 ± 0.32 6.07 ± 0.20 5.91 ± 0.17	0.31 (0.73) n.s.
Phone App Compliance	Non-Responders Responders No Treatment	81.14 ± 3.35 80.39 ± 2.67 75.77 ± 4.06	0.73 (0.49) n.s.
Treatment Compliance	Non-Responders Responders No Treatment	90.7 ± 3.34 95.9 ± 3.84 n/a	**unpaired t-test: t-stat (p-value)** 0.73 (0.47) n.s.
2c: Medication Usage at Study Entry			
Medications for Pain		**# Patients (n = 63 total)**	**# Patients/Group**
NSAIDs or acetaminophen		24	7 Non-Responders 10 Responders 7 No Treatment
No treatment reported for pain		32	10 Non-Responders 11 Responders 11 No Treatment

Combination of NSAIDs with other drugs • **Gaba-ergic anticonvulsants** • **Muscle relaxers** • **Opioids/narcotics**	6	2 Non-Responders 3 Responders 1 No Treatment
Gaba-ergic anticonvulsants	1	1 No Treatment
2d: Medicine Quantification	**Groups**	**MQS Score ± SEM**
	Non-Responders Responders No Treatment	2.90 ± 0.76 3.22 ± 0.81 3.20 ± 0.78

Table 6: Placebo response is invariant to potential confounds. Initial pain levels, study compliance, and previous medication use are all variables that could impact placebo response; we therefore tested each of these measures to verify that the placebo effect in our cohort could not be explained by these potential confounds. **2a**. No significant group differences were observed in pain characteristics measured with NRS, MPQ-sf sensory or affective subscales, PainDetect, or NPS at study entry (visit 1). **2b.** There were no significant group differences in average pain rated using the electronic app over the 2-week baseline period, the percent compliance when using the electronic app to rate pain, and compliance when taking the study agent. Compliance was considered across the duration of the study (phone app). One patient from the non-responder group failed to bring back his medications for either of the treatment periods, and thus study agent compliance could not be calculated for this patient. **2c.** The Medication Quantification Scale (MQS) was used to calculate participant's pain medication usage at study entry; an MQS score is based on the kinds of medications a patient is taking, the dosage of those medications, and the number of medications in total. The MQS was broken down into medications reported by participants to treat their pain, which included no treatment at all, NSAIDS or acetaminophen, and a combination of anti-inflammatories with other drugs, such as gaba-ergic medicines, muscle relaxers, and opiates/narcotics. There were no differences in the amount of patients per group using these medications (Pearson chi^2 (6) = 3.1601; p = 0.79). **2d.** Likewise, there were no differences between groups in total MQS score at visit 1; F(2,60) = 0.05, p = 0.95 n.s.; one-way ANOVA). All participants were required to discontinue all medications used to treat their pain for the duration of the study beginning at visit 1. Unless otherwise stated, non-responders = 19; responders = 24; and no treatment = 20.

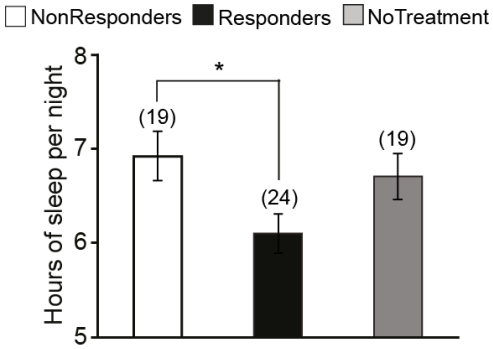


Figure 10: Placebo responders reported sleeping less than placebo non-responders. Numbers of hours of sleep per night obtained from the personal health information questionnaire were compared between the groups. There was a significant group difference ($F_{(2,59)}$ = 3.43; p = 0.04) where the placebo responders reported significantly less hours of sleep per night compared to non-responders. * p < 0.05 after Bonferroni correction.

The temporal dynamics of placebo response in RCT

Patients visited the lab on six occasions over 8 weeks and underwent identical scanning protocols on four of these visits (**Figure 11a**). Additionally, a battery of questionnaires was collected at each visit to capture patient's psychological profile and current emotional and pain states. Throughout the duration of the trial, the patients were asked to use a visual analogue scale (VAS) displayed on a smartphone app to continuously monitor their pain two times per day. Baseline pain levels were initially measured with the app for 2 weeks, after which patients were randomized either into a treatment or no treatment arm. Those in treatment were allocated either to active pills (Naproxen 500mg + Esomeprazole 20mg bid) or placebo pills (lactose bid) in a double-blinded fashion, while those in the no treatment arm did not receive medication but underwent the same procedures as participants in the treatment group. Treatment was introduced for two weeks followed by a one-week washout and reintroduced a second time for two weeks followed by a second one-week washout. Here, only brain images collected at visit 2 (pre-treatment baseline scan) and questionnaires collected at visit 1 (screening) were used to study the likelihood of clinical placebo response in chronic back pain.

Figure 11b shows how the permutation test was used to classify individuals into responders or non-responders based on their pain ratings. **Figure 11c** displays the time course of the percent change in pain measured with the phone app throughout the clinical trial for each group. On average, as expected, placebo responders showed an initial pain diminution that partially reversed back towards baseline levels during the first washout period before decreasing again as soon as the second treatment was introduced. Importantly, no-treatment and non-responder groups showed no change in back pain intensity throughout the trial. A closer examination of individual pain ratings ordered in time indicated that placebo responders displayed an instantaneous pain reduction at the first rating entered after the start of the treatment (**Figure 11d)**. Remarkably, the opposite effect was not true, as discontinuing the treatment did not cause the pain to instantaneously return to baseline levels (**Figure 11e).** Better understanding the temporal dynamics of clinical placebo response in the patients' natural environment could be critical in dissociating analgesia attributed to an active agent from that caused by a placebo in an RCT.

To investigate consistency of response across different assessments, we further compared the smartphone app pain measurements with other pain outcomes – participants' pain memory (a numerical verbal scale of their average pain over the last week), the numeric rating scale (NRS, a single value used to quantify pain changes in clinical trials [127]), the McGill Pain Questionnaire (MPQ) affective and sensory scales, the PainDetect, and the Neuropathic Pain Scale (NPS). Principal component analyses were performed on the pain outcomes at each treatment period to demonstrate that %change in pain questionnaires were clustered into 2-3 independent components dissociating pain outcomes measured with numerical scales (phone app, memory, and NRS) from outcomes assessing the qualities and characteristics of the chronic pain as measured with MPQ, PainDetect, and NPS (**Figure 11f**). Correlations between these measures can be found in **Table 7**.

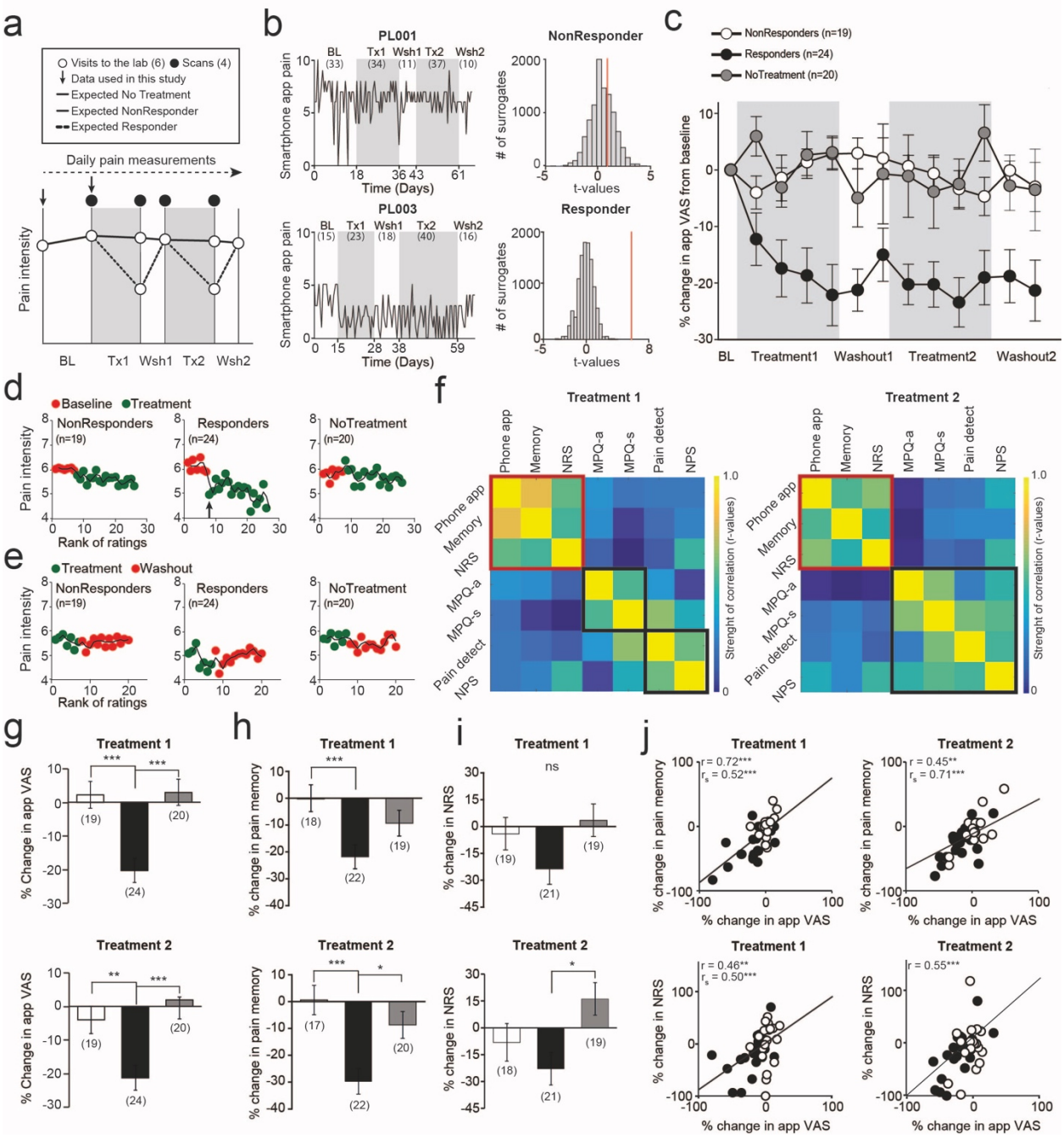


Figure 11: Characteristics of clinical placebo response. **a.** Experimental design and time line: CBP patients entered a six-visit 8-week trial. Although questionnaires were administered at every visit (white circles) and scans were collected at 4 visits (black circles), only the visit 1 questionnaires and visit 2 brain scans were used in this study (black arrows). **b.** Time series of smartphone app pain ratings in 2 patients (values in parentheses are the total number of ratings in a given period). Permutation tests randomised app pain ratings during baseline with those entered during each treatment period to detect significant decreases in back pain and identify placebo responders ($p<0.05$; PL001 was classified as a non-responder, while PL003 was a responder). **c.** Group-averaged %change in app ratings of back pain intensity is subdivided into two bins per week, averaged within each group, and displayed across the trial. **d.** Individual pain ratings ranked in chronological order for the last 7 ratings during baseline (red) and the

first 20 ratings of treatment 1 (green) averaged within each group. In responders, the placebo analgesia was instantaneous after the introduction of the treatment (black arrow). **e.** In contrast, the cessation of the treatment yielded little change in the patients' pain ratings. **f.** Principal component analyses indicated that pain outcomes were clustered in several factors at both treatment periods. The placebo response was specifically captured by pain outcomes of Factor 1 relying on numerical scales (squared in red). **g-i.** Significant differences in %change in app pain intensity were observed during the last week of both Treatment periods and a similar pattern and magnitude of back pain intensity diminution was observed with two unbiased outcomes: pain memory and NRS. **g.** Scatterplots show strong correspondence between %change pain app ratings with pain memory, and NRS. All post hoc comparisons were Bonferroni corrected: * $p< 0.05$; ** $p<0.01$ *** $p<0.001$. Scatterplot present the Pearson correlation coefficient (r) as well as the Spearman correlation coefficient when one of the variables was not normally distributed (r_s). Number of subjects is shown in parenthesis. BL: baseline; Tx1: Treatment 1; Wsh1: Washout 1; Tx2: Treatment 2; Wsh2: Washout 2.

Here, placebo treatment specifically impacted the intensity of pain quantified by questionnaires using numerical scales. Significant differences in %change in smartphone app pain intensity were observed during the last week of Treatment 1 ($F_{(2,60)} = 12.85$; $p < 0.0001$; one-way ANOVA) and Treatment 2 ($F_{(2,60)} = 9.72$; $p = 0.0002$; one-way ANOVA), during which placebo responders reported decreased back pain intensity compared to non-responders and no-treatment groups (**fig. 1g**). A similar pattern and magnitude of back pain intensity diminution for treatments 1 and 2 was observed with pain memory and NRS. For pain memory, there were significant group differences for Treatment 1 ($F_{(2,56)} = 5.40$; $p = 0.007$; one-way ANOVA) and Treatment 2 ($F_{(2,56)} = 9.39$; $p = 0.0003$; one-way ANOVA; **fig. 1f).** For NRS, group differences were trending during Treatment 1 ($F_{(2,50)} = 2.47$; $p = 0.09$; one-way ANOVA) and significant during Treatment 2 ($F_{(2,48)} = 4.36$; $p = 0.02$; one-way ANOVA; **fig. 1g**). Change in the memory of pain and NRS strongly correlated with the app pain ratings **(fig. 1j)** for both treatment periods**,** indicating that participants' perceived changes in pain were related to their spontaneous pain ratings.

Together, our results suggest that placebo response impacts a single dimension of chronic pain, as the intensity systemically diminished across all numerical scales leaving the intrinsic qualities of the pain unchanged. We conclude that changes in the qualities of the chronic pain may not be as salient in a placebo treatment compared to an active treatment. Further studies will however be necessary to test this proposition.

Pain Outcomes by Treatment Session (% change in score)	Treatment 1	Treatment 2
	VAS (phone app)	VAS (phone app)
NRS	**r = 0.46, p = 0.003, n = 40**	**r = 0.55, p = 0.003, n = 39**
Pain Memory	**r = 0.72, p <0.001, n=40**	**r = 0.45, p = 0.004, n = 39**
NPS	r = 0.10, p = 0.51, n = 43	**r = 0.39, p = 0.01, n = 41**
MPQ-sensory	r = 0.06, p = 0.71, n = 42	r = - 0.06, p = 0.67, n = 41
MPQ-affective	r = 0.21, p = 0.23, n =34	r = - 0.16, p = 0.39, n = 31
PainDetect	r = 0.05, p = 0.75, n = 41	r = 0.13, p = 0.42, n = 39

Table 7: Comparison of pain outcome measures. All pain-related scores were first converted to percent change from baseline for both treatment sessions. Table shows Pearson correlations between the percent change using the smart phone app (from which we based our stratification of placebo responders/non-responders) with percent change in other pain outcomes. Pain questionnaires using numerical scales such as the verbal pain memory report and the Numeric Rating Scale (NRS) significantly correlated with the phone app. Of the remaining pain outcomes, only the Neuropathic Pain Scale (NPS) showed modest correlation with the VAS from the phone app, which was for treatment 2 only. Correlation coefficients, p-values, and corresponding number of patients are provided.

Blinding of the analyses (a reminder)

Given the lack of reproducibility in scientific findings [254] and the importance of transparency in data analysis, we followed recommendations by MacCoun and Pearlmutter [255] and employed cell scrambling to blind our neuroimaging and questionnaires data and minimize bias. All analyses performed on brain imaging data and questionnaires therefore followed a pre-determined plan of analysis presented in **Figure 12** and were performed three times according to the blinding procedure (one for each classifier; one correct classifier and two scrambled classifiers) as explained in the methods.

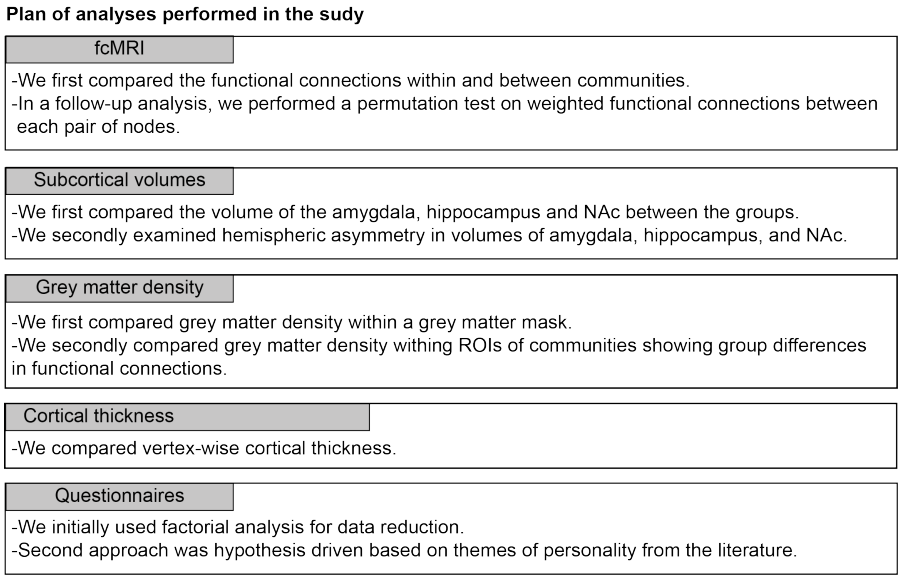


Figure 12: Pre-determined plan of analyses. The plan outlined here presents every analysis performed on this data; this plan was decided upon prior to study completion. All analyses were performed blindly using 3 classifiers to minimize the effects of bias and expectations. No analyses were performed outside of the plan until post-hoc tests were run, which were completed after un-blinding the data.

Functional circuitry of the lateral frontal cortex predisposes patients to placebo response

The traditional mapping of a behavioral response to a single brain area or a set of regions is often an overly simplistic approach to understanding cognitive function, especially in the case of complex phenomena such as the placebo effect. We instead applied techniques used in network science to construct brain network from resting state Pearson correlations computed from the average time course between brain regions (parcellation scheme from [258], **Figure 13a-b**).

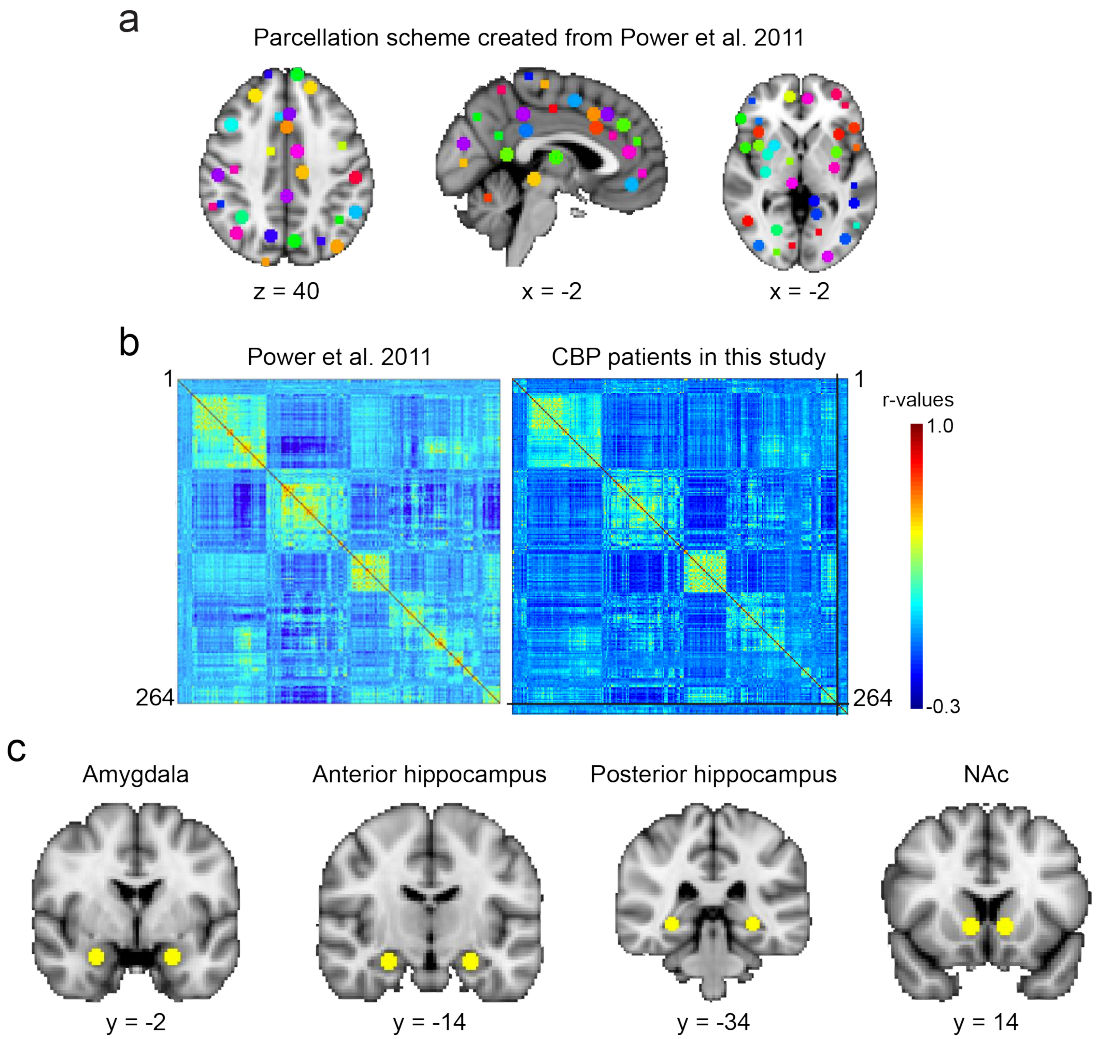


Figure 13: Parcellation of the brain into 272 ROIs. a. In this study, we used the parcellation scheme from Power et al 2011[258]. They constructed functional brain networks from 5-mm radius ROIs located in 264 regions representing the putative functional area of the brain. These regions were identified from a meta-analysis on task-based fMRI and complemented by using the center of gravity of cortical patches constructed from resting state functional connectivity. **b.** The similarity between the averaged connectivity matrices of 106 healthy subjects from Power's study with the averaged connectivity matrices from our 63 CBP patients can be visually appreciated. The ROIs extending below the black line (ROIs 265-272) represent the ROIs that were manually added in subcortical regions of interest displayed in **c**. below. **c.** These limbic ROIs represented 5-mm radius parcels were added in amygdala, bilateral anterior hippocampus, bilateral posterior hippocampus, and bilateral nucleus accumbens (NAc).

A modularity analysis was performed to segregated nodes of the network into 6 communities but all statistical analyses were restricted to the default mode network (DMN), the sensorimotor, the lateral prefrontal, and the posterior communities because they were analogous to placebo-related networks of interest derived from an independent data set of osteoarthritis patients (OA) exposed to placebo

treatment in an RCT (**Figure 14a-b**). Subcortical limbic regions were manually added to the networks because of their prominent role in chronic pain [45] and placebo response [99] (**Figure 13c)**. This approach has many advantages, including increasing statistical power by limiting the number of comparisons, preventing over-fitting of the data, and identifying hypothesis-driven biological mechanisms.

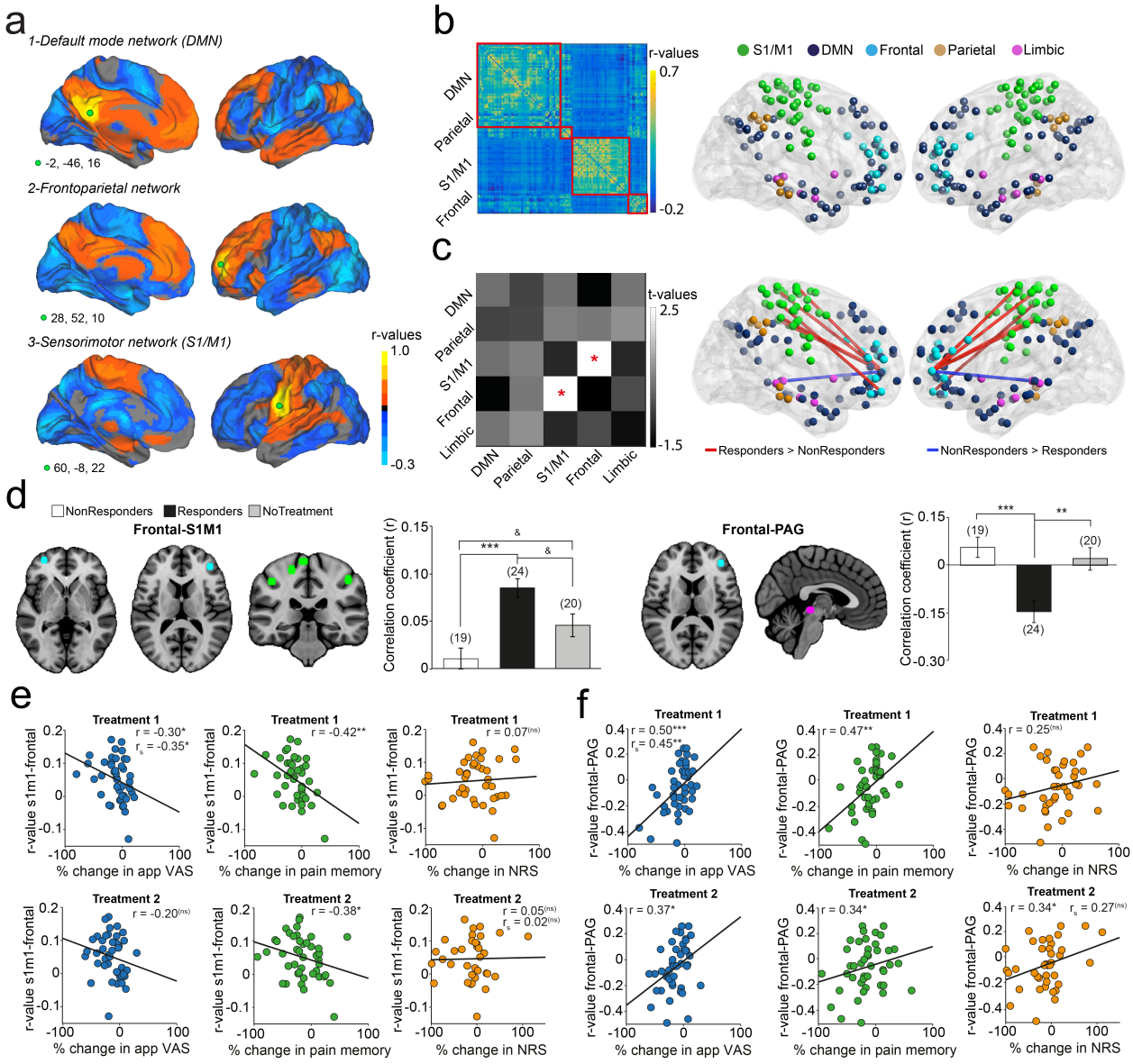


Figure 14: Lateral frontal circuits predisposing patients to placebo response. a. Our previous results indicated that degree counts in the middle frontal gyrus, the anterior cingulate cortex, the posterior cingulate cortex, and sensorimotor cortex predicted placebo response in osteoarthritis patients (Pascal). Seed-based coactivation maps derived from 1000 healthy subjects (http://neurosynth.org) were generated from the functional connectivity between these ROIs, represented by a green seed, with other regions of the brain and used as a localiser to restrict our resting state functional analyses. **b.** The connectivity matrices was restricted to communities overlapping with the networks of interest presented in

a. **c.** Group comparisons indicated that the averaged functional connectivity between the 12 ROIs of the frontal network and the 38 ROIs of the sensorimotor networks were significantly stronger in placebo responders compared to nonresponders. Network based statistics using randomised weighted connections of all possible connections of the network indicated stronger connectivity between nodes located in the sensorimotor community with nodes from the frontal community and weaker connectivity between a node located in the DLPFC and the PAG (FDR-corrected $q < 0.05$). **d.** The averaged correlation coefficient (r-value) of significant connections reported in **c.** is presented for each group. **e,f.** Scatterplots show moderate but significant relationships between the strength of functional connectivity in these frontal circuits with pain outcomes measured by the phone app (blue) and memory of pain (green), but not by NRS (orange). All post hoc comparisons were Bonferroni corrected: $^{\&}$ $p<0.10$; * $p< 0.05$; ** $p<0.01$ *** $p<0.001$. Scatterplots present the Pearson correlation coefficient (r) as well as the Spearman correlation coefficient when one of the variables was not normally distributed (r_s). Number of subjects is shown in parenthesis.

Placebo responders displayed stronger functional connectivity across the lateral prefrontal community and sensorimotor community compared to non-responders ($t_{(41)} = 2.50$; uncorrected $p = 0.018$; two sided unpaired t-test; **Figure 14c)**. Performing a permutation test on the weighted connections between all possible pairs of nodes revealed that placebo responders displayed stronger connections between nodes pertaining to the sensorimotor and frontal communities (frontal-S1M1) and weaker connections between a node located in the dorsolateral prefrontal cortex (DLPFC) and a node located in the periaqueductal gray (PAG; frontal-PAG) (**Figure 14c).** The average r-values of these significant connections strongly dissociate placebo responders from non-responders and indicate that the no treatment group likely represents a mixture of responders and non-responders as would be expected (frontal-S1M1: $F_{(2,60)} = 11.67$; $p < 0.0001$; one-way ANOVA; frontal-PAG: $F_{(2,60)} = 10.41$; $p < 0.0001$; one-way ANOVA; **Figure 14d)**. Importantly, the strength of the functional connectivity of these systems correlated with the amplitude of placebo analgesia measured by the app ratings and pain memory, and the NRS (**Figure 14e-f**). These results were invariant to potential confounds such as head motion (**Figure 15**). The two other classifiers representing scrambled codes from our blinding procedure yielded no significant group differences.

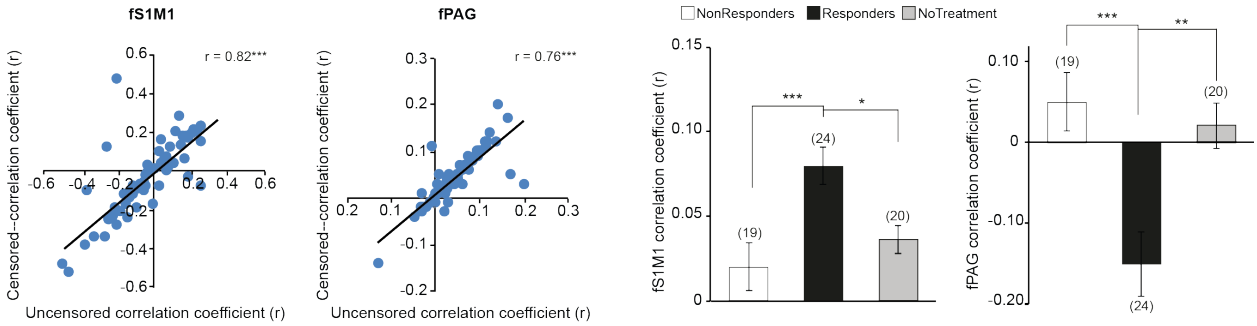
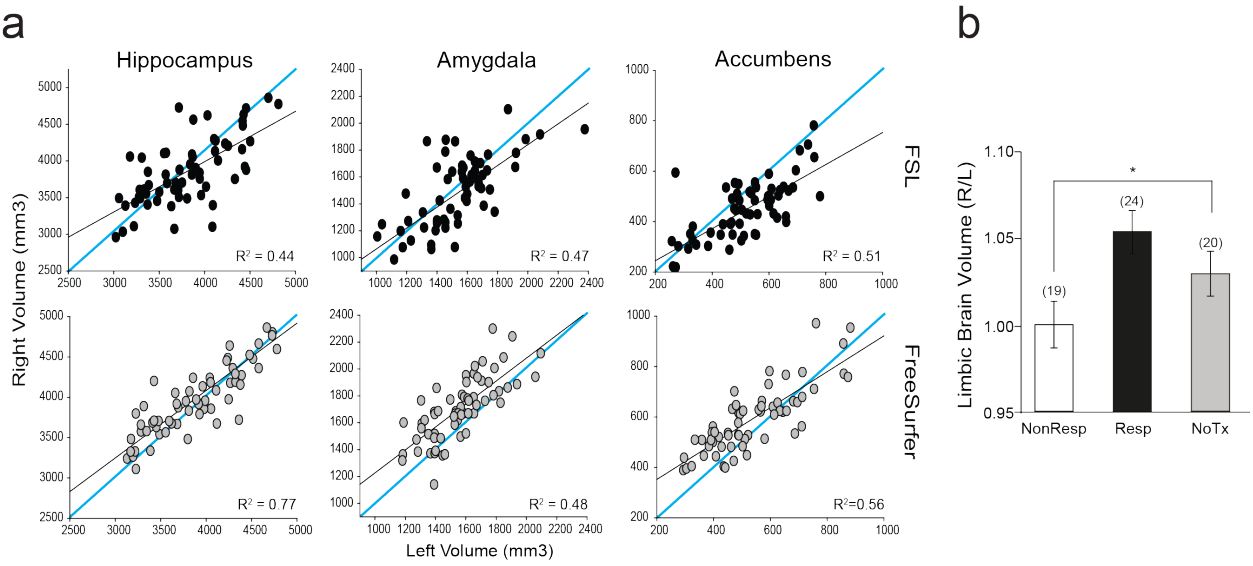


Figure 15: Functional results were invariant to volume censoring. Any given volume along with the one preceding and following it were removed from the time series if *i.* the derivative of the sum of absolute head displacement exceed 0.2 mm, *ii.* the root mean squared of the derivative of global BOLD signal intensity exceeded a z-score of 2.3, or *iii.* the standard deviation of the BOLD signal intensity exceeded a z-score of of 2.3 (this procedure is suggested by [257]). The scatterplots show very robust correlation between the original functional connectivity in frontal-S1M1 (fS1M1) with the one obtained after volume censoring. Volume censoring did not change our initial result of stronger functional connectivity between frontal-S1M1 and weaker functional connectivity in frontal-PAG in placebo responders presented in **Figure 14**. * p< 0.05; ** p<0.01 *** p<0.001.

Trait characteristics from brain anatomy predetermine placebo response

The volumes of the NAc, amygdala, and hippocampus were examined because of their involvement in motivation, learning, and contextual memory and because they represent risk factors for developing pathological emotional states [278, 279] and chronic pain [55]. Volume comparison of subcortical volumes between placebo responders and non-responders yielded no significant results. Inter-hemispheric laterality of limbic structure volume, however, indicated that placebo responders showed leftward asymmetry compared to non-responders, which is in accordance with the proposition that subcortical structural lateralization plays a role in psychopathologies [280]. Importantly, we replicated our finding when using another segmentation software (Freesurfer) (**Figure 16a**). **Figure 16b** shows the statistical differences between the 3 groups ($F_{(2,60)}$ = 4.14; p = 0.02; one-way ANOVA). The two other classifiers representing scrambled codes according to our blinding procedure yielded no significant group differences in volumes or asymmetry.



Supplementary Figure 5: Limbic volume asymmetry found using both FSL and Freesurfer. Subcortical volumes were segmented and extracted bilaterally for 3 regions of interest: hippocampus, amygdala, and nucleus accumbens; this was done twice using two different softwares – FSL and Freesurfer. **a.** For both FSL and Freesurfer, volumetric asymmetry between hemispheres was found for all regions. Scatterplots show Pearson correlations between right and left volumes of the hippocampus, amygdala, and nucleus accumbens for FSL (black, top) and Freesurfer (gray, bottom) for all patients. X-axis = left hemisphere; y-axis = right hemisphere. Blue lines represent a 1/1 ratio where left and right volumes would be equal to each other; values above this line indicate more rightward asymmetry, whereas values below indicate more leftward asymmetry. **b.** The volumes from all 3 regions were summed for each hemisphere and represented as a ratio, with total right limbic volume in the numerator and total left limbic volume in the denominator. Freesurfer's segmentation replicated the significant differences between responders and non-responders initially observed using FSL $F_{(2,60)} = 4.42$; $p = 0.02$; one-way ANOVA between 3 groups). Results from only the correct classification (classifier #3 for subcortical volume analysis) is shown for all plots. * $p < 0.05$ after Bonferroni correction.

Differences in cortical properties between responders and non-responders were further assessed with voxel-based morphometry (VBM) of grey matter density and cortical thickness. The initial VBM analysis yielded no significant differences between responders and non-responders (**Figure 17a-b**). In a follow-up analysis, we extracted grey matter density for nodes pertaining specifically to the sensorimotor community and the frontal community, since the relationship between these two communities indicated a functional predisposition to placebo response. This follow-up analysis indicated that non-responders displayed denser grey matter within the right postcentral gyrus compared to responders and no treatment groups ($F_{(2,60)} = 6.95$; $p = 0.0019$ one-way ANOVA; **Figure 17c**).

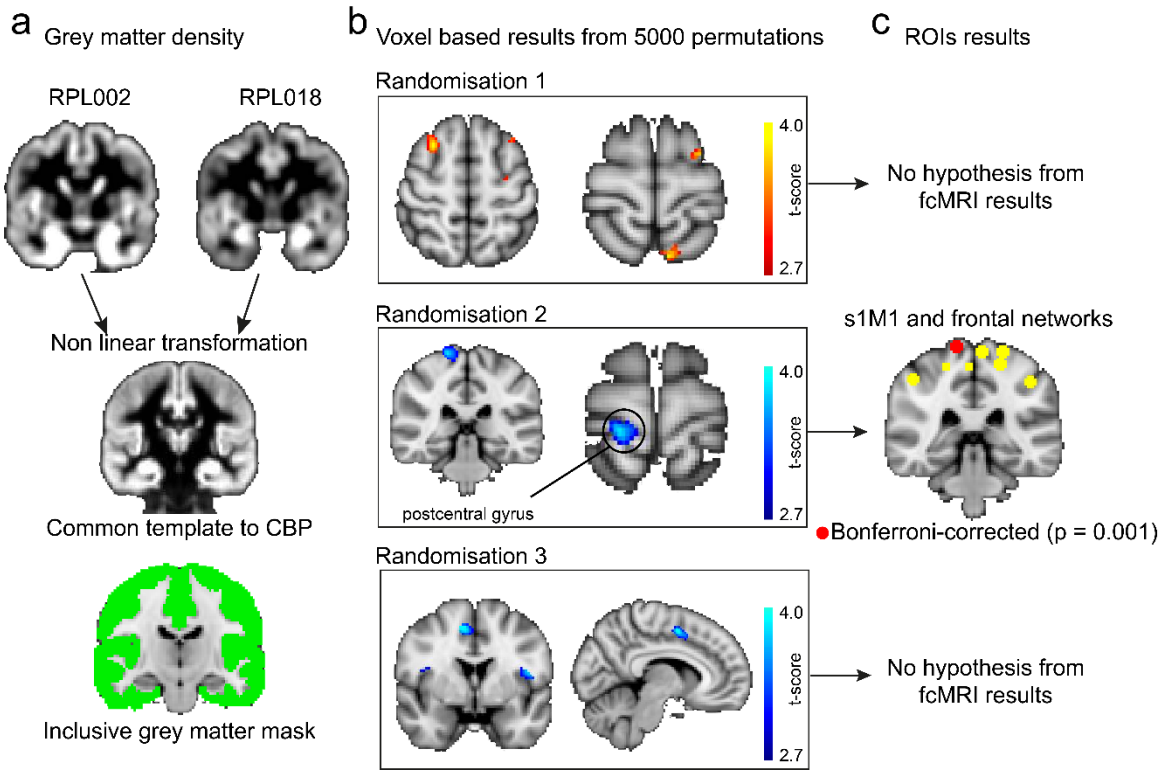


Figure 17: Placebo non-responders showed higher grey matter density in the postcentral gyrus (Classifier 2). Grey matter density was examined using voxel-based morphometry (VBM). **a.** All T1-weighted images were first brain extracted and then segmented into grey matter, white matter, or cerebrospinal fluid using FSL software. A common grey matter template was generated for CBP by registering and averaging all grey matter images. The grey matter image of each participant was then registered to the common template using non-linear transformation. A voxel-wise permutation test (fsl randomize) was used to generate group differences within an inclusive grey-matter mask. All three classifier codes generated results that did not survive threshold free cluster enhancement (TFCE). As a follow-up analysis, we extracted the grey matter within the ROIs that were most strongly functionally connected in placebo responders (classifier 2). The results indicate higher grey matter density in one ROI within the sensorimotor network located in the postcentral gyrus (Bonferroni corrected p = 0.001 (0.05/50 ROIs)).

In addition, anatomical differences between responders and non-responders were finally examined using vertex-based cortical thickness where non-responders showed thicker cortex in the right superior frontal gyrus compared to placebo responders and no treatment groups ($F_{(2,60)} = 6.74$; $p = 0.002$; one-way ANOVA). The two other classifiers representing scrambled codes generated no significant group differences. Unlike functional connectivity which fluctuates depending on attention levels, metabolites, and other external factors [281], the presence of differences in all three of these anatomical properties before treatment provide us with strong evidence that the placebo response stems, in part, from a

hardwired brain predisposition invariant to external factors such as expectations and context. A summary of the 3 anatomical findings can be found in **Figure 18** below.

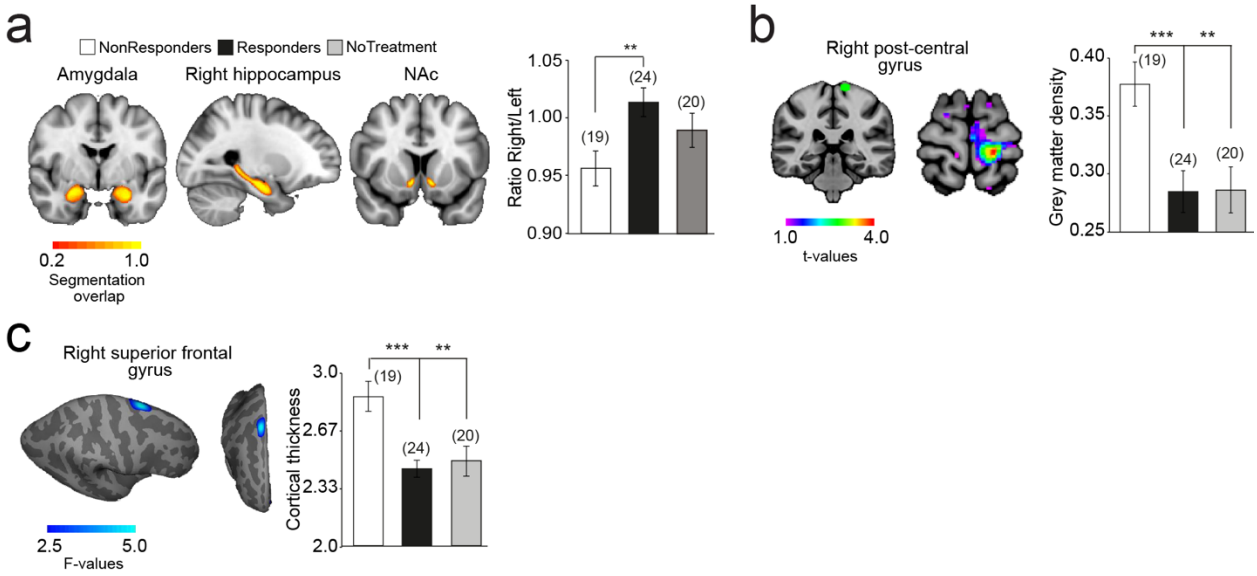


Figure 18 Summary of the anatomical characteristics predetermining placebo response. a. Heat maps display the overlap of automated segmentation of the nucleus accumbens (NAc), the amygdala, and the hippocampus performed with FSL software across all patients. The placebo responders displayed a leftward asymmetry in the volume of these subcortical structures. **b.** Grey matter was denser in the node located in the post-central gyrus in non-responders compared to responders. **c.** Vertex-wise analysis indicated that placebo non-responders had thicker cortex compared to responders (FDR corrected $q < 0.05$). The bar graph shows the GMD values for all 3 groups.

Emotion-related psychological traits predetermine placebo response

We administered an exhaustive battery of 47 subscales from 29 questionnaires that captured both the psychological properties previously associated with placebo response in healthy participants as well as those suggested to be involved in pain chronification. Because of the high amount of data, the analysis plan included a data driven approach relying on PCA (the approach failed to dissociate responders from non-responders and was dropped; see methods for detail) and a hypothesis driven approach organizing the data by themes. According to the latter method, 29 items from 13 questionnaires were organized into 4 themes chosen apriori: *theme 1*: attitudes and emotions toward pain; *theme 2*: positive affect/personality; *theme 3*: negative affect/personality; and *theme 4*: emotional regulation/awareness. Separate logistic regressions were performed on items within each theme, and a combination of 7 items explained a significant amount of variance in placebo response: Pain

Catastrophizing scale (PCS) rumination and PCS helplessness from *theme 1*, Beck Depression Index (BDI) and Loss Aversion Questionnaire (LAQ) from *theme 3*, and Emotional Regulation (ERQ) suppression, Multidimensional Assessment of Interoceptive Awareness (MAIA) emotional awareness, and MAIA not-worrying from *theme 4*.

These items were then combined into one final questionnaire model that predicted likelihood of response using backwards, stepwise logistic regression. LAQ was removed first (p = 0.16) followed by BDI (p=0.05). The resulting model explained 54% of the variance (pseudo R^2 = 0.54, p <0.001) in placebo response (**Table 8)**. The scrambled code from the 2 other classifiers of our blinding procedure indicated that classifier 2 failed to explain variance in outcome, and classifier 3 had only one item – ERQ_suppression (z = 2.13, p = 0.03) explaining little of the variance (pseudo R^2 = 0.08, p = 0.02).

Classifier	Item	Coef.	SEM	z-stat	p-val of item	[95% CI]		LR chi2	p-val of model	Pseudo R2	Log likelihood
C1	*PCS rumination*	0.71	0.32	2.18	0.029	0.07	1.34	31.64	P< 0.0001	0.536	-13.69
	PCS helplessness	-0.64	0.28	-2.32	0.021	-1.18	-0.1				
	MAIA emotion	2.62	0.98	2.68	0.007	0.70	4.54				
	ERQ suppress	-0.41	0.18	-2.34	0.019	-0.76	-0.07				
	MAIA notworry	-1.84	0.70	-2.60	0.009	-3.22	-0.45				
C2	n/a	-	-	-	-	-	-	0	n.s.	0	-30.45
C3	*ERQ suppress*	0.12	0.06	2.13	0.03	.01	0.23	5.44	0.02	0.089	-27.73

Table 8: Final questionnaire model (Classifier 1). The logistic regressions from the theme analysis resulted in 7 potential self-report items, including 2 subscales from PCS (rumination and helplessness from theme 1), 2 subscales from MAIA (emotional awareness and not-worrying from theme 4), 1 subscale from ERQ (suppression from theme 4), the BDI (theme 3), and the LAQ (theme 3). All of these items were entered into a backwards, step-wise logistic regression explaining stratification for placebo responders and non-responders for each of the 3 classifiers. For classifier 1 (the correct one, C1), BDI and LAQ were excluded, leaving 5 items explaining over half of the variance in placebo response. Classifier 2 (C2) had no significant items that predicted response, and Classifier 3 (C3) had only one item – ERQ_suppression – significantly predict response, explaining only ~9% of the variance in the data.

A model predicting clinical placebo response in RCT

Figure 19a shows that the psychological traits associated with placebo propensity were largely dissociable from the brain markers, with little correlation between the two dimensions. To better understand how brain biomarkers complemented the identified psychological characteristics, we

introduced each significant brain marker with the 5 significant questionnaires in a series of multiple regressions presented in **Table 9**. The results indicated that functional connectivity of frontal-S1M1 and frontal-PAG were independent contributors to response, while the subcortical asymmetry and grey matter density in the post central gyrus were not kept in their respective models. It also indicated that emotional awareness and emotional suppression were key psychological parameters in explaining the placebo response, as these were the only two questionnaire items to consistently survive all regressions regardless of the brain parameter.

Variables	Surviving Item	Coef.	SEM	z-stat	p-val of item	[95% CI]		LR chi2	p-val of model	Pseudo R2	Log likelihood
Post-central gyrus gray matter density + 5 self report items	PCS rumination	0.71	0.32	2.18	0.029	0.07	1.34	31.64	P< 0.0001	0.536	-13.71
	PCS helplessness	-0.64	0.28	-2.32	0.021	-1.18	-0.10				
	MAIA emotion	2.62	0.97	2.68	0.007	0.70	4.54				
	MAIA not-worry	-1.84	0.70	-2.6	0.009	-3.22	-0.45				
	ERQ suppression	-0.41	0.18	-2.34	0.019	-0.76	-0.07				
Subcortical Asymmetry + 5 self report items	PCS rumination	0.71	0.32	2.18	0.029	0.07	1.34	31.64	P< 0.0001	0.536	-13.71
	PCS helplessness	-0.64	0.28	-2.32	0.021	-1.18	-0.10				
	MAIA emotion	2.62	0.97	2.68	0.007	0.70	4.54				
	MAIA not-worry	-1.84	0.70	-2.6	0.009	-3.22	-0.45				
	ERQ suppression	-0.41	0.18	-2.34	0.019	-0.76	-0.07				
Frontal-S1M1 connectivity + 5 self report items	ERQ suppression	-0.27	0.12	-2.33	0.02	-0.49	-0.04	33.81	P< 0.0001	0.572	-12.61
	Frontal w/ S1M1 connectivity	36.18	12.44	2.91	0.004	11.79	60.58				
	MAIA emotion	1.95	0.77	2.52	0.01	0.44	3.46				
Frontal-PAG connectivity + 5 self report items	Frontal w/ PAG connectivity	-11.85	4.30	-2.75	0.006	-20.3	-3.42	31.61	P< 0.0001	0.535	-13.71
	ERQ suppression	-0.33	0.12	-2.72	0.007	-0.56	-0.09				
	MAIA emotion	1.92	0.70	2.74	0.006	0.54	3.29				

Table 9: Creation of the final multivariate model. After un-blinding, the data from the correct classifier in each data set was used to create a final multivariate model. To avoid over-fitting, we split the data according to the brain data, with each of the 4 significant neuroimaging results (post-central gyrus gray matter density, subcortical asymmetry, frontal-s1m1 functional connectivity, and frontal-pag functional connectivity) being entered in separate backwards stepwise models with the 5 significant questionnaire items from **Table 8** (PCS rumination and helplessness, MAIA emotion and not-worrying, and ERQ suppression). Shown here are the results of all 4 regressions (n=43 observations in each). For two models – gray matter and asymmetry – the brain parameters did not survive the model and all 5

questionnaire items remained. For the other 2 regressions, the two different functional connectivity measures, (shown in blue) survived their régressions and eliminated 3 questionnaires each. The same 2 questionnaires were consistent (p <0.05) in all 4 models : MAIA-emotion and ERQ_suppress (shown in red). Given this, we entered these 2 questionnaires along with the 2 surviving functional neuroimaging parameters into a final multivariate model presented in **Figure 19**.

Finally, we combined the 2 surviving functional connectivity imaging parameters with the 2 surviving questionnaires as the signature of placebo response in RCT (**Figure 19b**). All variables remained significant in a logistic regression (p<0.05) and explained 71% of the variance. We next used support vector machine (SVM) as a classification algorithm and determined that our predictive signature showed cross-validated prediction accuracy of 83.4% in a leave-one-subject out procedure (**Figure 19c)**. Critically, applying this signature to the no-treatment group indicated a mixture of responders (n=11) and non-responders (n=9) matching the actual proportions of real responders and non-responders observed in the treatment group. Similar levels of accuracy were obtained using alternative classification methods such as a logistic regression based on Mahalanobis distance (accuracy of 85.7%) or Naïve Bayes model (accuracy of 76.1%). **Figure 19d** shows the posterior probability distribution plots of the Bayesian model for each pair-wise combination of variables.

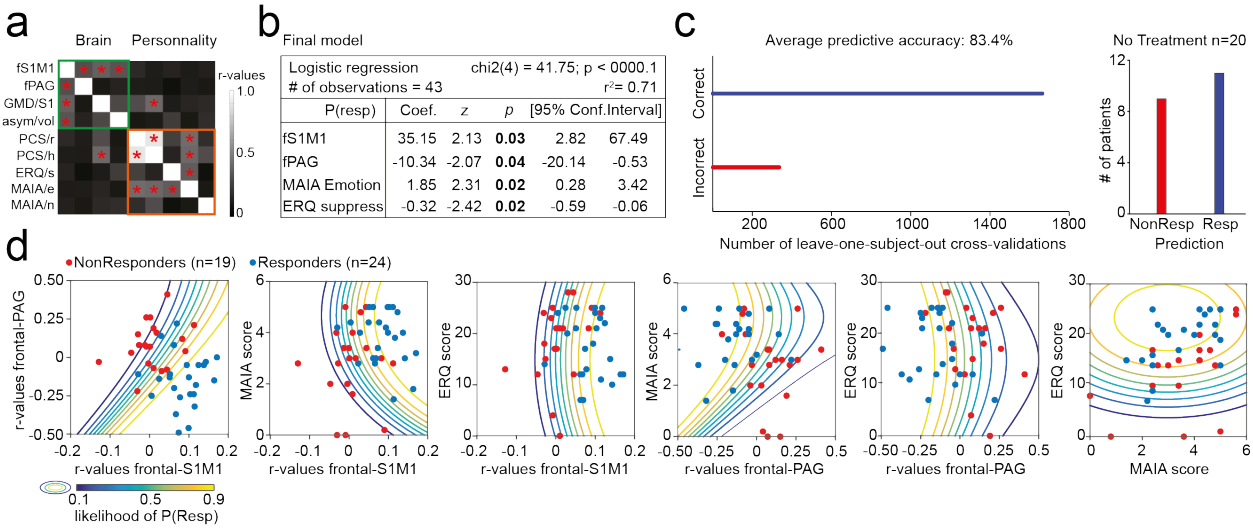


Figure 19: A signature predicting placebo response from the brain and personality. The correlation matrix shows poor correspondence between the psychological characteristics of placebo response in CBP patients and the multimodal brain parameters predisposing patients to placebo response. **b.** Logistic regressions were used to demonstrate that only functional connectivity between lateral frontal cortex with S1M1 and PAG explained unique variance unexplained by our psychological characteristics. Our final model indicates that these four parameters were independent contributors explaining 71% of the variance

in placebo response. **c.** Support vector machine (SVM) using a leave one subject out procedure obtained 83.4% accuracy in predicting placebo response according to the four significant parameters and the histogram indicate the probability of placebo response in the no-treatment group. **d.** Posterior probability distribution plots for the pair-wise combination of the four model variables. For each plot, the probability of response is indicated by the color of the circular lines, where dark blue indicates a low probability of response and light yellow indicates a high probability of response. The points in the graphs are the training data off of which the probability functions were defined (blue = responders and red = non-responders).

Discussion

This is the first randomized partially-blind clinical trial that was specifically designed to identify factors contributing to placebo propensity in individuals suffering from CBP. The fact that the no treatment arm showed minimal regression to the mean and the fact that placebo response was time-locked and sustained with treatment periods indicated a true placebo response rather than natural recovery or symptoms fluctuations. We first demonstrated that emotion-related characteristics are important predictors of the response, which may be specific to chronic pain populations. We secondly showed that placebo response was predetermined by anatomical properties invariant to external factors and by functional connectivity between frontal regions with the sensorimotor community and the PAG. We then demonstrated that the likelihood of placebo response in RCT could be accurately predicted using a signature developed from functional connectivity and personality traits.

Several pitfalls have been raised when trying to predict complex behaviors like placebo response, the most important being that the placebo effect studied in healthy individuals has been highly variable and showed lack of reproducibility[282]. Moreover, conflicting results have emerged from the literature because of the multiplicity of experimental designs, plans of analysis, and pain outcome measures. Here, many of these limitations were accounted for by incorporating several methodological strategies. First, we used the settings of a RCT during which patients were exposed to two treatment periods each separated by a washout period to better determine the propensity of placebo response to more than one treatment regimen. Importantly, our RCT included a no treatment arm documenting the natural fluctuations of chronic pain. Second, smart phone technology was used to track patients' pain twice a day throughout the duration of the study and was compared to other independent pain outcomes in order to better define and understand the response. Third, all functional analyses were limited to networks of interest-identified

apriori in two independent cohorts of chronic pain patients[250]. Finally, all analyses were performed blindly using one real code and two scrambled codes to minimize the effects of bias and expectations in the findings. Therefore, this study is novel not only in terms of its advancement in the understanding of placebo response but also in terms of its methodological soundness.

Previous work in healthy individuals demonstrated that placebo response is known to recruit endogenous pain pathways acting upon the opioid system to regulate descending inhibition throughout the periaqueductal grey (PAG) [109, 110] a mechanism that can be reversed by the administration of naloxone [11, 111]. Besides these anti-nociceptive circuits, which reduce spinal responses to pain [112], the placebo effect is also dependent on limbic circuitry and higher-order frontal mechanisms involved with context generation and de-coding, expectations of treatment outcome, and emotional appraisal of events [99]. The relative contribution of each of these cognitive, emotional, and anti-nociceptive systems to the clinical placebo response remained unknown. Here, we showed that placebo analgesia in an RCT depended on a pre-existing complex circuitry coupling the lateral frontal cortex with the sensorimotor system, one of the major projection sites from the spinothalamic tract, and the PAG modulating endogenous descending inhibition. Moreover, the PAG also encodes aversive prediction errors (PEs) [283] and likely computes the mismatch between the expected pain relief from treatment pills and the actual episodes of spontaneous pain. During long lasting placebo analgesia observed in RCT, higher connectivity between lateral frontal cortex in placebo non-responders may therefore reflect a higher ability to learn from PEs that nullify the effects of expectation, whereas placebo responders (showing anti-correlation) may fail to adapt their expectation and continue to experience a decrease in pain intensity.

Interestingly, none of the often-cited traits in placebo literature - dispositional optimism, anxiety, extraversion, neuroticism, agreeableness, or loss/harm avoidance – were able to successfully differentiate placebo responders from non-responders in our cohort. Instead, placebo response was primarily driven by both an increased emotional awareness (noticing how one's body feels and changes with emotions and experiences) and by a decrease in worrying about discomfort (noticing one's pain but not becoming upset from it), both from MAIA. These resulting subscales fit well with those from PCS indicating higher rumination (i.e., they thought more about pain) and lower helplessness (i.e., they weren't held back by their pain) in placebo responders. Placebo responders had lower scores in expressive

suppression of emotions on the ERQ than their non-responder counterparts, an observation that is complementary to the PCS and MAIA findings and indicates less of an attempt to hide or reduce expression of emotions. Altogether, our results revealed that placebo responders appear to have an ability to recognize subtle cues in the body regarding emotional and physical well-being, to remain attentive to these cues and emotions by not ignoring or suppressing them, and to choose to accept these states as opposed to becoming worried or burdened by them.

We finally developed a predictive signature of placebo response in clinical trials based on the brain functional connectivity and the personality traits of the patient. This is a crucial step towards understanding the mechanisms of clinical placebo response observed in a clinical context. The growing literature indicates that placebo response can be predicted, which can pave the way for using placebo as a therapeutic option with limited side effects across multiple chronic pain conditions. This also represents an advance that may beneficially impact phase 3-drug testing through the identification and subsequent removal of placebo responders from clinical trials to improve accuracy in efficacy assessments of novel pain medications.

CHAPTER 4: RESULTS FOR STUDY 2

Semantic language properties underlying the placebo response in chronic low back pain patients

Demographics

Of these initial 125 participants enrolled, 66 people completed all aspects of the study, including an exit interview at the final visit (**Figure 20a**). Of these individuals, 4 received active treatment and were removed prior to any analysis for a final sample size of n = 62 CBP patients. These participants were well matched in key demographics, including 37 men (48.8 ± 1.9 years of age; 12.7 ± 4.2 years of education) and 25 women (41.84 ± 2.7 years of age; 13.2 ± 3.0 years of education). 20 of these participants were assigned to the no treatment arm; of the remaining 42, the permutation test of pain ratings classified 19 as placebo non-responders and 23 as placebo responders. A summary of these demographics can be found in **Table 10.**

Group	N (62 total)	Age (years)	# Females (%)	Income Level[&]	Education (yrs)	Pain Duration (years)
Non-Responders	19	44.7 ± 14.3	6 (31.6 %)	2.5 ± 1.3	12.2 ± 3.4	5.0 ± 9.0
Responders	23	46.9 ± 11.4	9 (39.1%)	2.2 ± 1.6	12.7 ± 3.8	3.9 ± 4.1
No Treatment	20	46.2 ± 13.2	10 (50%)	3.0 ± 1.8	13.8 ±3.9	4.80 ± 5.1

Table 10: Demographics for Study 2: Demographics of participants, divided by group, are shown. Unless noted below, numbers represent average values ± SD. Years of education are based on an American system, K-12, with 12 indicating a high school diploma and anything beyond that, either technical or higher-level education. & = Income level is represented by the numbers 1-6 which indicate different yearly income brackets: 1 = <$10,000; 2 = $10,000-25,000; 3 = $25,001-$50,000; 4 = $50,001-$75,001; 5 = $75,001-$100,000; 6 = >$100,000. There were no significant differences between groups in any of the measures shown in this table.

Interviews were well controlled and not influenced by potential confounds

The interviews were audio recorded, transcribed, preprocessed, and analyzed according to the methods described in **Chapter 2**, summarized in **Figure 20b.** The interview script can be found in **Table 2**.

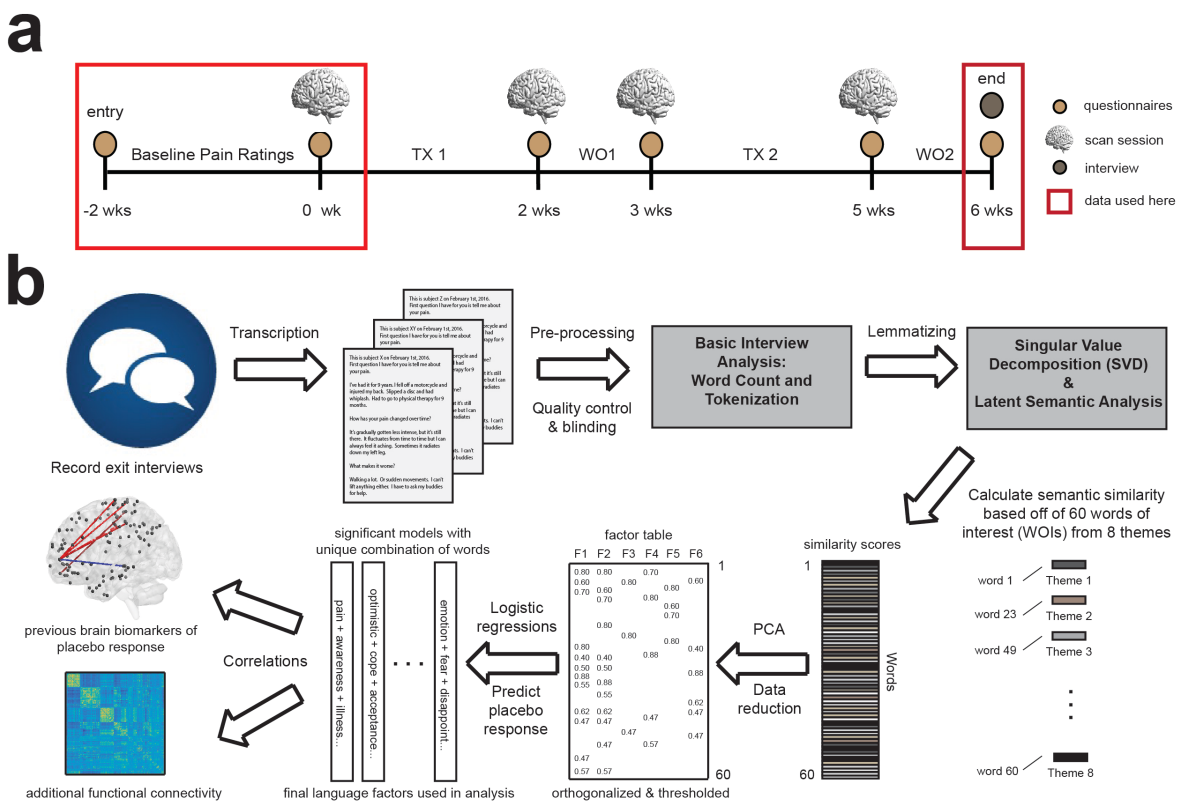


Table 20: Study 2 design and procedures. **a**. Schematic of the study overview is shown. As part of a randomized, partially-blinded clinical trial, participants with chronic low back pain (CBP) participated in 6 visits spread out over 8 weeks, completing a total of 4 brain scans, 2 consecutive treatment and washout periods, 6 batteries of questionnaires, daily pain and mood ratings, and a final exit interview. For the purpose of this paper, only the data from the 1st, 2nd, and 6th visit (marked by red boxes) are used to analyze language properties post-placebo response and correlate them with brain and psychological properties pre-treatment. **b**. Flow diagram of interview preprocessing, quality control, and analysis. A total of 62 CBP patients were interviewed. These interviews were recorded and transcribed into text documents. Following transcription, a subset of transcripts was manually checked for spelling and verbal accuracy as part of the quality control process. Transcripts and corresponding questionnaire and neuroimaging data were then blinded and preprocessed as described in the methods section. A basic word count analysis was performed to verify groups did not differ in key elements of their interviews, and transcripts were tokenized, lemmatized, and converted to numbers corresponding to their numerical position in our data dictionary. Latent Semantic Analysis (LSA) was performed through the utilization of singular value decomposition (SVD) in order to convert the interviews into semantic space and study how patients' language was related to the English lexicon. From the resulting matrices, the average semantic distance between all words in each participant's interview to 60 words of interest (WOIs, **Table S3**) was calculated to create 60 semantic similarity scores for each participant. These scores were averaged within groups and entered into a principal component analysis (PCA) for further data reduction. PCA components were orthogonally rotated and thresholded at loadings of +/- 0.4 to keep only the most significant elements. Remaining words within each component were entered into backward, stepwise logistic regressions (one for each component) predicting placebo responders (**Table S4**). Words that provided significant logistic models (i.e., explained a significant amount of variance in response) were combined via their model equations into final language factors (**Table 4**) that were used as vectors to be correlated with previous markers of placebo propensity, additional self-report measures, and resting state functional connectivity.

The average length of an interview was 27.1 ± 10.3 minutes, with the shortest being 13 minutes and the longest being 66 minutes in length. We conducted a simple word-count analysis as a control investigation to make sure that none of our participants groups differed in the general properties of their interviews, parameters which might influence or bias semantic language differences studied. For each interview, verbosity, vocabulary, and lexical diversity was calculated, averaged within group, and compared between groups. There were no statistically significant differences between groups in any of the 3 measures (**Figure 21a**, averages provided in **Table 11**), indicating that our patients were relatively balanced when it came to the amount of talking that they did and the number of different kinds of words they used.

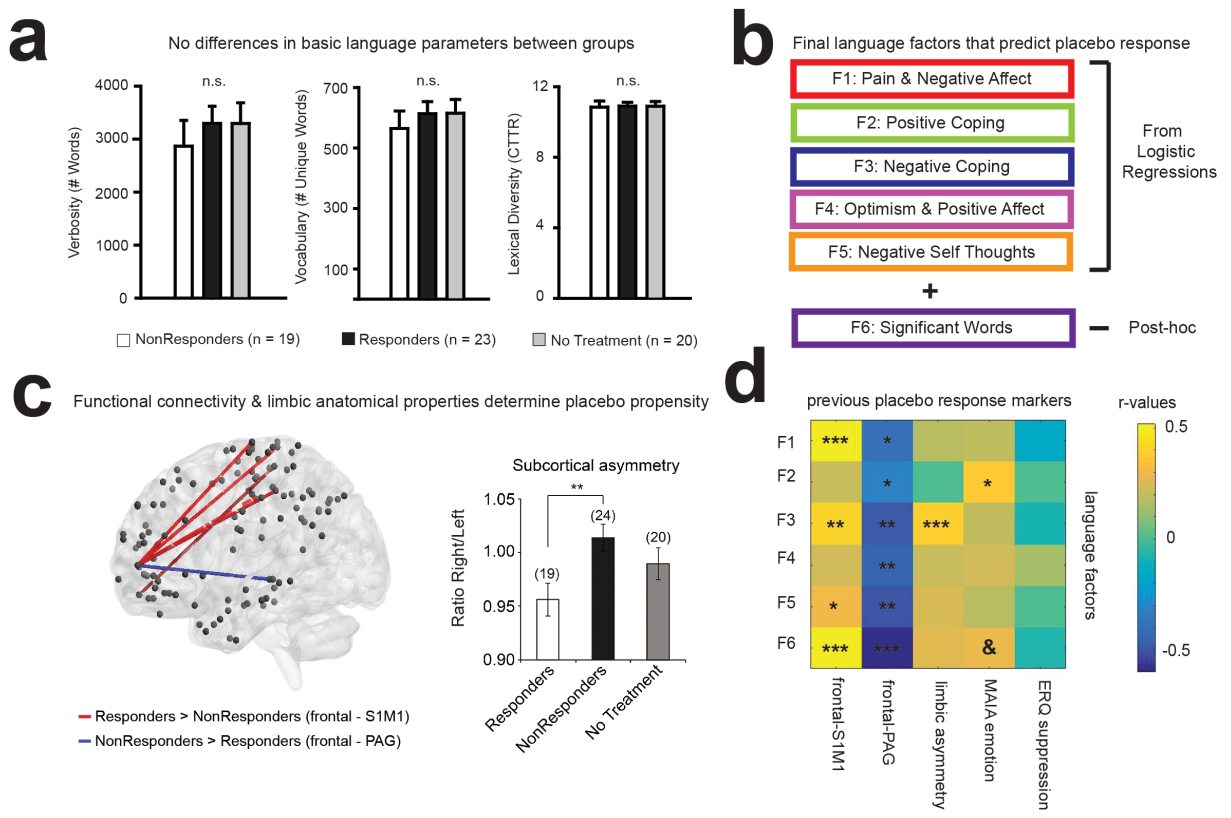


Figure 21: Semantic language properties correlated to previous markers of response. a. Basic word count analyses were used as control measures and revealed no significant differences between groups. Overall verbosity scores, total vocabulary used, and diversity of language was similar for non-responders, responders, and no-treatment (verbosity: one-way ANOVA: $F(2,59) = 0.37$, $p = 0.69$; vocabulary: one-way ANOVA: $F(2, 59) = 0.35$, $p = 0.71$; lexical diversity: one-way ANOVA: $F(2,59) = 0.01$; $p = 0.98$). **b.** Logistic

regressions performed on semantic similarity scores from 60 words resulted in the 5 significant factors (F1-F5) shown here; provided names are descriptive and used to summarize the combination of words within each factor. To increase our odds of finding significant relationships between language semantics and previous psychological and biological variables, we created an additional factor (F6) post-hoc that combined words with semantic differences that were significantly different between responders and non-responders. **c.** Summary of key neuroimaging findings from **Study 1** differentiating responders from non-responders. Left: Placebo responders had increased functional connectivity between their lateral frontal cortex and sensorimotor regions (S1M1, shown in red) compared to non-responders and decreased functional connectivity between lateral frontal areas and the descending pain modulating region of the periaqueductal gray (PAG, shown in blue) compared to non-responders. Right: Volumes of three limbic regions (hippocampus, amygdala, and nucleus accumbens) were summed within hemisphere and divided to create a ratio of subcortical limbic asymmetry. Responders had more leftward limbic asymmetry than non-responders, with individuals who were not treated falling in between both groups. **d.** Language factors from B were correlated with the 3 neuroimaging parameters from C predicting placebo response in addition to questionnaire subscores from 2 questionnaire measures also significantly distinguishing responders from non-responders in **Study 1**. Matrix is colored to show corresponding Pearson correlation coefficients between variables, with asterisks and hashtags indicating level of significance. Additional predictors previously identified are shown in **Table 19.** MAIA = Multidimensional Assessment of Interoceptive Awareness (emotion subscale); ERQ = Emotional Regulation Questionnaire (suppression subscale); $p<0.05$*, $p<0.01$**, $p<0.001$***, # = trending with $0.05<p<0.10$.

It also suggested that the interviews were well controlled regarding the total amount of information in them, meaning that any changes seen in other analyses were not due to differences in these basic parameters. Unsurprisingly, each of these measures were correlated to education and income levels reported by our participants (**Table 12**). However, we did not regress out these effects as neither education nor income differed between the 3 groups (education: one-way ANOVA, $F(2,59) = 0.95$, $p = 0.39$; income: one-way ANOVA, $F(2,58) = 1.28$, $p = 0.28$). Additionally, none of these basic language parameters correlated to pain or mood reported at the day of the scan, although positive mood's relationship was trending such that the better the mood, the more talkative they were (also **Table 12**). These results are negligible as there were no differences between groups in pain (one-way ANOVA: $F(2,59) = 2.39$, $p =0.10$), positive affect (one-way ANOVA: $F(2,59) = 2.01$, $p = 0.14$), or negative affect (one-way ANOVA: $F(2,58) = 0.74$, $p =0.48$) at the time of interview.

Group	Verbosity	Vocabulary	Lexical Diversity
Non-Responders	2866.4 ± 2124.4	564.6 ± 251.3	10.8 ± 1.5
Responders	3297.4 ± 1553.2	613.3 ± 191.4	10.9 ± 1.0
No Treatment	3295.1 ± 1746. 7	615.1 ± 203.7	10.9 ± 1.2

Table 11: Basic word count analysis: Interviews were tokenized according to the methods and a basic word count analysis was performed. There were no differences between groups in verbosity (# of total words), vocabulary (# of unique words), or lexical diversity (vocabulary/sqrt(2*verbosity)), corrected type-to-token ratio), indicating that the interviews were well controlled between participants. Numbers are average values ± SD.

Parameter	Education	Income[&]	NRS (visit 6)	PANAS + (visit 6)	PANAS – (visit 6)
Verbosity	r = 0.29, p = 0.023, n = 62	r = 0.22, p = 0.086, n = 61	r = 0.021, p = 0.873, n = 62	r = 0.23, p = 0.087, n = 62	r = 0.08, p = 0.542, n = 61
Vocabulary	r = 0.37, p = 0.003, n = 62	r= 0.24, p = 0.058, n = 61	r = 0.033, p = 0.797, n = 62	r = 0.24, p = 0.063, n = 62	r = 0.06, p – 0.650, n = 61
Lexical Diversity CTTR	r = 0.40, p = 0.001, n = 62	r = 0.26, p = 0.045, n = 61	r = 0.07, p = 0.566, n = 62	r = 0.22, p = 0.079, n = 62	r = -0.06, p = 0.673, n = 61

Table 12: Correlations of language with demographics, pain, and mood. The 3 basic language parameters were correlated to demographics of interest and pain and mood at the day of the interview. Education, income level, and positive mood were all either significantly related or marginally related (trending) to these language measurements. Pearson correlation coefficients are provided; red indicates significant correlations at $p<0.05$ and blue indicated trending correlations at $0.05<p<0.10$. & = indicates that income was not normally distributed and the r-values provided in this column are actually Spearman rho values and their associated p-values. NRS = Numeric Rating Scale (measuring pain intensity from 0 to 100); PANAS = positive and negative affect scale (+ = positive subscale and - = negative subscale, measuring mood); visit 6 = final study visit (same day as interview).

9 features identified through LSA could not categorize responders from non-responders

We utilized the TASA corpus to construct a dictionary representing a lexicon of the English language weighed by the frequency of each word in the database and truncated at 300 dimensions. We used this dictionary to construct a reduced semantic representation for each participants interview, forming a matrix (Xi) whose elements were the 300 vectors of every word in each interview. We performed SVD to decompose Xi into its 3 component matrices and extracted the 9 features with the strongest singular values to compare between groups. Alone, none of these features were different between responders and non-responders (unpaired ttests were all not significant). We trained a support vector machine (SVM) classifier on the combination of all 9 features using leave-one-out-cross-validation (LOOCV) to measure how well it could differentiate the two groups. Average accuracy over 1000 iterations was quite poor at a mere 39%. When limiting the classification problem to only the top 3 features instead (in order to minimize potential noise), average accuracy went up to 55%, but this was still at chance level, suggesting that basic LSA failed to differentiate responders from non-responders.

Semantic similarity of interviews to 60 words differentiated responders from non-responders

8 themes hypothesized to be important for placebo response and/or chronic pain were created and functioned as search lights in the literature looking for words of interest; 60 words from ethnographies, journals, and questionnaires were chosen that fit within these themes (**Table 13**).

Initial Theme & Color	Word	Reason for Inclusion or Examples of a Word Could Mean	Reference
1. Mindfulness	understand	part of illness representation is what participants understand about their pain from a medical perspective	[284, 285]
	awareness	interoceptive awareness important for clinical placebo analgesia; MAIA questionnaire	[236, 237]
	notice	MAIA notice subscale	[236, 237]
2. General mindset	stress	related to PCS and PASS questionnaires, some subscales of which are important for clinical placebo analgesia	[226, 227, 230, 231]
	worry	Part of MAIA subscale wording…	[236, 237]
	anxiety	related to PCS and PASS questionnaires, some subscales of which are important for clinical placebo analgesia	[284, 285] [249]
	acceptance	related to CPAQ questionnaire	[227] [286]
	content	related to CPAQ questionnaire	[227]
	control	Do they feel like they have control of their pain or that their pain controls them?	[170]
3. Expectations	afraid	Related to PCS; are they afraid they won't get better or that the pain will never end?	[230]
	distrust	Based on previous experiences or beliefs, do they distrust the medical system, physicians, or science?	[284, 285]
	doubt	Do they experience a lot of doubt – either in their treatment or in their pain?	[284, 285]
	uncertain	Are they uncertain about the future with respect to their pain or do they think it will either diminish, stay the same, or get worse? Are they uncertain about different kinds of treatment?	[284, 285]
	disappoint	Where they disappointed with previous medical treatment or physician interactions?	[284, 285]
	surprise	Are they easily surprised by how well a treatment does or does not work for them? Did the pain take them by surprise?	[284, 285]
	believe	A word that has been associated with placebo responses in the past; belief systems include expectations of treatment, belief in what caused their pain, belief in reliability of science, beliefs about how well their own actions or medicine can control their pain	[99, 284, 285]

	trust	Given beliefs and understanding, do participants trust in medicine in science? Do they trust the research staff's knowledge and in the process of discovery?	[284, 285]
	hope	Do they have hope that their pain will go away, that they will eventually find a medication that works, or that their participation in the study will be beneficial for others?	[284, 285, 287, 288]
	open	The only of the Big 5 personality scores that was different between responders and non-responders was openness from the IPIP NEO FFI.	[244]
	optimistic	A common characteristic of placebo responders in healthy controls but not CBP patients in the LOTR; does language provide a different measure of optimism that relates to placebo propensity?	[287, 288]
4. Body Status and Physical Impact	chronic	Describing the nature of their pain; to what extend do the words used also capture how long-term this experience has been for patients?	[30]
	discomfort	The extent to which their condition gives them discomfort (something slightly less than pain but still abnormal and attention grabbing)	[249]
	force	A possible factor that affects pain intensity; weight-bearing, pressure, etc ; alternatively, do they feel forced by	[170, 284, 285]
	illness	To what extent do they view their pain as a symptom or an illness in and of itself?	[168, 170]
	magnify	PCS subscale that is important for clinical placebo analgesia: magnification	[230]
	pain	Main topic of discussion so in everyone's interviews; how well does this word represent the overriding sensation of being at the core of people's experience	[284, 285]
	sensitive	Pain sensitivity; factor that might affect pain and would relate to PSQ	[232]
	well	The opposite of chronic; how good do people feel despite being in chronic pain (may encompass both physical and mental well-being)?	[227]
	cope	Positive or negative ways that people manage their pain; could include keeping active, seeking support, making adjustments or modifications to their posture or activities, etc…	[228, 284, 285]
	healthy	The opposite of illness; how good to people feel despite being in chronic pain (physical)?	[227]
	recover	Since placebo response involves temporary relief or recovery from pain, do participants' interviews show similarity to this word?	[30]

	resilient	Since placebo response may be trait-like and part of similar mechanisms of resilience, do participants interviews show similarity to this word?	[30]
5. Mood and Emotional Impact	fear	Are they fearful of certain medications or of what is causing the pain?	[27] [244]
	angry	Are they angered by their pain and what it has done to their lives?	[27] [244]
	sad	Are they saddened or depressed by their pain and what it has done to their lives?	[27] [244]
	unhappy	Are they unhappy overall with the state of their pain (more vague than sadness or anger)?	[284, 285] [27] [244]
	emotion	Another common topic of conversation. What is the emotional impact of people's pain…how does pain affect people's emotions and likewise, how do their emotions influence pain? Could be negative or positive. Also related to the MAIA emotion subscale, which was predictive of placebo propensity.	[284, 285]
	happy	Are they generally happy with themselves or their lives? Might be indicative of extraversion, which has been positively correlated with placebo in healthy people.	[244]
	cheerful	Do they have a generally cheerful demeanor? Might be indicative of extraversion, which has been positively correlated with placebo in healthy people.	[244]
	joyful	Are they filled with joy because of something in their lives (context dependent, as opposed to personality dependent, and stronger than happy or cheerful)	[244]
6. Behaviors and Cognition	avoidance	Do they avoid certain activities because of their pain?	[226, 227, 230, 231] [249].
	aversion	Do they have an aversion to certain medications because of their pain? Do they have an aversion to loss? May be related to LAQ.	[83, 246]
	impulsive	Pain patients have been shown to display more impulsive behaviors tied to changes in cognition and emotional processing, which might impact response. Are their interviews related to this word? May be related to LAQ?	[83] [289]
	restless	Are they tired of their pain? Do they feel on edge because they don't have answers?	[27]
7. Societal Impact	burden	Do they feel like their pain is a burden on them or that they are a burden to their families/friends that they rely on?	[249].
	disability	Do they identify as disabled because of their pain or do they acknowledge their pain but don't think it hinders them in anyway?	[168]

	loss	Do they feel that pain has caused them to lose out on social activities or physical activities? Do they feel a loss of self or a loss of body? Do their word resonate with loss avoidance or aversion (something that has been tied to CBP before).	[83, 284, 285]
	limit	Do they feel physically limited or socially limited? Do they feel limited in terms of access to health care?	[284, 285]
	stigma	Do they feel stigmatized because of their pain status (do they feel like people think they are drug seeking or are disabled?)	[171, 249, 290]
	embody	How much is pain embodied for them? Obviously they feel in their body, but has it become deeper than that?	[170]
	identity	Has pain infiltrated their identity and how they see themselves?	[170, 284, 285]
	life	Is pain associated with their life in such a way that it's familiar/like family or is it just a nuisance?	[170, 284, 285]
	empathy	Do they know anyone else going through pain and have empathy for these individuals? Do they have empathy for the doctors and scientists who try to treat or cure pain to little avail? Do they experience empathy from others?	[123, 168]
8. Emotional Control and Strategies	internalize	Related to suppress; do they tend to internalize their suffering and expression of pain or do they tend to share it? Do they view it as internally driven or externally perturbed?	[238]
	ruminate	PCS rumination subscale, which is important for clinical placebo response	[230]
	secret	Due to either cultural upbringing or previous negative experiences talking to people about their pain, how secretive are individuals about their pain experience? To what extent to they share their story or keep it to themselves? To what extent does it remain a secret to the listener even when it is told?	[249]
	suppress	Do they suppress their emotions? Related to ERQ suppression subscale, which is important for clinical placebo response	[238, 249]
	private	Do they tend to keep emotions or expressions of pain private as opposed to sharing them with other people?	[249]
	confront	Do they tend to confront the pain (active coping in the form of acknowledging it or seeking treatment) instead of ignoring it or passively coping with it.	[228]
	express	Do they express their emotions and figure out what to do with them? Related to ERQ reappraise and eACS	[238, 240]

Table 13: 60 words of interest chosen from 8 themes. These themes were based on concepts that might that might relate to the chronic pain experience, placebo response, or both based on previous findings and paper. Note that the themes are relatively arbitrary in that some words could fit into more than one theme and they were not based on linguistic calculations: therefore, they are only for organizational/visual purposes. The words within each theme were chosen from pain ethnographies and narratives, previous studies that involved interviewing pain patients, peer-reviewed journals, and questionnaire measures used in the study.

The semantic similarity between all the words in an interview and each of these 60 words was calculated as described in the methods section and averaged within interview. Thus every participant had an average semantic similarity score for each of the 60 words, representing how close his/her interview was to each of these concepts. These 60 semantic similarity scores were then averaged within groups and compared between responders and non-responders in a series of unpaired t-tests. This identified 11 words that were significantly different between the two groups (shown in **Table 14**).

Word	Average Similarity (NonResponder)	Average Similarity (Responders)	Statistics
Awareness	0.0090 ± 0.0018	0.014 ± 0.0015	t = -2.19 (df = 40), p = 0.03
Emotion	0.0100 ± 0. 0017	0.015 ± 0.0014	t = -2.09 (df = 40), p = 0.04
Stigma	-0.0038 ± 0.0004	- 0.0065 ± 0.0008	t = 2.80 (df = 40), p = 0.01
Well	0.022 ± 0.0022	0.017 ± 0.0012	t = 2.18 (df = 40), p = 0.04
Empathy	-0.0012 ± 0.0009	0.014 ± 0. 0008	t = -2.14 (df = 40), p = 0.04
Fear	0.012 ± 0.0013	0.017 ± 0.0011	t = -2.70 (df = 40), p = 0.01
Identity	-0.0053 ± 0.0009	-0.0020 ± 0.0011	t = -2.21 (df = 40), p = 0.03
Loss	0.012 ± 0.0008	0.014 ± 0.0008	t = -2.22 (df = 40), p = 0.03
Magnify	- 0.0074 ± 0.0008	-0.0054 ± 0.0006	t = -2.15 (df = 40), p = 0.04
Disappoint	0.012 ± 0.0009	0.014 ± 0.0008	t = -2.15 (df = 40), p = 0.04
Force	0.00025 ± 0.0003	-0.00087 ± 0. 0003	t = 2.65 (df = 40), p = 0.01

Table 14: **11 words differed between responders and non-responders**. Semantic similarity values were calculated for 60 words of interest according the methods in the manuscript. Of these 60 words, 11 were significantly different between responders and non-responders. Average similarity scores with SEM are provided for each word for both groups; statistics represent un-paired ttests between the two groups (not corrected for multiple comparisons).

To further reduce the amount of words investigated in this study, we entered all 60 similarity measures into a PCA, thresholding at a component loading of +/- 0.4 to retain the strongest word relationships. This resulted in 6 statistically relevant language components containing subsets of the words of interest. For each of the 6 components, the words within them were used as covariates in

logistic regressions predicting the binary outcome of being a responder or not. These models showed which combinations of word similarities explained a significant amount of variance in placebo response (shown in **Table 15**); those words that did not meet inclusion criteria were discarded.

Factor	Surviving Parameters	Coefficient	p-val	Eliminated Parameters from Component	Model Information
1	stress believe avoidance aversion awareness discomfort illness loss pain sensitive	-176.26 -99.84 -58.73 257.08 318.06 60.74 -77.85 192.72 -56.09 -280.04	0.257 0.427 0.641 0.163 0.082 0.627 0.563 0.226 0.638 0.059	fear angry anxiety emotion sad suppress	n = 42 p = 0.0492 pseudo R^2 = 0.32
2	acceptance awareness cope empathy express limit optimistic identity	-177.41 288.07 -208.07 302.78 -355.57 150.27 -113.57 277.18	0.209 0.072 0.131 0.113 0.120 0.375 0.436 0.064	understand stress anxiety emotion life	n = 42 p = 0.0485 pseudo R^2 = 0.27
3	content confront disability impulsive internalize magnify secret unhappy	-93.42 -83.84 72.46 -271.35 -87.01 444.70 -126.40 245.28	0.472 0.389 0.503 0.109 0.530 0.021 0.190 0.111	uncertain worry	n = 42 p = 0.0486 pseudo R^2 = 0.27
4	afraid trust healthy hope	213.24 104.34 116.02 -90.50	0.018 0.245 0.183 0.371	happy open resilient restless	n = 42 p = 0.0428 pseudo R^2 = 0.17
5	burden embody stigma disappoint private	103.51 -55.10 -384.10 124.03 -123.01	0.395 0.611 0.043 0.249 0.370	confront doubt surprise	n = 42 p = 0.0292 pseudo R^2 = 0.22
6 significant	awareness emotion stigma well empathy fear identity loss magnify disappoint force	442.04 -65.53 -1630.87 -110.72 -296.12 479.09 18.11 288.05 832.91 113.16 -2302.02	0.465 0.888 0.133 0.565 0.616 0.434 0.978 0.601 0.236 0.738 0.062	n/a	n = 42 p <0.0001 pseudo R^2 = 0.80

Table 15: Creation of language factors. Semantic similarity scores from 60 words of interest were entered into a PCA, identifying 6 potential factors. After thresholding, words within each factor were entered into a backward, stepwise logistic regression explaining placebo response. Once a model accurately predicted response at $p<0.05$, removal of words was stopped. Those words remaining in each model were combined with their associated model equation to form language factors, which were the primary regressors in our analyses. 5 out of the 6 components resulted in significant models, creating 5 language factors (the 4th non-significant factor consisted of the words "control", 'hope", "joyful", "ruminate", "unhappy", and "life". An additional factor was created post-hoc from the 11 words significantly differentiating responders from non-responders (**Table 14**) and is referred to here as factor 6 (significant words).

5 of the 6 components provided significant models, and the equations associated with each of the final models were used to create 5 new language factors that represented the unique combination of words that (a) were semantically related to the interviews and (b) differentiated responders from non-responders. The factors were created from each model's final unique equation. Each factor was assigned a name associated with the words that was purely theoretical and was provided to help interpret the factors' relationship with the placebo effect and additional self-report and brain measurements. In summary, factor 1 (f1) had words primarily associated with "pain and negative emotions", f2 with "positive coping mechanisms", f3 with "negative coping mechanisms", f4 with "optimism and positive emotions", and f5 with "self-referential thoughts", many of which were negative (**Figure 21b,** summarized in **Table 16**). Out of the original 60, 34 words were kept in at least one of these factors; **Figure 22a-b** shows how the 34 words from each factor were related to one another in semantic space.

Language Factors	F1: Pain & Negative Affect	F2: Positive Coping	F3: Negative Coping	F4: Optimism & Positive Affect	F5: Negative Self Thoughts	N.S. no name	F6: Significant Words
Word	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Post-hoc
Understand		0.6447					
Stress	0.4669	0.5489					
Believe	0.6331						
Well							2.18 (0.035)
Acceptance		0.7848					
Afraid				-0.4795			
Fear	0.8110						-2.70 (0.010)
Angry	0.6356						
Anxiety	0.6961	0.5300					
Avoidance	0.6770						
Aversion	0.9029						
Awareness	0.6046	0.7471					-2.19 (0.034)
Burden					0.6440		
Chronic							
Cheerful							
Content			0.7933				
Control						0.4184	
Cope		0.8300					
Confront			0.6689		0.6025		
Disability			0.5840				
Discomfort	0.8881						
Distrust							
Trust				0.4526			
Doubt					0.6317		
Embody					0.4676		
Emotion	0.7958	0.5086					-2.09 (0.043)
Empathy		0.7245					-2.14 (0.039)
Express		0.8812					
Force							2.65 (0.012)
Happy				0.6285			
Healthy				0.6819			
Hope				0.4427		0.6671	
Illness	0.8260						
Impulsive			0.5419				
Internalize			0.8461				
Joyful						-0.5598	
Loss	0.7491						-2.22 (0.032)
Limit		0.7892					
Magnify			0.4221				-2.15 (0.038)
Notice							
Open				0.6211			
Optimistic		0.6982					
Pain	0.8991						
Recover							
Resilient				0.4493			
Restless				-0.6954			
Ruminate						-0.8090	
Sad	0.6005						
Secret			-0.6414				
Sensitive	0.7419						
Stigma					0.5460		2.80 (0.008)
Suppress	0.6369						
Uncertain			0.7037				
Unhappy			0.4205				
Worry			-0.4611			0.5573	
Identity		0.7759					-2.21 (0.033)
Disappoint					-0.6450		-2.15 (0.037)
Life		0.5642				-0.4431	
Private					0.7354		
Surprise					-0.6011		

Table 16: **Final language factor table.** 60 words of interest from 8 themes (colored accordingly on the left) were entered into a PCA. This identified as many as 6 significant components; the table was rotated and thresholded at loadings of ± 0.4 to identify the strongest words in each component (surviving words shown with their loading values). These words were then entered into 6 backwards stepwise logistic models until each model explained significant variance in placebo response ($p<0.05$, **Table 15**). 5 of the original 6 components provided words with significant models (the 6th factor did not, and was discarded – in gray). The remaining words (colored in the table) were combined using each model equation to form new "language factors". These factors were given names based on the words within them and are shown at the top. In addition, we created a post-hoc factor that combined the 11 significant words from **Table 14** into a 6th factor, which we refer to as the "significant words" factor (in purple); un-paired t-test results shown for this post-hoc factor: signed t-value (p-value).

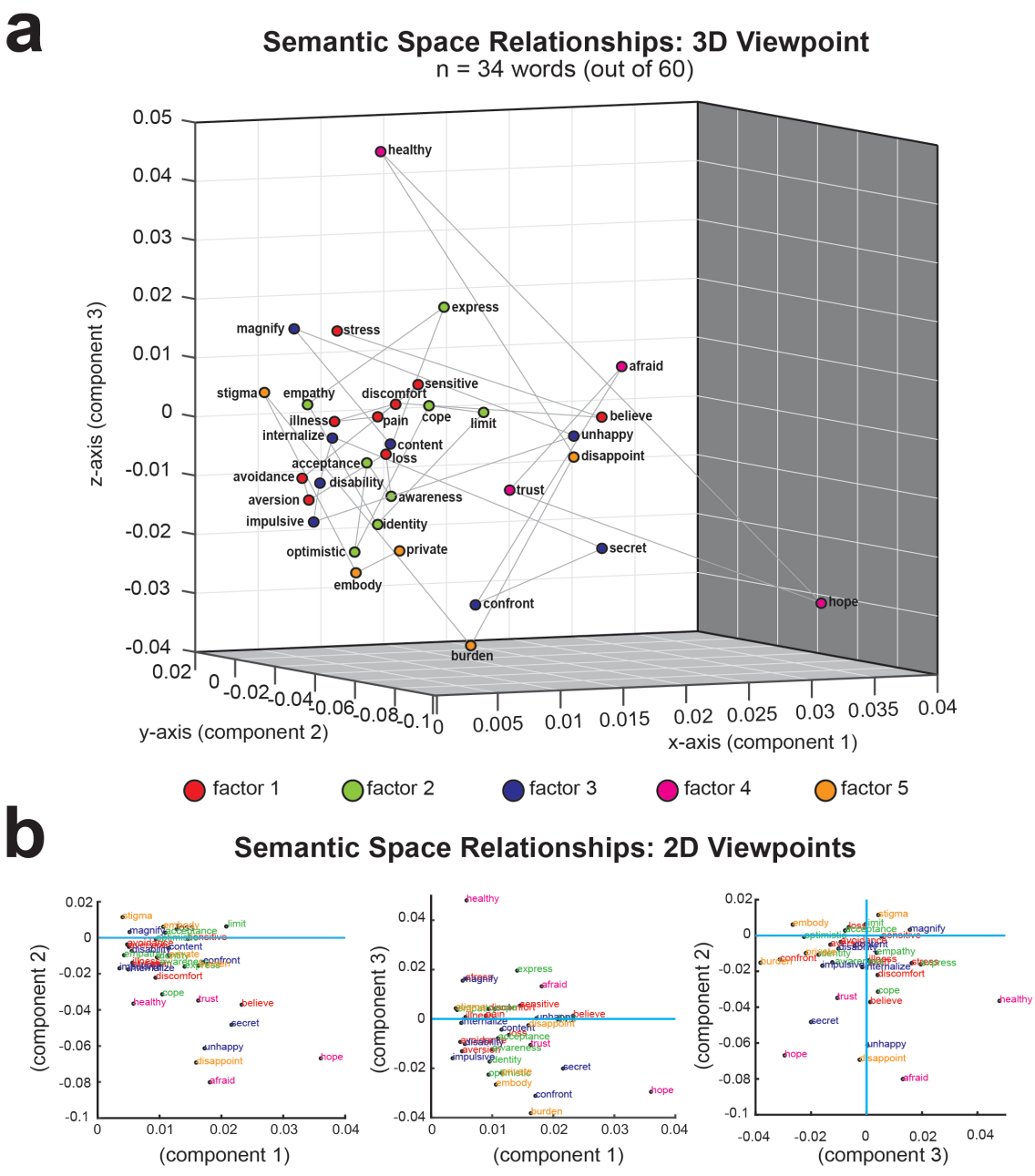


Figure 22: Semantic relationships between 34 words from 5 factors. A. Depiction of the final set of words plotted in semantic space across the first 3 features (components) of the TASA-based lexicon is shown. The closer a word is to another, the closer it is in meaning and context to that word or other words in a cluster, according to the lexical feature space defined here. Note that the coordinates of the words are not determined by our interviews. Word dots are colored according to the factor to; here, color is determined by our interviews, as they represent the average semantic distance to the interviews, as well as the results from the PCA which divided the words into factors. One can appreciate, for example, that the words "illness", "pain", and "discomfort" (in red) are all close to each other in the lexicon and cluster together based on our interview data (i.e., they are all in the same factor). In contrast, the words "stigma" (orange) and "empathy" (green) are semantically close in the lexicon but have clustered into separate factors according to our data. Lines are non-informative; they were kept for better special visualization. **B.** The same data as in (A), but depicted in 3 pair-wise 2D plots. Note that the last plot (right) is the most-commonly used to depict these semantic relationships. Blue lines indicate 0 axis through origin; colors of words indicate factor. Words from post-hoc factor 6 are not shown.

An additional 6th factor was created outside of the PCA by combining the 11 significant words together in a logistic regression. The equation created from the un-thresholded regression was also used as a regressor for future analyses, since our main aim was to see if we could increase the semantic signal of the model to increase our chances of finding brain regions associated with the words; this factor is referred to as the "significant words" factor (also shown in **Table 16**). We also entered the 11 words into a backwards stepwise regression predicting placebo propensity, thresholded at $p < 0.05$, to identify the words that were significantly predictive of response propensity (i.e., both the model and the regressors had p-values<0.05). A combination of 4 words – "awareness", "force", "stigma", and "magnify" – survived this latter regression and were shown to significantly contribute to almost 70% of the variance in placebo response (pseudo $r^2 = 0.68$, $p < 0.0001$); this factor (6&) is shown in **Table 17.** We entered these 4 surviving words into the support vector for pattern analysis. When this model was cross-validated, it was successful, achieving an average of 84% accuracy in dissociating responders from non-responders. Due to the large amount of variance explained and the relatively high level of accuracy obtained, this model was also related to additional aspects of the data, explained below, and is referred to as "factor 6&". These results indicate that placebo response in chronic pain patients can be captured by interviews and that placebo responders differ from non-responders along specific semantic properties.

Surviving Parameter(s)	Coefficient (SEM)	p-value	95% C.I.	Eliminated Variables (p > 0.05)
Entered in all 11 words with significant differences in semantic similarity values				
awareness	413.72 (180.37)	0.022	[60.2, 767.2]	identity, emotion, disappoint, empathy, loss, well, fear
magnify	705.20 (315.45)	0.025	[86.9, 1323.5]	
stigma	-918.61 (422.04)	0.030	[-1745.8, -91.42]	
force	-1862.609 (729.16)	0.011	[-3291.7, -433.5]	
N = 42; Pseudo R^2 = 0.68; LR chi^2= 39.61; log likelihood = -9.12; p < 0.0000				
f(x) = -5.11 + 413.7*awareness + 705.2*magnify - 918.6*stigma – 1862.6*force				

Table 17: A combination of 4 words significantly explained the placebo response. A backwards, step-wise logistic regression was utilized to study how much variance in response propensity was explained by semantic similarity to words of interest. 11 words whose average semantic similarity scores were shown to be different between responders and non-responders (posthoc factor 6**)** were entered as independent regressors predicting the binary outcome of responder (1) versus non-responder (0). 4 of the words survived this regression at p<0.05 and explained almost 70% of the variance in response.

Data-driven linguistic analyses dissociated responders from non-responders but failed validation

As an alternative approach to selecting words of interest apriori, we tested whether semantic similarities in general (over the entire lexicon) could be identified that would classify response group. We first created a matrix that represented the average semantic similarity of each person's interview to every word in the dictionary; this matrix was then broken up into its 3 component matrices using SVD. There was a significant difference in the loading on the second component between the responder and non-responder conditions ($t_{(40)}$ = -2.15; p = 0.038), explaining approximately 8% of the variance in response (logistic regression, pseudo R^2 = 0.078, p = 0.032). However, in the machine-learning analysis, LOOCV yielded only an average accuracy of 53%. Like the initial LSA result, this finding may indicate that the interview data, despite preprocessing and controlling for verbosity, may still have significant amounts of noise in the signal, a prospect which we will address in the discussion.

Language factors correlated to neurological and psychological predictors of placebo response

Because the interview took place at the end of the study post-treatment, we could not definitively say whether the semantic properties seen were predictive of placebo propensity or a result of placebo response. To try to answer this question, the 6 language factors were correlated to the main findings from a previous report that identified neuroanatomical, neurophysiological, and psychological predictors of placebo propensity in the same cohort of individuals (**Figure 21c-d**): this included resting state functional connectivity between frontal regions and either sensorimotor (S1M1) regions or the PAG; limbic asymmetry (the ratio of left and right gray matter volumes of the amygdala, hippocampus, and nucleus accumbens); and questionnaire subscales regarding awareness and regulation of emotion (MAIA emotion and ERQ suppress, respectively). With the exception of the ERQ, all other identified predictors correlated with at least one of the 6 factors, many of them with more than one. Additional parameters were also found to be predictive of placebo response in different models of propensity, including grey matter density and cortical thickness in S1M1 regions, as well as 3 more questionnaire subscales that also dealt with emotion and thought processing (2 from the Pain Catastrophizing Scale, PCS – rumination and helplessness subscales - and another from MAIA, the notworry subscale). For this subset of data, only cortical thickness and PCS helplessness showed any relationship to the language factors. All correlations can be found in **Table 18.** These results suggest that the semantic language properties are related to both personality and neurobiology. Moreover they also indicate that these properties might be predictive of or at least contribute to response, as they correlated well with both biomarkers and psychological traits of propensity.

F	Frontal-S1M1	Frontal-PAG	Limbic Asym.	MAIA Emotion	ERQ Sup.	GM Density	Cortical Thick.	PCS rum.	PCS help.	MAIA notwor.
1	r=0.42 p=0.0004 n=42	r=-0.37 p=0.02 n=42	r=0.18 p=0.26 n=42	r=0.20 p=0.20 n=42	r=-0.14 p=0.36 n=42	r=-0.13 p=0.40 n=42	r=-0.26 p=0.11 n=40	r=-0.008 p=0.96 n=42	r=-0.14 p=0.38 n=42	r=-0.25 p=0.12 n=42
2	r=0.21 p =0.18 n=42	r=-0.31 p=0.046 n=42	r=0.007 p=0.97 n=42	r=0.37 p=0.01 n=42	r=0.0008 p=0.996 n=42	r=-0.23 p=0.14 n=42	r=-0.46 p=0.003 n=40	r=-0.056 p=0.73 n=42	r=-0.207 p=0.19 n=42	r=0.10 p=0.514 n=42
3	r=0.40 p=0.009 n=42	r=-0.49 p=0.001 n=42	r=0.40 p=0.009 n=42	r=0.19 p=0.22 n=42	r=- 0.09 p=0.56 n=42	r=- 0.19 p=0.32 n=42	r=-0.25 p=0.12 n=40	r=- 0.089 p=0.57 n=42	r= 0.208 p=0.19 n=42	r=-0.13 p=0.411 n=42
4	r=0.22 p=0.15 n=42	r=-0.44, p=0.003 n=42	r=0.21 p=0.18 n=42	r=0.24 p=0.13 n=42	r=0.15 p=0.36 n=42	r=0.09 p=0.58 n=42	r=-0.50 p=0.001 n=40	r=0.059 p=0.71 n=42	r=0.036 p=0.82 n=42	r=-0.13 p=0.421 n=42
5	r=0.31 p=0.049 n=42	r=-0.45 p=0.003 n=42	r=0.25 p=0.11 n=42	r=0.19 p=0.24 n =42	r=0.01 p=0.94 n=42	r=-0.17 p=0.27 n=42	r=-0.26 p=0.10 n=40	r=0.112 p=0.48 n=42	r=-0.026 p=0.87 n=42	r=-0.18 p=0.254 n=42
6	r=0.52 p=0.0005 n=42	r=-0.60 p<0.000 n=42	r=0.27 p=0.09 n=42	r=0.29 p=0.06 n=42	r=-0.07 p=0.67 n=42	r=-0.20 p=0.19 n=42	r=-0.46 p=0.003 n=40	r=-0.055 p=0.73 n=42	r=-0.269 p=0.085 n=42	r=-0.08 p=0.60 n=42
6 &	r=0.50 p=0.0008 n=42	r=-0.56 p<0.000 n =42	r=0.15 p=0.35 n =42	r=0.28 p=0.07 n=42	r=-0.08 p=0.62 n =42	r=-0.20 p=0.20 n=42	r=-0.45 p=0.003 n=40	r=-0.127 p=0.42 n =42	r=-0.272 p=0.081 n=42	r=0.006 p=0.97 n=42

Table 18: Correlation of language factors with previous indicators of placebo propensity. Pairwise correlations between the language factors and 10 of the key neuroimaging and psychological parameters shown to predict or contribute to placebo response in a previous study. Correlation coefficients, p-values, and sample sizes are provided for the treatment group only. Red indicates significant ($p<0.05$) correlations and blue indicates trending correlations ($0.05<p<0.10$). Factors 1-5 were derived from a PCA analysis; factor 6 was built post-hoc from words that significantly differed between responders and non-responders; and 6& = factor including only the surviving (significant) 4 words from the factor 6 regression. (Abbreviations: S1M1 = sensory motor; PAG = periaqueductal gray; Asym = asymmetry; Sup = suppression; GM = gray matter; Thick = thickness; rum = rumination; help = helplessness; notwor = notworry)

Language factors also correlated to additional psychosocial parameters

We also wanted to see if the language factors were related to additional self-report measures that did not explain a significant amount of variance in placebo response on their own. All factors were correlated to the questionnaires measuring personality traits and psychological states at either visit 1(screening visit) or visit 6 (time of interview). The resulting Pearson correlation coefficients and statistics are shown in **Table 19**. Interestingly, almost all factors showed some significant relationship to the "openness", one the Big 5 personality traits; this trait was marginally different between responders (average: 46.4 ± 1.45 SEM) and non-responders (average: 41.8 ± 2.02 SEM; unpaired t-test, $t_{(40)} = -1.90$, trending at $p = 0.06$) Additionally, factor 2 (whose word similarity scores corresponded, in general, to positive coping mechanisms) was related to many of the psychological parameters that participants filled out, including optimism (Life Orientation Test, Revised – LOTR) and attributions of health control

(Multidimensional Health Locus of Control – MHLC), among other measures. These findings also point to the ability of language to pick up on more subtle aspects of different personalities or mindsets, some of which may contribute to placebo response to smaller extents or in more complicated ways. Moreover, given that the majority of these questionnaires occurred prior to the interview indicates that the semantic language properties studied here might be trait-like.

Factor	1	2	3	4	5	6	6&
Measures and Stats (V1)	CPAQ-activity (0.27#) openness (0.42**)	LOTR (0.43**) LAQ (-0.27#) FFM-observe (0.30#) FFM – describe (0.29#) FFM-aware (0.30#) MAIA-selfregulation (0.29#) agreeable (0.26#) extraversion (0.28#) openness (0.36*)	CPAQ-painwillingness (0.31*)	openness (0.32*)	conscientious (-0.36*)	openness (0.37*)	openness (0.34*) FFM-observe (0.26#)
Measures and Stats (V6)	none	MHLC-chance (-0.33*) MHLC-others (0.27#) MHLC-doctors (0.32*)	none	none	none	none	none

Table 19: Correlation of language factors with additional psychosocial measures. Pairwise correlations between language factors and other psychological and personality-related variables for the treatment group. Language factors were correlated to additional questionnaire measures collected either before treatment randomization (V1, visit 1) or at the same time as the interview (V6, visit 6). For each factor, all questionnaire scales or subscales that were significantly or marginally correlated are shown (r-values in brackets, followed by asterisks indicating level of significance). Red indicates significant

($p<0.05$) correlations and blue indicates trending correlations ($0.05<p<0.10$). * = $p<0.05$; ** = $p<0.01$; # = trending. Factors 1-5 were derived from a PCA analysis; factor 6 was built post-hoc from words that significantly differed between responders and non-responders; and 6& = factor including only the surviving (significant) 4 words from the factor 6 regression. For all calculations, N = 42 (except MHLC, N=41). (Abbreviations: CPAQ = Chronic Pain Acceptance Questionnaire; LOTR = Life Orientation Test, Revised; LAQ = Loss Aversion Questionnaire; FFM –Five Facets of Mindfulness; MHLC = Multidimensional Health Locus of Control (form C)); extraversion, openness, agreeable, and conscientious were all subscales of IPIP NEO-FFI (measure the Big 5 personality dimensions).

Language factors identify functional brain connections associated with placebo response

Since the language factors were associated with many of the already identified predictive markers of placebo response, we were interested in investigating the extent to which they could identify neural signatures of placebo propensity on their own. To do this, we constructed resting state networks from 272 ROIs in 8 different communities spread over the entire brain (**Figure 23a**), and correlated the 6 language factures to each of the connections of the weighted network. The resulting correlations were thresholded at a significance level of $p<0.001$ to identify the strongest links that were most related to language (**Figure 23b**, brains). These functional connections represent the networks important for the semantic relationships underlying the words in each factor. Surviving connections were grouped according to which community they belonged, and a within- and between-community differentiation of these links was calculated (**Figure 23b,** node diagrams). For all networks, there was very little within-community links related to language factors; the majority of connections involved communication between communities.

To detect if any of these language-related functional connections also differentiated propensity, we extracted all connections, averaged them within group, and compared between placebo responders and non-responders. Those that were significantly different between the two groups were corrected for multiple comparisons using FDR. The remaining functional connections were one-way ANOVAs, Bonferroni corrected, to compare between all 3 groups. Brains in **Figure 23c** show the final functional connections that differentiated responders from non-responders in each language factor; red indicates connections that were higher for responders than non-responders, whereas blue indicates connections that were greater in non-responders than responders. Bar graphs in **Figure 23c** show the number of connections according to this differentiation. **Table 20** shows the final nodes and links for each factor, as

well as the associated statistics (node and module ids can be found in **Appendix II**. These results show that not only were language factors able to discover novel functional connections related to placebo response propensity (prior to the interview), but also identify some of the same connections that were previously found to differentiate groups (such as frontal to PAG and frontal to S1M1).

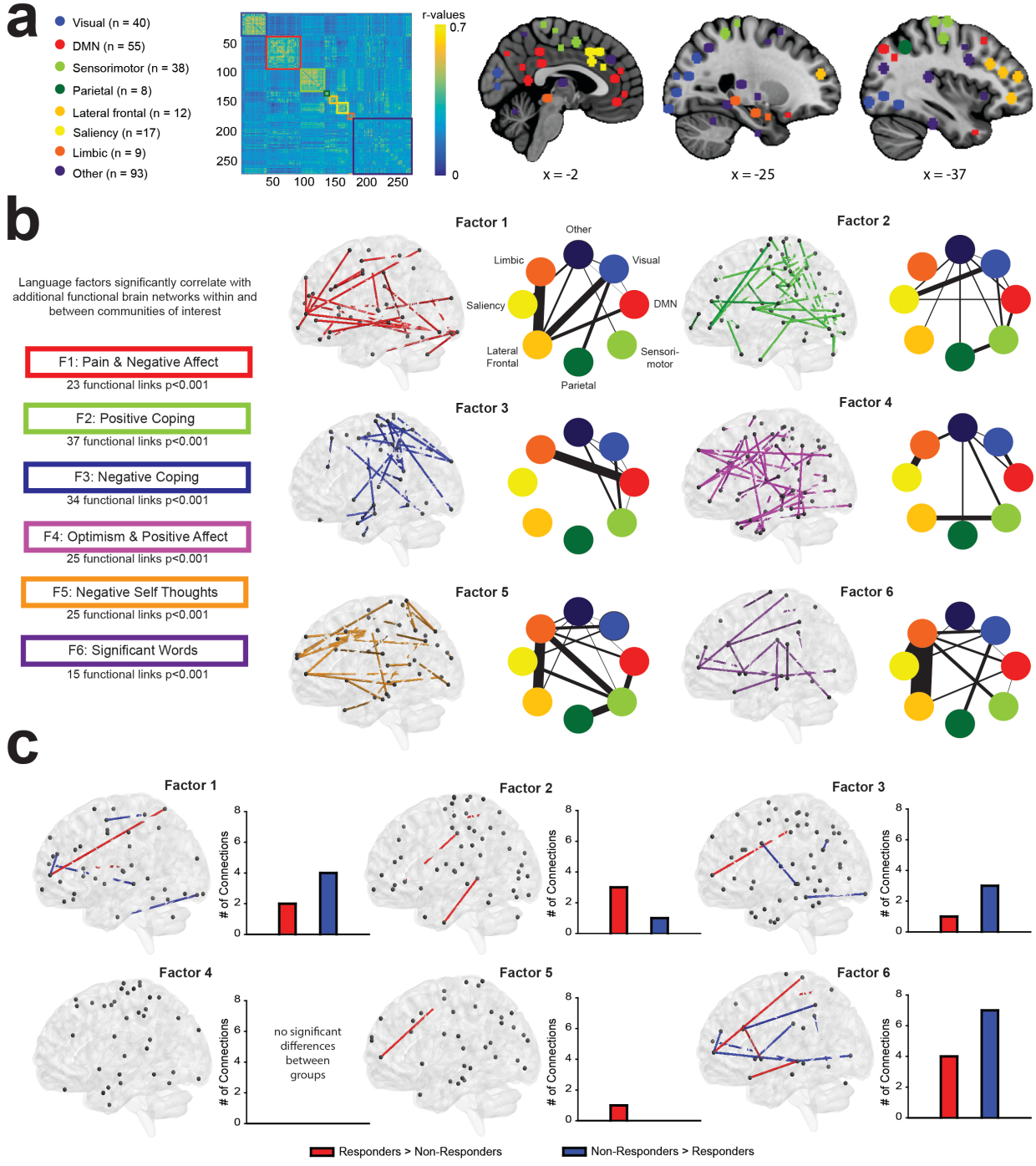


Figure 23: Language factors are sufficient in identifying functional connections that differentiate placebo responders from non-responders. **a.** A whole brain connectivity analysis was used to investigate whether semantic language properties could stand in as surrogates for brain biomarkers of placebo propensity collected prior to randomization and treatment commencement. Shown here are the 8 functional communities of interest made up of 272 total ROIs spread across the entire brain (number indicates # of ROIs within each community) and their associated connectivity matrix. Each r-value of this matrix represents the weighted strength of the functional connection between each region with one another, and colored squares are provided to show the within module connectivity for each of the 8 communities. Circles projected on the brain masks show examples of spherical ROIs from each community. **b.** The six language factors from **Figure 21b** were correlated to each of the 264X264 functional connections and thresholded at $p<0.001$ to identify the most significant functional connections. Left panel shows each of the factors with the number of associated functional links that survived this threshold. Right panel shows brain masks with functional links colored according to which factor with which they significantly correlate (e.g., red lines are functional links that correlate with F1 and pink lines are functional links that correlate significantly with F4, etc). Next to each brain, a circle graph is depicted to display the within- and between-community connectivity for each of the surviving links. Colored circles represent each of the 8 communities of interest, lines around the circles represent the percentage of total possible connections that are within that community, and lines connecting circles represent the percentage of total possible connections that are between those two communities. Line diameters are weighed according to their percentage, with thicker lines indicating a larger % of functional connections, thinner lines a smaller % of function connections, and no line indicating no connections within or between a community or set of communities. **c**. The functional links shown in B were compared between placebo responders and non-responders to see if they differentiated the two groups. Those significant connections were then corrected for multiple comparisons using false-discovery rate (FDR). Shown on each brain are the FDR corrected functional links that significantly differ between responders and non-responders. Red links indicate that responders had higher functional connectivity between these regions than non-responders, whereas blue links indicate that responders had decreased functional connectivity between these regions (or, inversely, that non-responders had increased connectivity between these regions than responders). Bar graphs show number of surviving links associated with increased and decreased connectivity; all but one factor (F4) showed at least one significant link. Semantic language properties were able to identify differences in functional connectivity between responders and non-responders not previously reported, in addition to detecting some of the connections that were already discovered to be involved in response (such as those between lateral frontal regions, S1M1, and PAG), indicating that language parameters can be used to capture placebo response propensity and can result in findings that are physiologically valid.

Factor	Modules Connectivity	Nodes	q-values	Greater in R or NR	oneway ANOVA	noTX stats (Bonferonni)
1 Pain and Negative Affect	DMN with frontal	115, 197	0.006	NR	F(2,59) = 8.03, p=0.008	n.s. (middle)
	other with other	61, 207	0.009	R	F(2,59) = 5.02, p=0.0097	n.s. (middle)
	frontal with limbic (PAG)	218, 226	0.005	NR	F(2,59) = 6.55, p=0.0027	n.s. (middle)
	visual with other	141, 247	0.01	NR	F(2,59) = 4.37, p = 0.017	n.s. (middle)
	sensorimotor with other	25, 261	0.01	NR	F(2,59) = 4.41, p = 0.016	n.s. (middle)
	frontal with other	197, 263	0.009	R	F(2,59) = 5.40, p = 0.007	n.s. (middle)

2 Positive Coping	sensorimotor with sensorimotor	28,49	0.02	NR	F(2,49) = 5.35, p = 0.0073	n.s. (middle)
	sensorimotor with other	26, 210	0.01	R	F(2,59) = 7.39, p = 0.0014	# with NR (middle)
	other with DMN	77, 250	0.048	R	F(2,59) = 3.75, p = 0.0293	n.s. (middle)
	DMN with other	97, 251	0.02	R	F(2,59) = 5.98, p = 0.0043	* with R (middle)
3 Negative Coping	DMN with DMN	88, 89	0.04	NR	F(2,59) = 4.34, p = 0.174	# with NR (middle)
	other with other	2, 126	0.02	NR	F(2,59) = 5.04, p = 0.0095	* with NR (middle)
	other with limbic (PAG)	187,226	0.02	NR	F(2,59) = 6.34, p = 0.0032	** with NR (smallest)
	frontal with other	197,251	0.02	R	F(2,59) = 7.53, p = 0.0012	* with R (middle)
4 Optimism and Positive Affect	no connections survived	-	-	-	-	-
5 Negative self thoughts	frontal with sensorimotor	29, 197	0.008	R	F(2,59) = 6.52, p = 0.0028	n.s. (middle)
6 Significant Words	saliency with other	58,175	0.03	R	F(2,59) = 7.84, p = 0.001	** with NR (highest)
	sensorimotor with frontal	18, 197	0.002	R	F(2,59) = 10.62, p = 0.0001	*** with NR (highest)
	DMN with frontal	115, 197	0.002	NR	F(2,59) = 8.03, p = 0.00-8	n.s. (middle)
	visual with parietal	150, 204	0.03	NR	F(2,59) = 5.05, p = 0.0095	* with R (middle)
	DMN with limbic (PAG)	132, 226	0.01	R	F(2,59) = 4.16, p = 0.02	n.s. (middle)
	visual with limbic (PAG)	173, 226	0.03	NR	F(2,59) = 3.28, p = 0.0447	n.s. (middle)
	frontal with limbic (PAG)	197, 226	0.002	NR	F(2,59) = 10.91, p = 0.0001	** with R (middle)
	frontal with limbic (PAG)	218, 226	0.002	NR	F(2,59) = 6.55, p = 0.0027	n.s. (middle)
	DMN with other	97, 151	0.005	R	F(2,59) = 5.98, p = 0.0043	* with R (middle)
	other with other	212, 259	0.01	NR	F(2,59) = 4.65, p = 0.0133	n.s. (middle)
	DMN with limbic (NAc)	221, 265	0.03	NR	F(2,59) = 3.95, p = 0.0243	# with NR (middle)

Table 20: Summary of functional connections identified by language that differentiated groups. Functional connections identified from language factors were compared between groups. Those that survived correction for multiple comparison with FDR (brains in **Figure 23c**) are shown here, ordered by their original language factor. Also shown are the between and within module assignments for each

connection, their node identity (see **Appendix II** for coordinates), their FDR adjusted p-values (q-values) between connections comparing responders and nonresponders, indication of strength between 2 placebo groups (NR = responders, R = responders, bar graphs in **Figure 23c**), and group statistics for all 3 groups (one-way ANOVA). The last column represents the post-hoc Bonferroni statistics to see how the no treatment (noTX) group compares to NR and R, which are significantly different. Here, the significant differences are indicated, along with overall placement of the group (e.g., if middle, their values were between NR and R); significance: * = $p<0.05$; ** = $p<0.01$, *** = $p<0.001$; # = $0.05<p<0.01$ trending; n.s. = not significant.

Discussion

In this study, we employed semantic language analyses to investigate whether patient narratives can be used to dissociate placebo responders from non-responders and identify psychological and neurobiological parameters associated with response propensity. We utilized the setting of an RCT investigating placebo response in patients with CBP that captured both neuroimaging and personality data pre- and post-treatment. At the end of this trial, patients completed an exit interview which was used to extract language properties for various analyses; importantly these interviews were controlled for content and length, and did not differ between groups in verbosity, vocabulary, or lexical diversity measures. A basic latent semantic analysis (LSA) failed to classify responders from non-responders; however, a more specific analysis that calculated semantic similarity between interviews with 60 words of interest identified 11 words that significantly differed between groups. Of these, 4 words predicted propensity, explaining 68% of the variance in response with high accuracy (84%). Using PCA and logistic regressions, these 60 words were clustered, thresholded, and reduced to 6 significant language factors. The majority of these factors significantly correlated with previously identified biomarkers of placebo propensity, in addition to previously discovered personality traits. Moreover, these factors identified novel resting state functional connections associated with the meaning of participants' language, as well as discovered additional functional links that differentiated responders from non-responders pre-treatment. Finally, language factors also were able to identify the same brain nodes and connections that classified responders from non-responders in a previous analysis.

To our knowledge, this is the first study using semantic speech characteristics to study placebo-related alterations in chronic pain (although previous research has looked at word count measures of pronoun use between participants treated with either placebo and homeopathic remedies [291].

Additionally, while other researchers have recently used LSA to identify semantic components of neural activity in ECoG and fMRI data[292], we are the first to combine neuroimaging data, self-report measures, and automated language approaches together as part of a randomized clinical trial. Here we show that the language of placebo responders differs according to how semantically close their spoken words were to key words capturing the chronic pain and response experience. 4 words – “awareness”, “force”, “stigma”, and “magnify” significantly differed between responders and non-responders, explained over two-thirds of the variance in response, and were highly accurate in differentiating the two groups. The combination of these words is not easy to interpret, but we can see that together, they capture physical, psychological, and social components of placebo response. Additionally, some of the words also correspond well to previously identified personality traits of clinical placebo response in CBP – for example, responders have been shown to have increased emotional and interoceptive awareness (MAIA emotion) compared to non-responders, and here we also see that responders’ language is closer, on average, to the word or concept of “awareness” than nonresponders. These findings suggest that language can be used to further identify and define differences between patients in a clinical context, and that interviews may be important tools for future clinical trials.

Specific combinations of semantic similarity scores were also able to capture differences between groups. All language factors correlated to at least one of the key psychological or biological markers of placebo propensity in CBP. These included structural brain differences, functional brain differences, and personality profiles. These relationships show the that semantic differences seen in the interviews are significantly related to the biology and the mental state of the participant, suggesting that automated language analyses can provide results that are both physiologically and psychologically valid. Additionally, these relationships indicate that the semantic differences seen at visit 6 between responders and non-responders are not due to any contextual influences during the interview – instead, they imply that these properties either predetermine placebo propensity (since they are correlated to traits of response) or are caused by placebo response (and are picked up during the interview after all treatment and washout sessions).

The semantic language factors were also able to identify important functional connections on their own (without prior knowledge of significant connections or relevant communities). Whole brain network

analyses revealed different sets of functional connections in accordance with each of the 6 factors. These links are interesting in that they establish the brain regions and properties associated with different components of meaning, findings which are important in and of themselves. Importantly, nodes involved in these semantic links are shown to be scattered throughout the brain as opposed to located within traditional regions such as Broca's or Wernicke's area. This observation directly mirrors findings from a recent and exciting study by Huth et al [268] showing that semantic meaning is distributed across the whole brain in distinct patterns. Additionally, some of the language-associated links were also able to differentiate responders from non-responders; many of these links were novel (not found in previous reports) but importantly, many of them were the same as those shown to be predictive of response (e.g., frontal to S1M1 and frontal to PAG). The latter result highlights the power of language as a tool to discover brain mechanisms associated with chronic pain experience and its relief.

Our findings have important implications. First, they provide a general proof of concept that (a) specific properties of language can be used by themselves to identify clinically meaningful differences in cohorts of interest, and (b) can be combined with more traditional approaches to provide comparable results or even validate previous findings. Second, given the limitations of existing methods (discussed in the introduction), interviews and language analyses could be used either as an adjunct to other approaches when appropriate (for example, to better characterize or explain differences seen in self-report measures) or as surrogate measures for approaches which may not be as easily implemented or cost-effective (such as neuroimaging). Despite high computational demand, language measurements like the ones used in this paper are easy to implement and relatively automated.

CHAPTER 5: RESULTS FOR STUDY 3

Memory of chronic pain is biased by left posterior hippocampus morphology

Demographics

72 CBP (divided into a discovery and a validation dataset, **Table 21**) were asked to rate their pain and mood for the duration of an 8-week clinical trial studying placebo response. In this study, we only analyzed one week of interest that preceded the administration of any treatment (**Figure 24a**). During that week, patients were asked to rate their pain and mood 2 times per day using their smartphone app (**Figure 24b-c**) and provide a verbal rating about their pain experienced over the last week (i.e. pain memory). On average, participants were compliant when entering their pain and mood ratings with their app during this time (average compliance = 77.7 ± 21.1% SEM for discovery group; 76.9 ± 20.9% SEM for validation group).

	Discovery Group (CBP)				Validation Group (CBP)				Healthy Controls (CON)				Persisting Subacute Back Pain (SBPp)			
Number of participants	48				21				22				21			
Females (%)	17 (35.4)				9 (42.9)				9 (40.9)				11 (52.4)			
Number of participants in follow-up:	25				8				-				-			
	mean	SD	min	max	mean	SD	min	max	mean	SD	min	max	mean	SD	min	max
Age (years)	47.7	12.4	20	71	43.5	11.8	22	64	36.3	7.6	27	53	44.7	9.6	25	63
Duration of pain (years[&]) reported at Visit 1	4.6	6.8	0.5	40	4.7	3.5	0.5	13	-	-	-	-	9[&]	3.8	2.0	16
Time between scans (days)	-	-	-	-	-	-	-	-	376.1	13.1	359	421	379.1	21.9	352	445
Time before follow-up call (days)	259	64.7	126	369	83.5	37.8	10	115	-	-	-	-	-	-	-	-

Table 21: Demographics for Study 3. Basic demographics of all participants from all datasets. "&" = SBP participants entered study with subacute (<3 months) of back pain, and their pain duration is thus given in *weeks* instead of years; SD = standard deviation; "-" = not applicable for that dataset.

Chronic pain patients show memory bias in the setting of a clinical trial

As expected [293], pain and mood phone app ratings were anti-correlated for the majority of participants (75%) (**Figure 24d,** mean coefficient correlation: -0.52 ± 0.05 SEM; $t_{(47)}$ = -4.76, p<0.0001; one sample t-test), and the recalled pain was significantly higher than the average experienced pain monitored using the phone app (**Figure 25a**). The average memory discrepancy for the group (defined as the recalled pain minus the average experienced pain rated over the previous week) was 1.05 ± 0.18 SEM units on a 0-10 VAS scale. The self-reported pain memory was on average 18% higher than the average phone app ratings, and 37 out of the 48 people (77%) overestimated their pain intensity (**Figure 25b**). Importantly, age and gender identity were not related to memory bias, nor was pain duration, which did not correlate with pain memory or the extent of the discrepancy.

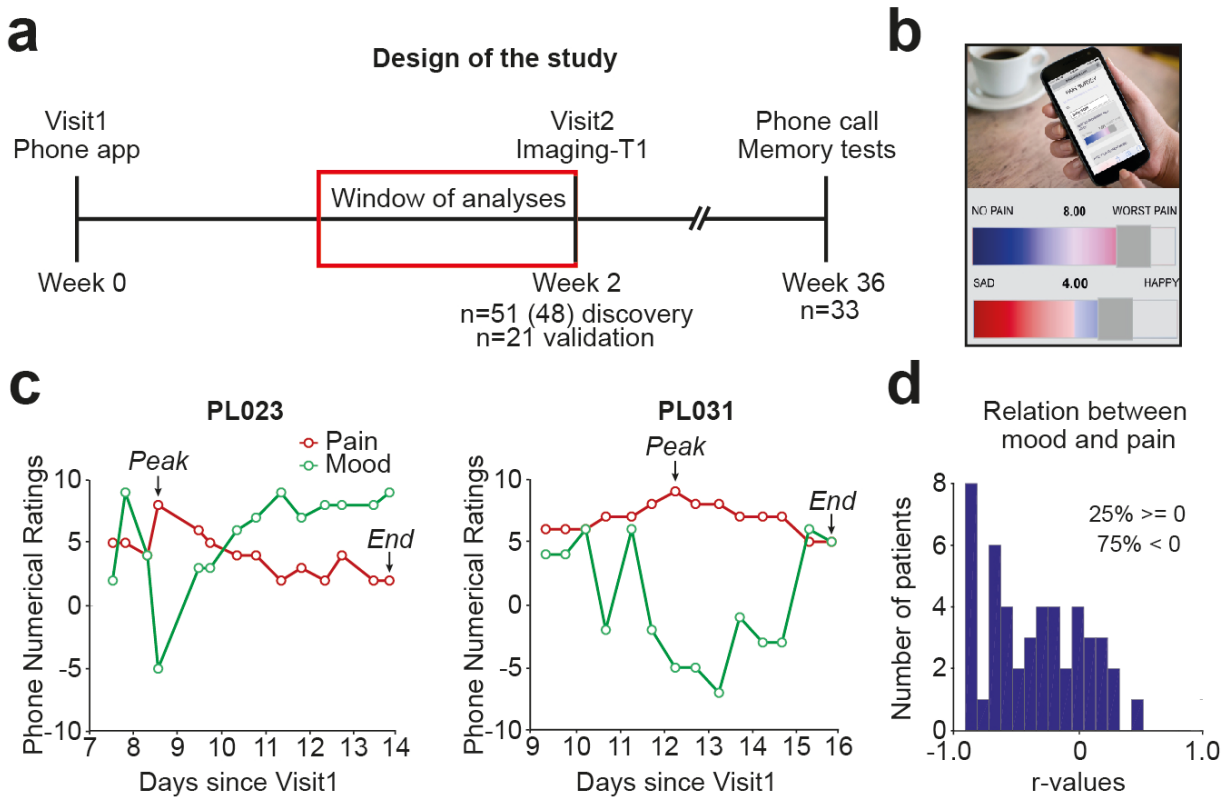


Figure 24: Study design and psychometric parameters derived from ratings of experienced CBP pain. **a.** Illustration of study design. Participants completed a battery of questionnaires at visit 1 (week 0) and were provided with a smart phone application (app) to track their back pain and mood twice a day for 2 weeks, after which they returned to complete an MRI scan and another set of questionnaires. Red box indicates that only the last 7 days of this rating period were used in analyses. A total of 72 people were

enrolled; 48 were used as a discovery group, 21 set aside for validation, and 3 excluded from analysis. After completion of the study, a subset of individuals (n=33) were contacted and asked follow-up questions to probe their memory of the study (on average 217 days after Visit 1). **b.** Example of the rating app. After entering in their participant IDs, patients rated how much pain they currently felt from 0 to 10 and the valence and magnitude of their current mood from -10 to +10. **c.** Examples of two participants' pain and mood ratings over one-week of the rating period are shown, with the peak and end indicated. **d**. Distribution of correlations between pain and mood ratings; the majority of participant's moods were negatively correlated to their pain intensity as expected.

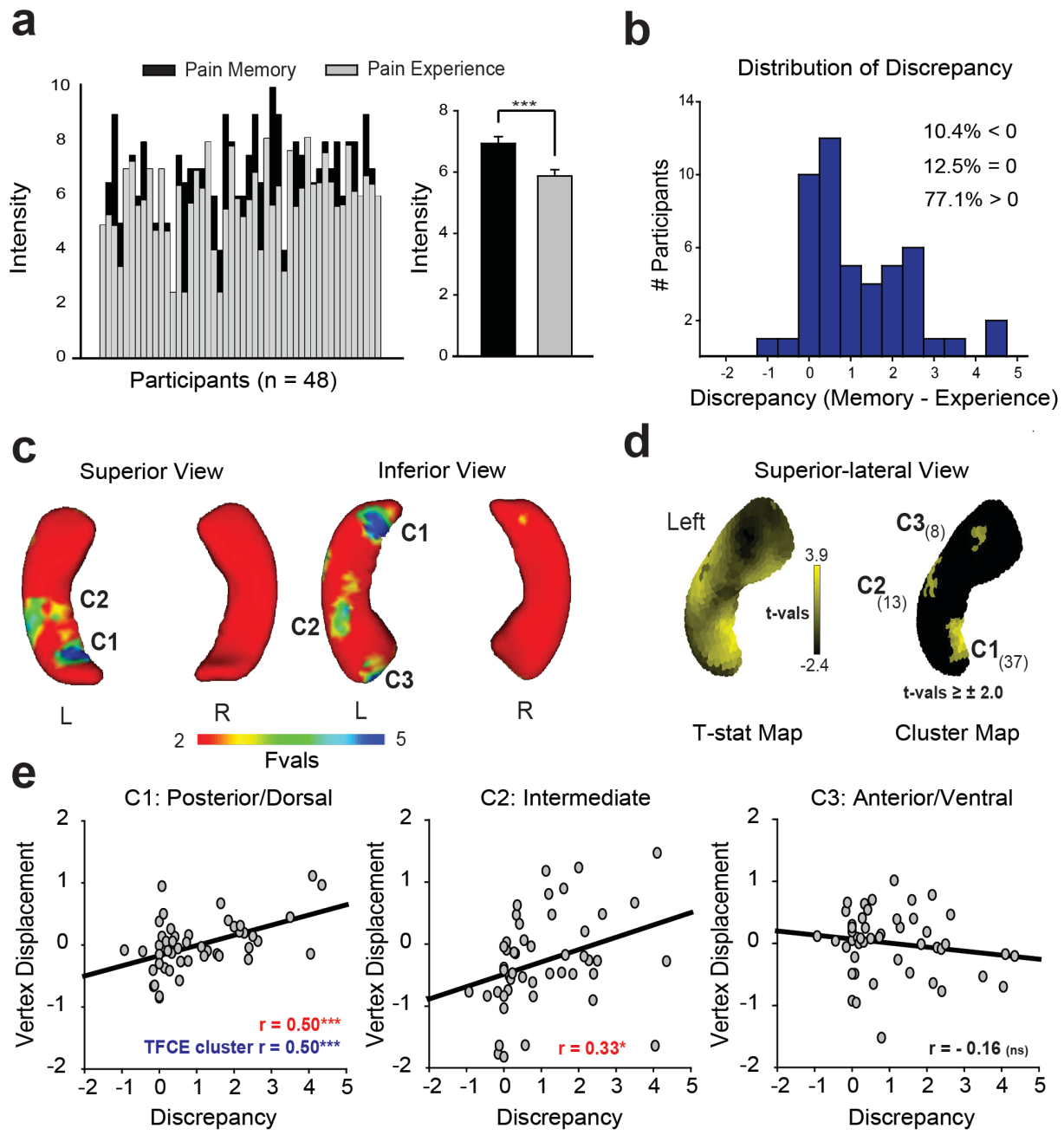


Figure 25: **Chronic back pain patients show exaggerated pain memory and this discrepancy correlates with left hippocampal shape displacement**. **a.** All participants' (discovery group) experienced (rated) pain from the app (averaged over one week, gray) plotted over their pain memory (black) during the same 7-day period. Visible black bars indicate a bias toward remembering higher pain than was actually experienced. Bar graph is group averaged result (average pain memory: 6.93 ± 0.22 SEM; average rated pain: 5.87 ± 0.21 SEM; $t_{(47)}$ = 5.75, paired t-test). **b.** Distribution of all participant's memory discrepancies (pain memory– rated pain). **c.** Left hippocampal shape displacement correlated to the memory discrepancy values displayed in **b**, with three areas (labeled C1-C3) being statistically significant on the surface output maps; colors on surface mesh display the strength of the correlation at each of the 732 vertices (F-values from uncorrected multivariate statistics). One area was located in posterior hippocampus, another in the intermediate region, and the third in the anterior hippocampus. Right hippocampus showed no relationship to discrepancy values. **d.** The group t-stat map (left) and the corresponding thresholded cluster map (right) for left hippocampus are shown. T-values are projected onto each vertex surface (circles), with colors ranging from black (negative values) to bright yellow (positive values). T-values were thresholded at ± 2.0. A k-means algorithm was used on the thresholded t-stat map to statistically restrict the three areas into distinct clusters that were most related to discrepancy (all vertices that were not clustered are depicted in black). The change in shape from each vertex (vertex displacement) in the clusters was extracted, averaged within the cluster, and correlated to discrepancy. Numbers in parentheses indicate # of vertices within that cluster. **e**. Correlations between left hippocampal vertex displacements and memory discrepancies for each identified cluster (C1-C3); cluster 1 (C1), corresponding the posterior hippocampus, and cluster 2 (C2), corresponding to the middle or intermediate hippocampus, showed a significant relationship between shape and discrepancy. Positive displacement values indicate an outward direction (expansion of shape) on average, whereas negative values indicate inward direction (shrinking of shape). For both C1 and C2, more outward displacement in these regions correlated to higher memory discrepancy. For all graphs, age and sex have been regressed as covariates of no interest. Of all clusters, only C1 survived TFCE cluster correction for multiple comparisons. * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$; ns = not significant

We further examined which experienced pain (**Table 22**) and mood (**Table 23**) parameters contributed to this discrepancy. As shown in Redelmeier and Kahneman [201], the psychophysical properties of the pain ratings of our CBP participants were highly correlated to one another and to their recalled pain (**Table 24 top rows**). Given the importance of the peak-end rule (pain will be remembered depending on its worst intensity and ending intensity), we examined the characteristics of pain and mood ratings. Unlike pain characteristics that highly correlated with memory of pain, none of the mood parameters correlated to recalled pain; however, both ending mood and ending pain were significantly anti-correlated to memory discrepancy (**Table 24, bottom rows**).

Phone App Ratings (VAS units, 0-10)				
	Discovery Group		**Validation Group**	
	mean	SEM	mean	SEM
Average Pain (1 week)[&]	5.87	0.21	6.26	0.29
Total Pain (AUC)	58.70	3.50	63.20	5.48
Peak Pain	7.69	0.22	8.00	0.35
End Pain	5.85	1.70	6.02	0.56
Self-Report and Questionnaire Measures				
	mean	SEM	mean	SEM
Average Recalled Pain (1 week, VAS 0-10)	6.93	0.22	6.52	0.33
NRS (0-100) at Visit 1	56.48	3.52	62.33	4.61
NRS (0-100) at time of recall (current pain, Visit 2)	57.25	2.96	56.38	6.21

Table 22: Characteristics of Pain. Pain measurements and pain-related questionnaire data. Top section shows the calculations based on the daily pain ratings participants completed over the previous week during their baseline period of the study. Bottom section shows the questionnaire measures used at Visit 1 (study entry) and Visit 2 (current, at the time of recall and scanning). VAS = Visual Analogue Scale; NRS = Numeric Rating Scale; SEM = standard error

Phone App Ratings (VAS units, -10 to +10)				
	Discovery Group		**Validation Group**	
	mean	SEM	mean	SEM
Average Mood (1 week)	3.64	0.54	4.31	0.64
Total Mood (AUC)	34.57	5.59	42.93	7.79
Peak Mood	6.71	0.47	7.05	0.65
End Mood	3.83	0.59	4.52	0.76
Self-Report and Questionnaire Measures				
	mean	SEM	mean	SEM
PANAS positive at Visit 1	35.31	1.13	34.33	1.78
PANAS negative at Visit 1	18.62	1.10	18.48	0.99
PANAS positive at Visit 2 (current mood)	34.49	1.03	33.33	1.89
PANAS negative at Visit 2 (current mood)	19.13	1.05	19.62	1.62

Table 23: Characteristics of Mood and Cognition. Mood rating measurements and mood-related questionnaire data. Top section shows the calculations based on the daily mood ratings participants completed over the previous week during the baseline period of the study; calculations match those for pain in **table S2**. Bottom section shows questionnaire data for mood at Visit 1 (study entry) and Visit 2

(current, at the time of recall and scanning). PANAS = Positive And Negative Affect Scale; N = 48 for all discovery group calculations for except the following, which had missing data points: PANAS positive V1 = 47, PANAS negative V1 = 46, PANAS negative V2 = 47; SEM = standard error

	memory	peak pain	end pain	ave pain	auc pain	pe pain	peak mood	end mood	ave mood	auc mood	pe mood
recalled pain (memory)		0.66**	0.43**	0.65***	0.43**	0.53**	-0.18	-0.14	-0.16	-0.19	-0.18
peak pain	0.66***		0.37*	0.75***	0.57***	0.68***	0.05	0.15	0.09	0.03	0.15
end pain	0.43**	0.37*		0.73***	0.28#	0.91***	-0.05	0.03	0.01	-0.05	0.04
average (ave) pain	0.65***	0.75***	0.73***		0.59***	0.85***	-0.06	0.10	0.01	-0.07	0.11
total (auc) pain	0.43**	0.57***	0.28#	0.59***		0.44**	-0.11	-0.15	-0.16	0.02	-0.11
peak*end (pe) pain	0.53**	0.68***	0.91***	0.85***	0.44**		0.01	0.11	0.07	0.004	0.12
peak mood	-0.18	0.05	-0.05	-0.06	-0.11	0.01		0.68***	0.79***	0.77***	0.68***
end mood	-0.14	0.15	0.03	0.10	-0.15	0.11	0.68***		0.84***	0.74***	0.94***
average (ave) mood	-0.16	0.09	0.01	0.01	-0.16	0.07	0.79***	0.84***		0.93***	0.86***
total (auc) mood	-0.19	0.03	-0.05	-0.07	0.02	0.004	0.77***	0.74***	0.93***		0.78***
peak*end (pe) mood	-0.18	0.15	0.04	0.11	-0.11	0.12	0.68***	0.94***	0.86***	0.78***	

	memory	peak pain	end pain	ave pain	auc pain	pe pain	peak mood	end mood	ave mood	auc mood	pe mood
discrepancy	0.47***	-0.07	-0.32*	-0.37*	-0.16	-0.33*	-0.16	-0.29*	-0.2	-0.15	-0.33*

Table 24: Summary of relationships between psychometric pain and mood app ratings. Top panel indicates that the pain and mood properties extracted from individuals' daily ratings are highly correlated to one another. Bottom panel shows how each of these properties correlates with a person's memory bias (discrepancy = recalled pain-average pain). The Pearson correlation coefficients (r) are shown at their significance level (2-tailed) with # = trending at 0.05<p<0.10, * = p < 0.05, ** = p < 0.01, and *** = p <0.001. AUC = area under the curve; N = 48 for all correlations. The variable "peak*end (pe)" refers to averaging peak and end pain or peak and end mood together according to Kahneman and collegues' peak-end rule.

Shape displacement of the left posterior hippocampus is related to memory discrepancy

The volume of the left (average volume = 3755.5 ± 481.7 mm^3) and the right (average volume = 3855.1 ± 527.0 mm^3) hippocampus were invariant to memory discrepancy. However, a vertex-wise shape analysis in relation to pain memory bias uncovered left hippocampal areas correlated with memory discrepancy (**Figure 25c-e**); Cluster 1 (C1) corresponded to the posterior hippocampus, cluster 2 (C2) primarily to the intermediate hippocampus with some posterior overlap, and cluster 3 (C3) to the anterior hippocampus. From these, only C1 survived threshold-free cluster enhancement (TFCE) correction for multiple comparisons (17 out of 37 vertices), indicating that only posterior hippocampal shape distortion is related to pain memory discrepancy.

Regression models for pain memory

Multi-factor multiple regression analyses were used to test if the psychophysical qualities (peak, end, total, and average) examined from the experienced pain and mood ratings, current pain and mood parameters from the day of memory assessment (given [190-196]), and hippocampal shape displacement can explain the pain memory values. We incrementally tested 4 separate multivariate models; each regression built off the previous one to explain the memory of pain. The first model only used the parameters initially studied by Redelmeier and Kahneman [201] --peak pain, end pain, average pain, and AUC-- and showed that pain memory significantly depended on the peak pain and also the average pain over the week. The second model entered these 2 variables with mood variables and indicated that the mood at the end of the rating period explained unique variance beyond peak pain and average pain. A third model combined these 3 surviving rating parameters current pain (Numeric Rating Scale, NRS) and mood (Positive and Negative Affect Scale, PANAS) scores on the day of memory assessment; this model indicated that the current pain and emotional state failed to significantly contribute to pain memory and in turn did not provide a better alternative to the second model. The fourth and final model was the most comprehensive, incorporating experienced pain, experienced mood, and hippocampal morphometry. This showed that C1 hippocampal shape displacement in combination with average experienced pain accounted for 55% of the variance in pain memory (**Table 25).**

Model	Surviving Parameter(s)	Coefficient (SEM)	p-value	95% C.I.	Eliminated Variables (p > 0.10)
1	**Basic Pain Model (App Data)**				
	peak pain	0.396 (0.16)	0.02	[0.07,0.73]	End pain, total (auc) pain
	average pain	0.369 (0.17)	0.04	[0.02,0.71]	
	N = 48; R^2 = 0.48; Adj R^2 = 0.46; $F(2,45)$ = 21.03; p < 0.0000				
2	**Basic Pain + Mood Model (App Data)**				
	peak pain	0.438 (0.16)	0.01	[0.12,0.76]	auc mood, average mood, peak mood
	average pain	0.362 (0.16)	0.03	[0.03,0.69]	
	end mood	-0.09 (0.04)	0.03	[-0.17,-0.01]	
	N = 48; R^2 = 0.54; Adj R^2 = 0.51; $F(3,44)$ = 17.17; p < 0.0000				
3	**App Data + Current Variables (Pain + Mood)**				

	peak pain	0.322 (0.17)	0.06	[-0.02,0.66]	PANAS negative, PANAS positive, NRS
	average pain	0.453 (0.17)	0.01	[0.11,0.79]	
	end mood	-0.717(0.04)	0.08	[-0.15,0.01]	
	N = 47; R^2 = 0.53; Adj R^2 = 0.50; *F(3,43)* = 16.06; p < 0.0000				
4	**Comprehensive Model (Pain + Mood + Neuroimaging)**				
	average pain	0.734 (0.10)	p<0.0000	[0.53,0.94]	intermediate hippocampus (cluster 2), peak pain, end mood
	posterior hippocampus (cluster 1 uncorrected &)	1.336 (0.33)	p<0.0000	[0.66,2.01]	
	N = 48; R^2 = 0.57; Adj R^2 = 0.55; *F(2,45)* = 29.69; p < 0.0000				
	Final equation: recalled pain = f(x) = 2.61 + 1.34*cluster1 + 0.73*average pain				

Table 25: Summary of memory model results. Multiple regression analyses were performed to indicate what independent variables significantly influenced participants' memory of their pain (recalled pain = dependent variable). Regressions were run in a hierarchical manner such that each model built off of the previous one; only those variables that survived the previous model(s) were entered into the subsequent one. Groupings of variables were chosen apriori based off of previous literature; the first model tested the established peak-end rule of explaining memory with pain ratings, the second added to the peak-end rule with mood ratings (which can interact with the memory of pain, particularly at the end of a painful event), the third incorporated current pain and mood (which have also been shown to bias recall), and the fourth and final regression entered all surviving pain and mood parameters with the significant hippocampal shape displacement to create a comprehensive model accounting for behavior and neuroanatomy. "&" indicates that the posterior hippocampal cluster that survived TFCE correction for multiple comparisons (17 out of 37 vertex coordinates) was also tested in the final model; the results were identical. Adj R^2 = adjusted R^2, which represents the coefficient of determination after accounting for the number of predictors in the model; SEM = standard error; C.I. = confidence interval; auc = area under the curve

Validation and predictive utility of the model

To test the reliability and generalizability of our results, we attempted replication of our main findings in the CBP patients (n = 21) reserved for validation. Over half of these patients (n = 12, ~57%) displayed a discrepancy biased toward an overestimation of their pain, although average memory and experienced pain outcomes were not different (**Figure 26a)**. The vertex displacement from the left posterior hippocampus was extracted and averaged for this validation CBP group within C1, using the 37 coordinates defined from the discovery group (**Figure 26b, left**). The final model from the discovery group (Model 4 – C1 displacement + average rated pain) was used to predict these new participants' recalled pain values. **Figure 26b (right)** shows that the predicted values were strongly correlated to the actual

values reported. These results indicate that our model, and its component outcome measures, is reliable in explaining and predicting memory of pain.

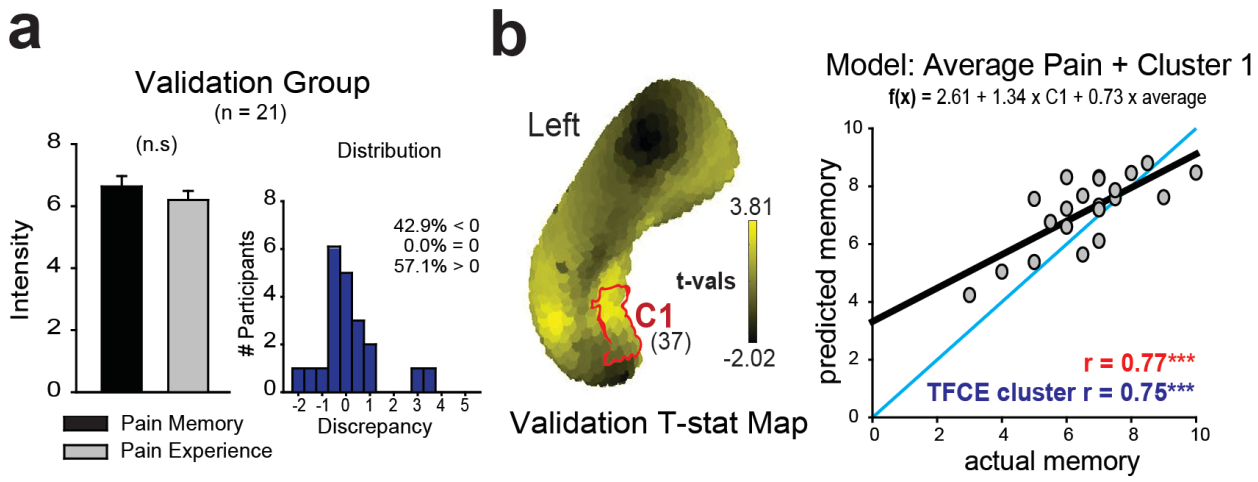


Figure 26: Validation and replication of left posterior hippocampus shape displacement for pain memory discrepancy. a. Unlike the discovery group, the validation group (n=21) did not result in a statistically significant difference between pain memory and pain experience (left panel, memory = 6.52 ± 0.33 SEM; experience = 6.26 ± 0.29 SEM; $t_{(20)}$ = 1.08, p = 0.29, paired t-test), although the majority of individuals still reported higher pain memory than experience, indicating a memory bias also existed (right panel). **b**. To validate the model from the discovery group, the vertex displacement from the 37 vertices in discovery cluster 1 (C1) were extracted from the left posterior hippocampus of the validation group (left panel, red outline). Participant's average pain and C1 vertex displacement values were entered into the model equation (provided at the top of the graph, right panel) to predict their memory of pain. The correlation between the predicted memory from the equation parameters and the actual memory reported is shown – these values were significantly correlated, validating the model. An identity line, indicating a perfect correlation, is shown in blue. * = p<0.05; ** = p<0.01; *** = p<0.001; ns = not significant.

Posterior hippocampal shape is stable over 1 year and with development of chronic pain

It has been established that the hippocampus is prone to dynamic changes in shape and volume as part of normal human development and aging [294], prolonged pain impacts the morphometry of the hippocampus [45], and that sub-regions of the hippocampus are differentially associated with stress and anxiety [295]. To test whether the left posterior hippocampal C1 shape displacement was related to pain memory in our CBP patients was influenced by a general aging process or a consequence of chronic pain, we compared the displacement of vertices corresponding to C1 in two independent data sets (**Figure 27a**). We tested the stability of C1 in healthy controls (CONs) between two scans, which were approximately 1 year apart (376.1 ± 13.3 days). The average change in displacement within this region was 0.06 mm ± 0.06 SEM; this change in shape was not significant as none of the vertices had t-stats

exceeding the threshold of ± 2.0 (minimum absolute t-value in the region = 0.27; maximum absolute t-value = 1.61), and no displacements survived TFCE, indicating that C1 was not an area that significantly changed in this timeframe. We next compared the C1 vertex displacement in 21 participants with persisting subacute back pain (SBPp) between two scans that were also 1 year apart (379.1 ± 21.9 days). These individuals entered the study with a new onset of back pain (less than 3 months) and ended up having their pain persist over the year, thus developing chronic pain and essentially becoming similar to our CBP cohort. The average change in displacement within C1 for SBPp between scans was 0.09 mm ± 0.10 SEM; none of the vertices in the cluster had t-stats surviving the ± 2.0 threshold (minimum absolute t-value = 0.29; maximum absolute t-value = 1.23), and none of the p-values survived TFCE. Additionally, there were no differences between CON or SBPp (**Figure 27a**) in either the average displacement of the region at scan 2 (CON: 0.03 mm ± 0.13 SEM; SBPp: 0.05 mm ± 0.23 SEM; unpaired t-test: $t_{(41)} = -0.065$; $p = 0.95$) or the average change in displacement between scans (unpaired t-test: $t_{(41)} = -0.29$; $p = 0.77$). These results indicate that shape displacement in C1 of the posterior hippocampus seems hardwired, as this region does not appear to change over 1 year and is not influenced by the stress of persistent pain over 1 year.

Pain memory bias is specific to reward/punishment personality traits in CBP

Although we discovered that the memory of pain is systematically biased and related to hippocampal shape, it was still unknown whether these biases and shape differences were generalizable to other memories. We contacted and questioned our participants regarding their memories of the study and also tested them for short-term memory (**Table 26**). We attempted to call back all participants whose anatomical data were analyzed; of these, 25 people from the discovery group and 8 from the validation group were reached by phone. Here, the two groups were combined to increase our sample size (total n=33). The average time between participant's last visit date and the phone contact was 216.7 ± 96.4 days, and importantly, none of the answers provided by participants significantly correlated to the length of this interim period, indicating that the amount of time between the study and the phone call did not impact obtained memory results. The participants' memory of baseline pain provided during the phone call was significantly correlated with the pain memory provided at their MRI visit ($r = 0.41$, $p = 0.019$,

Figure 27b); there was still a large discrepancy in recalled pain 216 days after the study, with over 75% of individuals remembering higher pain than actually experienced instantaneous pain (distribution also shown in **Figure 27b**). These results indicate that once a memory trace of painful events is distorted, this bias persists over a long time.

Recalled Values (compared to those expected for most participants)							
	Question Topic	mean	sem	*N*	expected[&]		
	Average **Pain** during first 2 weeks of study	7.0	1.22	33	-		
	Average **Mood** during first 2 weeks of study	3.5	0.60	32	-		
	# **Visits** (total) during the study	7.7	1.34	30	~ 6 visits		
	# Visits with **Scans** (total) during the study	4.5	0.78	33	~ 4 scans		
	Total **Compensation** ($) from all visits	372.0	64.76	32	~ $448.00		
	One phone **Rating**'s worth ($)	1.5	0.25	29	$0.25 each		
Discrepancy Values (Recalled during phone call – Actual during study)							
	Question Topic	mean	sem	*N*	%<0	%=0	%>0
	Average **Pain** during first 2 weeks of study	0.8	0.30	33	24.2	0.0	75.8
	Average **Mood** during first 2 weeks of study	-0.7	0.96	32	50.0	0.0	50.0
	# **Visits** (total) during the study	0.1	0.70	30	30.0	20.0	50.0
	# Visits with **Scans** (total) during the study	0.3	0.38	33	33.3	30.3	36.4
	Total **Compensation** ($) from all visits	-37.9	31.22	32	59.4	6.3	34.4
	One phone **Rating's** worth ($)	1.2[&&]	0.71	29	11.5	80.8	7.7
Current Measurements During Phone Call							
	Question Topic	mean	sem	*N*			
	Pain (VAS, 0 to 10)	5.4	0.94	33			
	Mood (VAS, -10 to +10)	5.3	0.94	33			
Behavioral Data							
	Questionnaire Scores from Visit 1	mean	sem	*N*[&&&]	Pain (r-val)	$ (r-val)	C1 (r-val)
	LAQ	62.9	1.55	33	-0.38*	0.21	-0.41*
	PCS rumination	7.3	0.70	33	0.15	0.16	0.05
	PCS magnification	4.0	0.46	33	0.25	-0.20	0.002
	PCS helplessness	8.6	0.79	33	-0.06	0.09	-0.12
	PSQ no-pain subscale	4.6	0.75	33	-0.14	-0.10	-0.11
	PSQ pain subscale	75.2	4.13	33	0.08	0.03	0.16
	PASS avoidance behavior	13.0	1.25	33	-0.16	-0.004	-0.12
	PASS cognitive anxiety	11.8	1.26	33	0.09	-0.12	-0.12
	PASS fear	7.4	1.11	33	0.14	-0.08	-0.09
	PASS physiological anxiety	6.0	0.95	33	0.02	-0.05	-0.20

Table 26: Summary of responses from follow-up call. Data corresponds to answers from questions given in **Table 3.** Discrepancy scores for every question were calculated, with positive scores indicating greater (over-estimated) recalled answers than what actually happened and negative scores indicating lower (under-estimated) recalled answers than what was experienced. Percentages indicate accuracy measurements: 0 = perfectly accurate memories, <0 = underestimation, and >0= overestimation. Other than pain memory, memory of total compensation was the only other measure that was not centered around 0 (showing a systematic bias). Average scores and sub-scores for 4 self-report measures are also provided (LAQ: Loss Aversion Questionnaire; PCS: Pain Catastrophizing Scale; PSQ: Pain Sensitivity Questionnaire; PASS = Pain Anxiety Symptoms Scale). R-values indicate the correlation with each of the scales and the 3 main parameters of interest: pain memory discrepancy (pain), total compensation discrepancy ($), and hippocampal shape displacement (posterior cluster C1). "-" = data not applicable for this measure; "&" = values are estimated based on what the majority of participants would have experienced but there are exceptions, as some patients did not complete the study and therefore would have had less visits and less scans, for example. Actual values differed from expected values and were calculated based on each participant's data. "&&" = this particular measurement had outlying values in it that drove the mean up (3 participants responded with values between $7 and $18); if those 3 values are removed (n=26), the discrepancy becomes -0.0008 ± 0.017 SEM, indicating high accuracy. "&&&" = there were no missing values for any questionnaire measures, resulting in a sample size of n=33 for all scores; however, one person was missing from the monetary ($) calculation, so this correlation had a sample size of n = 32 for all scores

To determine whether the pain memory discrepancy seen in our CBP participants was indicative of short-term memory (STM) impairment, we examined their scores on the Montreal Cognitive Assessment (MoCA) delayed recall task designed to detect cognitive impairments. On average, participants remembered 3 or 4 out of 5 words on this task (average = 3.70 ± 1.2 words, minimum: 0, maximum: 5), with only 15.2% (n = 5) remembering less than 3 words (**Figure 27c**), implying that the majority of our participants had no problems with STM [296]. Additionally, the number of words remembered did not correlate with the pain discrepancy at the MRI visit (Spearman's rho = -0.04, p = 0.83), with pain memory from the follow-up phone call (Spearman's rho = 0.05, p = 0.78), or with the vertex displacement of the C1 posterior hippocampus (Spearman's rho = -0.19, p = 0.29). These results suggest that neither CBP patients' pain discrepancy nor their posterior hippocampal C1 displacement are related to insufficient STM capabilities.

To test the relationship between pain discrepancy findings and other memories related to the experience of participating in the trial, we computed a discrepancy score for all queries administered during the follow-up phone call. Responses regarding mood, number of visits, number of MRI scans, and monetary compensation received per app rating all showed distributions centered around 0 (indicating response accuracy). Only memory of total compensation during the study was skewed with 59% of participants under-estimating the amount received (**Figure 27d** and **Table 26**). We correlated the current

memory bias in pain with these additional discrepancy scores. Again, only the recall of total monetary compensation showed a significant relationship with the current pain discrepancy (**Figure 27e**). This relationship was negative – that is, the more pain someone remembered having, the less monetary compensation they also thought they received, indicating a recall bias related to valence, reward, and/or punishment.

To identify personality characteristics that may be related to pain and monetary reward biases we examined their relationship to four personality trait outcomes (Loss Aversion Questionnaire – LAQ [246], Pain Catastrophizing Scale – PCS [230] , Pain Sensitivity Questionnaire – PSQ [232], and Pain Anxiety Symptoms Scale - PASS [231]; **Table 26, bottom**), collected at time of entry into the study. Of these, only LAQ scores were negatively correlated with both pain and money discrepancies (**Figure 27f**); moreover, both LAQ and money discrepancy positively correlated with hippocampal C1 shape displacement (**Figure 27g**). Therefore, to understand the co-dependencies between these variables, we tested the hypothesis that C1 shape displacement mediated the effects of personality (LAQ) and pain memory discrepancy. **Figure 27h** shows the paths and their standardized coefficients in this mediation model; there was a significant indirect effect from LAQ to pain discrepancy, indicating that posterior hippocampal C1 shape significantly mediated this relationship. Importantly, there was no significant mediation effect of C1 between LAQ and monetary discrepancy.

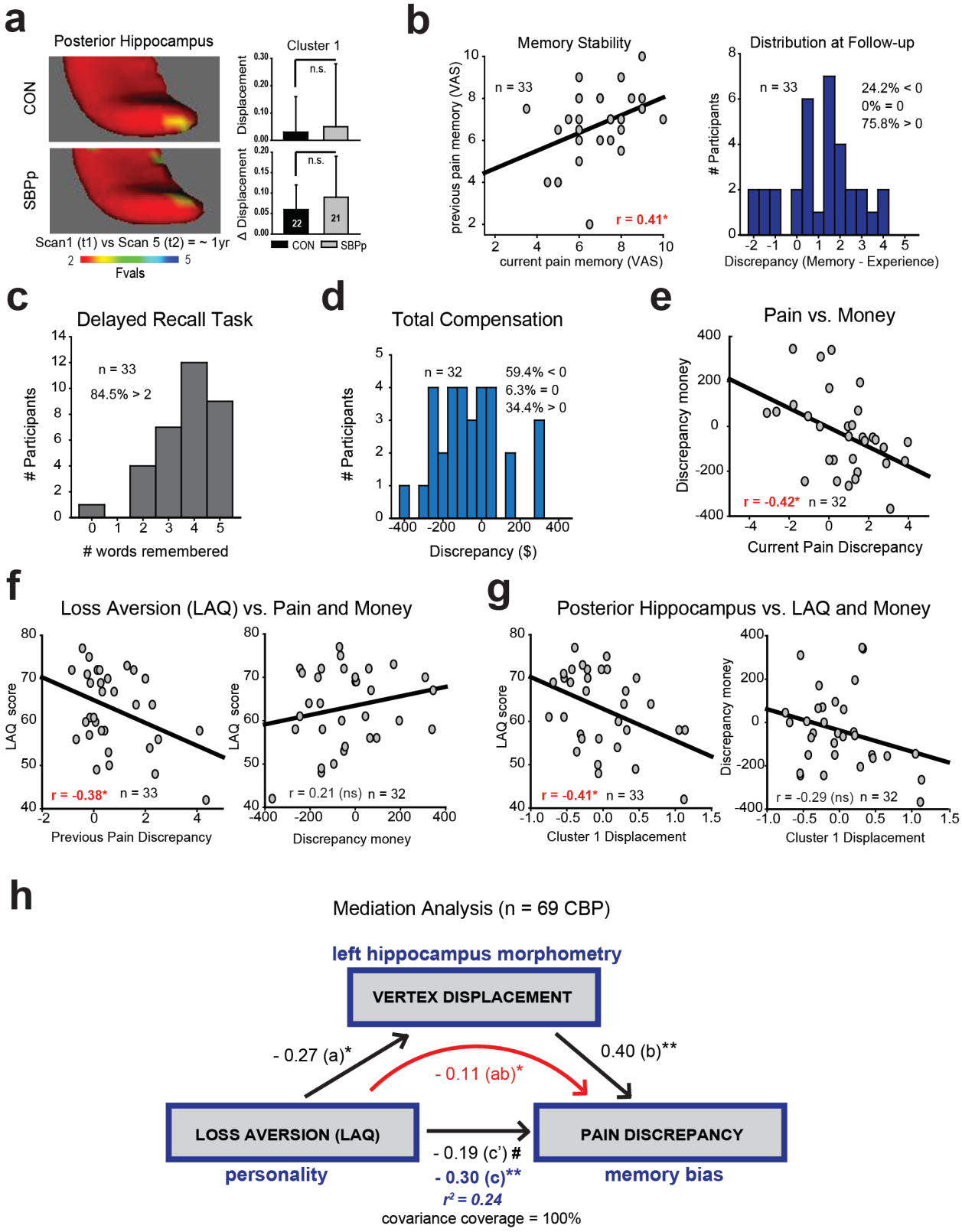

a
Posterior Hippocampus
CON
SBPp
Scan1 (t1) vs Scan 5 (t2) = ~ 1yr
2
5
Fvals
Cluster 1
Displacement
Δ Displacement
n.s.
n.s.
22
21
CON
SBPp
b
Memory Stability
n = 33
r = 0.41*
previous pain memory (VAS)
current pain memory (VAS)
Distribution at Follow-up
n = 33
24.2% < 0
0% = 0
75.8% > 0
Participants
Discrepancy (Memory - Experience)
c
Delayed Recall Task
n = 33
84.5% > 2
Participants
words remembered
d
Total Compensation
n = 32
59.4% < 0
6.3% = 0
34.4% > 0
Participants
Discrepancy ($)
e
Pain vs. Money
Discrepancy money
r = -0.42* n = 32
Current Pain Discrepancy
f
Loss Aversion (LAQ) vs. Pain and Money
LAQ score
r = -0.38* n = 33
Previous Pain Discrepancy
r = 0.21 (ns) n = 32
Discrepancy money
g
Posterior Hippocampus vs. LAQ and Money
LAQ score
r = -0.41* n = 33
Cluster 1 Displacement
Discrepancy money
r = -0.29 (ns) n = 32
Cluster 1 Displacement
h
Mediation Analysis (n = 69 CBP)
left hippocampus morphometry
VERTEX DISPLACEMENT
- 0.27 (a)*
- 0.11 (ab)*
0.40 (b)**
LOSS AVERSION (LAQ)
personality
- 0.19 (c') #
- 0.30 (c)**
PAIN DISCREPANCY
memory bias
r² = 0.24
covariance coverage = 100%

Figure 27: Longitudinal and follow-up analyses relate pain memory bias to loss aversion. **a**. Paired t-tests of hippocampal shape displacement between two scans separated by approximately 1 year. Surface map F-statistics for the left hippocampus for healthy controls (CON, n=22) and individuals with persisting sub-acute back pain (SBPp, n=21); images have been zoomed to focus on the posterior portion and the lack of significance in this region can be visually appreciated. Statistics performed on both groups showed no within subject differences due to time in cluster 1. Bar graphs indicate between subject statistics; there were no differences in the average displacement of this C1 region (top) or average change in C1 displacement (delta, bottom) between CON and SBPp. **b.** Recalled pain memory assessed 36 weeks after study completion (current pain memory) was significantly correlated to the memory of the pain assessed at the end of the week during the study (previous pain memory), with the majority of participants still maintaining a discrepancy biased towards exaggerated pain memory (distribution plot). **c.** Participants performed at or above average in the delayed recall task, with the majority remembering at least 3 words **d.** Discrepancies were calculated for all questions asked during the follow-up phone call (**table S6**). Only total compensation (the amount of money earned during the entire trial) showed a memory bias. **e.** Discrepancy of money and pain memories were anti-correlated, such that participants biased toward overestimated pain levels tended to underestimate the amount of money they received during the study. **f.** Loss aversion (LAQ) scores were significantly anti-correlated with pain discrepancy; the opposite relationship with money discrepancy was found but not significant. **g.** LAQ scores were also significantly negatively correlated to posterior shape distortion of the left hippocampus; discrepancy of money was not. **h.** After combining all participants from discovery and validation groups, a mediation analysis was used to measure the contribution of shape displacement of the hippocampus on the relationship between reward-related personality (loss aversion) and memory bias based on pain discrepancy scores. This effect was significantly mediated by left posterior hippocampal C1 shape (indirect pathway 95% CI: [-0.196,-0.022]; R^2 = 0.074 of unique variance). Path a = effect of LAQ on mediator variable (hippocampal shape); path b = effect of mediator on pain memory discrepancy; path c = effect of LAQ on pain memory discrepancy (*total effect*); path c' = effect of LAQ on pain memory after controlling for hippocampal shape (*direct effect*); path ab = amount of mediation produced by the hippocampal vertex displacement (*indirect effect*). * = $p<0.05$; ** = $p<0.01$; # = $0.05<p<0.10$ (trending)

Discussion

We identified psychometric, psychological, and neuroanatomical characteristics that underlie memory bias in self-reported intensity of chronic back pain. More than 70% of CBP patients exaggerated their pain memory by about 20% from experienced pain ratings, a bias that persisted many months after study completion. In accordance with Redelmeier and Kahneman's [201] peak-end rule, experienced peak and average pain ratings, with ending mood ratings, explained more than 50% of the variance in pain memory. We also demonstrated that shape displacement in the left posterior hippocampus, C1, was related to the discrepancy between pain memory and pain experience. This shape displacement, in combination with the average experienced pain, explained nearly 60% of the variance in pain memory, a finding that was validated in a separate group of CBP. Importantly, the shape of C1 was invariant over 1 year, unperturbed with development of chronic pain, and appeared to be independent from other likely confounds, including age, gender, duration of pain, short-term memory dysfunction, pain anxiety or catastrophizing, and generalized habits of exaggeration. Mediation analysis identified that C1 linked loss

aversion personality characteristics to back pain memory bias. Thus, we not only reproduce previous psychometric results found in pain and mood ratings but also identify a hardwired biological substrate responsible for distortions in pain memories, expanding on the role of the hippocampal mechanisms in chronic pain.

Our findings have important implications for understanding and treating pain. Given the reliance on self-reported numerical ratings of pain to influence the type and duration of treatment in chronic pain patients, our results emphasize that such retrospective measures are inaccurate and often are at least 10-20% higher than the patients' actual experience. Importantly, the magnitude of the discrepancy corresponds to thresholds often utilized in determining clinically meaningful interventions (~20% reduction in pain intensity). These results raise philosophical quandaries as to the relative significance of experience and memory, especially in clinical decision-making in pain management.

In the field of pain, the peak-end rule has identified that humans do not simply sum their pain over time to report a totality of experience but instead average their worst painful moment with their most recent level of pain. Replication of this memory shortcut in our participants' pain ratings highlights its robustness as a heuristic strategy utilized across healthy individuals and people in acute or chronic pain. Furthermore, several studies have demonstrated how mood during a painful event can influence recalled discomfort [191]. Kent's highly-cited study [197] showed that individuals who were highly anxious regarding dental examinations later rated their remembered pain as higher than experienced compared to individuals with lower anxiety; similar findings have also been reported in children [198]. Likewise, labor pain is retrospectively rated as less severe than was previously rated [190, 192-194], as is the pain of running a marathon [195], both of which are likely due to the impact of positive emotions at the end of each event. Our results also match these previous findings, as pain was shown to be anti-correlated with mood, and the mood at the end of the rating period also accounted for a significant portion of the variance in the reported pain memory. Thus the peak-end rule for both pain and mood are present to various extents in the data presented here.

The finding that a specific cluster on the surface of the left posterior hippocampus is associated with and predictive of memory discrepancy is novel. Although the structure and volume of the human hippocampus have been used to predict or significantly explain a variety of inter-individual differences,

including cognitive ability, psychiatric illnesses, and risk for chronic pain [45, 297, 298], this is the first time that the shape of the hippocampus has been utilized to predict memories of pain. The hippocampus can be subdivided functionally and structurally along the anterior-posterior (longitudinal) axis [299, 300], with distinctions made according to genetic expression, cell type patterns, and connectivity to other brain regions [301]. While the anterior hippocampus is more involved with emotional processing, the posterior hippocampus is primarily associated with conceptual or spatial memories, including recall of rules, contexts, language, and spatial navigation, although it is also more generally involved in learning, information processing, timing of repeated events, and memory retrieval and consolidation [301, 302]. Additionally, there is evidence that the posterior hippocampus's cells are better able to track continuous changes in time, especially for repeated stimuli [52], which might contribute to its role in remembering elements of pain ratings. We used a reverse-inference term-based meta-analytic approach to capture best-associated words with each of the clusters on the left hippocampal surface (Neurosynth [303]). The following words were associated with each of the clusters: C1: "encoding", "retrieval", "details", and "episodic memory"; C2: "encoding" and "mild cognitive impairment (MCI)"; and C3: "emotion", "amygdala", "fear", and "facial expression". These associations substantiate previous findings indicating functional differentiation of the hippocampus, with rostral sites typically more involved in emotional processes while caudal sites instead engaged in the encoding and retrieval of memories. Our results suggest that anterior and posterior hippocampus contribute in opposite ways to memory discrepancy, the former exaggerating while the latter minimizing the correspondence between the pain experienced and its memory. The general surface area of the hippocampus is thought to reflect the migration, proliferation, differentiation, and targeting of various cells as part of the neurodevelopment process; outward displacements of the surface may then reflect enhanced intra- or extracellular connectivity of this region [304], whereas inward displacement may represent decreased connectivity due to abnormal development or perturbations to the area. However, it is still unclear exactly how structure (in this case, differences in surface morphology) dictates function, and further future research is needed in the topic. Here, shape displacement in posterior hippocampus was related to memory discrepancies for reward (money) and punishment (pain) and furthermore mediated the relationship between loss aversion personality traits and pain memory bias. Critically, deformations within the identified cluster remained stable across time for two separate

participant groups and did not significantly differ between controls or SBPp. This suggests that the regional shape displacement seen in the posterior hippocampus cluster is hardwired and thus a deterministic individualized memory trait.

The asymmetry of this shape displacement is noteworthy, as only the left hippocampus showed any relationship with pain memory discrepancy. An accumulating body of evidence supports the notion that the hippocampus has hemispheric functional specialization, relatively preserved across species [305], indicating that this functional asymmetry interacts with its antero-posterior structural segregations to give rise to combined functional-structural specifications. For example, the left hippocampus is associated with verbal memory processes whereas the right is associated with more spatially-dependent memories [305]. Additionally, distribution of functional networks from the right and left hippocampus differ depending on location, with the right anterior and left posterior hippocampus exhibiting large, distributed functional networks, whereas the left anterior and right posterior segments are primarily confined to fronto-limbic networks [305]; and a longer and wider longitudinal axis in the left hippocampus significantly predicts working memory performance [304]. Regarding pain memories specifically, researchers investigating acute painful stimuli and associated memory of pain found that left hippocampal activity corresponded to remembering higher levels of pain, with no corresponding activity from the right side [306]. However, far more studies will be necessary to understand the role of hippocampal shape and laterality in chronic pain.

From our results, we doubt that a patient's bias in retrospective assessments is a reflection of a general tendency to exaggerate or catastrophize, due to the finding that other memories of the study were not biased in the same direction and to the fact that PCS, PSQ, and PASS were not correlated with discrepancy or shape displacement. Our results also suggest that the memory bias seen here is likely not specific to pain or negative memories, but instead might be more generalizable to highly salient or strongly valenced memories along a reward-punishment continuum. We showed that memories of pain were anti-correlated with memories of monetary compensation received and that both biases were related to loss-aversion scores; furthermore, shape displacement of the posterior hippocampus mediated the effect of loss aversion personality traits on the extent of pain memory bias. While the relationship between loss aversion and memory of pain or reward is complex and not well understood, research has shown that people often underestimate not only the amount of money they earned in reward tasks, but also the

number of times they receive money, indicating that people downplay monetary gains in general [307]. Additionally, we have previously reported that CBP patients show aberrant behavioral loss aversion with increased gain sensitivity in a gambling paradigm [83], suggesting that loss aversion and the experience of pain are intimately linked through neurophysiological and psychosocial mechanisms. The extent to which experiences and memories in other domains are also embedded in similar neurobiology or personality traits remains unknown and a critical topic for future exploration.

CHAPTER 6: GENERAL DISCUSSION

6.1 Summary of results

Here we show that clinically meaningful placebo analgesia in chronic back pain patients has identifiable neuroanatomical and personality traits, underlying neurophysiological bases, and specific ties to semantic language properties. Limbic volume asymmetry, region-specific gray mater density, and region-specific cortical thickness were all significantly different between responders and non-responders prior to treatment. Psychological traits related to interoceptive emotional awareness and emotional regulation were also shown to differ between responders from non-responders at baseline. Resting state functional connectivity of the lateral frontal cortex with either the sensorimotor community or the PAG also distinguished the two groups. Additionally, language properties of 60 words of interest extracted from participants' narratives in an exit interview were also able to differentiate responders from non-responders at the end of the study, with the semantic relationships of 11 words being significantly different between groups.

We further demonstrate that placebo propensity in an RCT can be predicted before randomization using multivariable models combining neuroimaging data with self-report measures; the two resting state functional connections identified (frontal-S1M1 & frontal-PAG), along with the two questionnaire subscales (MAIA-emotion & ERQ-suppress), together explained 71% of the total variance in response and classified the two groups with 84% accuracy. Likewise, we showed that placebo propensity can also be explained using a combination of words semantically related to the exit interview: the semantic similarity of patient narratives to "awareness", "force", "stigma", and "magnify" explained almost 68% of the variance in placebo response and was also over 80% accurate in classifying responders and non-responders when validating the model.

Finally, we reveal that the memory of chronic pain is systematically biased such that our patients reported more pain than what they actually experienced in the study. This discrepancy was driven by a combination of factors, including key components of momentary pain and mood, the morphometry of the hippocampus, and psychological traits related to reward learning. We show that the shape of the left posterior hippocampus, in combination with the psychometric properties of daily rated pain, explain over

50% of the variance in the reported memory of pain, a result which was reproduced and validated in a separate cohort of CBP patients. Moreover, this bias is stable in time, independent of health status and pain duration, and unrelated to abnormal short-term memory or other kinds of exaggeration tendencies. Interestingly, we also show that pain memory bias is anti-correlated with money memory bias and that both biases correlated to reported loss aversion personality traits, suggesting that our results might generalize to either highly salient memories or memories along a punishment-reward spectrum. Finally, we show that hippocampal shape significantly mediates the relationship between loss averse personality and pain memory bias, further solidifying its role in the mismatch between experienced and remembered pain.

All of these results have significant implications for the medical and scientific communities, as well as key limitations that warrant future investigation. Both of these topics are discussed below.

6.2 Scientific, clinical, and ethical impacts

6.2.1. *Study 1*

As discussed in the study-specific discussion, our results have important implications at a scientific and theoretical level by focusing in on mechanistic approaches to understanding placebo and propensity to respond to a placebo. However, they also have significant potential for clinical utility in at least 3 ways. First is the idea that we might be able to harness the power of placebos for medical use in patients as part of daily practice. For example, if we had a model that could predict with high accuracy that a person would respond to a placebo, doctors could prescribe them a placebo for a certain duration of time. This placebo could take whatever form the patient would feel the most comfortable with (e.g., a pill, a patch, an injection, yoga, meditation, etc) and it could be manipulated to mimic dose effects (e.g. to take 2 pills instead of 1 or to take a pill twice a day instead of once). The implications of this are huge – such practices could substantially aid patients by minimizing the number of harmful side effects experienced and lowering the cost of treatment all while obtaining meaningful clinical amelioration. In the instances where physicians may feel uncomfortable prescribing “only a sugar pill” (more on this below), conditioning paradigms in chronic pain may also prove to be vitally important in clinical placebo utility, such that placebos could be given in combination with active treatments to influence their effects. In a

previous study, the administration of a placebo has acted as a dose-extender of a drug effect that was dependent upon the treatment originally given (the conditioning drug). For example, a placebo provided after repeated administration of non-opioid drugs such as NSAIDS, like aspirin and Ketorolac, produces aspirin- or Ketorolac-like effects, respectively. Likewise, a placebo given after morphine conditioning produces opioid-like effects including not only the reduction of pain but also morphine-like induced adverse events and nocebo responses [115]. Since this initial study, the phenomenon has also been replicated in CB1 receptor antagonists including Rimonabant [101]. In these cases, it's thought that expectation plays a sort of "switch" or "gate controller", modifying pathways and selectively activating either endogenous opioid or cannabinoid systems depending upon prior drug exposure[101]. Importantly, it shows that placebos can be potent analgesics on their own or with other drugs. These results point to a learning effect and a powerful self-manipulation, where the brain and mind is recreating drug effects for its own benefit based on what it has expected and learned). They also highlight the plasticity of treatment response that go beyond just mechanisms of resilience; instead this flexibility can be channeled and trained to focus on specific kinds of self-preservation modes. This thesis provides the scientific foundation to begin such endeavors.

Second, identifying people with a high likelihood to respond to placebo would be extremely beneficial in clinical trials and investigations of new pain medications. As these studies are currently practiced, participants are randomized into either an active treatment group or a placebo treatment group, with placebo treatment viewed as the gold standard control. However, based on the results here and studies reported elsewhere, we could imagine that up to 50% of people assigned to active treatment might be potential placebo responders, meaning that they could respond to both an inactive or active treatment. This suggests that up to 50% of the responses in active treatment groups are actually the equivalent of placebo responses or some sort of interaction between active and placebo responses [139]. Given that there are currently no effective medications for treating chronic pain, it would appear that efficacy of drugs studied in clinical trials decreases after clinical trials when tested in the real world. This implies that efficacy measurements in clinical trials are inaccurate, biased, or both. Outside of contributions from things like fluctuating baseline pain levels [137] or report bias, one possible source of error or bias could be the lack of consideration for these placebo responders in the active treatment arm,

who essentially guarantee that the drug will work in more people. If we could use a predictive model to screen participants before these trials and classify the likelihood of placebo propensity, we might be able to remove them before randomization so that they don't receive treatment. This has the potential to drastically improve accuracy of drug efficacy assessments in these trials.

At the same time, anywhere from 36 to 52% of drugs tested (depending on the condition) fail to show superiority to a placebo[308], and this figure increases to 90% for drugs targeting neuropathic and cancer pain specifically [309]. The reason for this isn't necessarily because the drugs themselves are getting worse; instead, as reported previously in **Chapter 1**, the response to sham treatments is getting stronger, making it hard to prove a drug is more advantageous to a placebo. This leads us to the third potential contribution of **Study 1**'s results – if we can identify reliable psychological components governing and predictive of the placebo effect, we may be able to change the placebo response in clinical trials, perhaps by manipulating it either in the treatment arm so as to have a null effect or in the placebo arm to make the effects lower.

But with all of this clinical potential, we must also consider the ethical implications of the above possibilities. Regarding prescribing placebos or using them to augment active treatment effects, there are concerns about possible harm and consent. The Hippocratic Oath is a declaration made by future practicing physicians during their first year of medical school, and it covers the codes of conduct and bioethical responsibilities expected of doctors in reference to their patients. As a part of this covenant, physicians pledge to use all measures necessary to benefit a patient and to reject harm and avoid injury. They also promise to simultaneously avoid overtreatment and therapeutic nihilism, and state that they will remember they are treating more than a symptom and instead are caring for a sick person whose illness affects others and the economy (entities that their responsibilities as physicians are also related to). Although the placebo effect has been shown to be therapeutically beneficial, at levels that sometimes exceed the standard of care, many physicians feel as if administering placebos to patients is unethical as it might delay patients potent active medications (thus imposing potential harm) *and* it could involve dishonesty and a person's right to know and decide what treatment is best for them (thus breaching ideals of informed consent and autonomy).

Let's unpack these concerns. As discussed in **Chapter 1**, placebos are often seen at best as manipulators of psychology (but not pathology) and at worst as shams with little to no benefit. With this mindset, it is easy to see why some clinicians may feel uncomfortable prescribing placebos in place of medicinal treatments. Not only could they be harmful if they don't help the patient, in this case possibly delaying analgesia and causing prolonged suffering, but they also could be viewed as promoting (instead of combating) therapeutic nihilism, the idea that cures do more harm than good and that people should be encouraged to let the body heal itself [310]. Unfortunately, there are no real statistics about how often placebos work or don't work for participants – all we know is what I have already reported in this thesis (namely the placebos can achieve clinically meaningful analgesia and that placebo effects are rising in clinical trials). As for encouraging the body to heal itself, I am not convinced that this is necessarily a harmful attitude when promoted under the proper conditions. Of course, some diseases or symptoms do not seem to be good candidates for placebo interventions or at the very least less subject to as potent placebo effects as those seen in pain. Antibiotics for serious bacterial infections or chemotherapy for stage 2 and 3 cancers, for example, would not warrant ethical administration of placebo[134, 135], as these pathologies are life-threatening, have a higher likelihood of responding well to current treatments, and whose benefits likely outweigh the adverse events experienced. In contrast, chronic pain medications are not guaranteed to work and when they do, many only have short-term efficacy and/or serious side effects including dependency, thus questioning whether the limited benefits are greater than the risks. In fact, using placebos might help more than they hurt. For example, if placebos were administered in combination with stronger analgesics through conditioning paradigms, they would extend beneficial effects while simultaneously lowering harmful ones, such as addiction. Additionally, it might be possible that intermittent exposure to conditioned placebo responses could lower or help decrease tolerance to certain prescription drugs, since they would not be in a person's system at all times. Furthermore, we are already experiencing an epidemic of over-treatment and unnecessary medical care in the US. Looking at Medicare patients alone, as many as 42% of them per year receive at least one of 26 tests or treatments deemed useless or even harmful by a number of scientific and professional organizations[311]. Unnecessary care like this was estimated by the IoM to account for 30% of healthcare spending in 2010, equating to over 750 billion dollars a year (more than the nation's budget for K-12

education) [311]. There is evidence that a good chunk of these costs comes in relationship to pain treatments (medicinal or otherwise). For instance, one study found that for the 8 years between 1997 and 2005, national health care expenditures for back pain patients specifically increased by almost two thirds and at the same time, population surveys showed no improvement in the amount of pain reported by these patients [311]. Under this light, prescribing a placebo as a first try or as a compliment to already-existing medications appears to be ethically responsible, limiting deleterious health costs (physical and monetary) for the benefit of the patient and essentially promoting many of the aims of the Hippocratic oath (e.g., treating illness, combating overtreatment, and promoting the reduction of national economic burdens).

Regarding issues of autonomy, it is of course critical and necessary that informed consent would take place prior to usage of placebos in medical care. Currently, as many as 50% of practicing physicians report routinely prescribe placebos (or drugs that are viewed to be essentially similar to placebos, such as over-the-counter headache medicines and vitamins) without the knowledge of their patients [312, 313]. This is troubling, because without informed consent, trust in the medical system could be undermined, the patient-physician relationship compromised, and potential harm (including nocebo effects) could occur [314]. However, this does not mean that placebos should not be prescribed, nor does it necessarily mean that placebo usage would have to be limited to open-label (not deceptive and not concealed). Instead, placebos could be effective if the patient understands what a placebo is, knows it's possible they might receive one, agrees to this possibility, but cannot identify it and does not know the precise timing of its use. A physician could enlist patients' cooperation by explaining why using a placebo might be of benefit to the patient (e.g., obtain pain relief with less side effects and/or evaluate and compare effects of different medications) and obtain evidence of informed consent prior to using placebos. Such documentation could be built into already existing "medication contracts" or "pain treatment agreements" that pain physicians currently use to monitor patients use of certain drug classes, or it could be added as a new document as part of a physician's practice so as to provide evidence of these discussions.

More research is needed to investigate whether placebos work better when used covertly and under methods of deception or whether they are just as if not more efficient with honest and open administration. In general, open-label drugs tend to show stronger effects than concealed label drugs in a

variety of ailments, including pain relief. A 2001 study by Amanzio and colleagues [315] analyzed the effects of covert and overt administration of four widely-used pain killers (each with distinct mechanisms of action). Across all 4 medications, they found that the analgesic dose needed to reduce pain reports by 50% was higher with covert infusions than in overt infusions, proving not only how powerful the reinforcing effects of expectation can be, but also showing how generalizable these affects are, as they occurred across multiple classes of drugs [315]. Regarding placebo effects specifically, there is quite a bit of evidence that open-label placebos also work (meaning that a subset of individuals will still get relief even if they know they are receiving an inactive medication). One study showed that open-label placebos are superior to no-treatment in a trial monitoring patients with irritable bowel syndrome (IBS) [130, 155, 316], and another very recent indicated that adding open-label placebo to back pain participant's usual treatment decreased both pain and disability reports more than the treatment alone study [317]. Thus in some instances where placebos might be used clinically, consent might not even be an issue. Therefore, due to the prevalence of placebos, increasing evidence of their effectiveness, and physician's current use of them in practice, it's not only implausible to dismiss placebo responses as irrelevant to health, pathology, and recovery but in many cases, unethical to do so.

Of course, there are also ethical and social considerations to keep in mind regarding placebo responder identification in clinical trials. Take, for example, the idea of creating screening processes that involve similar predictive models identifying people with high likelihood of responding to placebo. Some might argue that we have an ethical responsibility of eliminating these individuals from trials or at the very least from active treatment groups so as to make drug efficacy assessments more accurate (which would benefit society as a whole, including pain patients). However, some might also question whether we have the right to do this and essentially deny these individuals access to pain relief. The other idea presented above was influencing responses to essentially make placebo effects null across all trial arms. Again, this would make trials more accurate while also possibly denying a certain subset of participants pain relief. Moreover, the idea that we could identify key parts of individual's psychology and manipulate their body's ability to respond or not to a medication makes some uneasy – again, do we have this right? In this case, one could argue that we could also exploit placebo responses in both directions depending on context, therefore providing increased pain relief to a subset of individuals too. Additionally, we regularly use

medicine or procedures to change what the body does, even lower it's actions (e.g., Metformin is a diabetes medication that reduces the body's natural ability to elevate blood sugar); what makes manipulating the mind that different at the end of the day than manipulating other aspects of the body?

And then there are other considerations that we simply do not have time to cover here. As one example, if placebos are eventually accepted as possible treatments, how much would they cost (how much could a drug company ethically charge for an inactive medication)? As another example, if we can predict future responders, would there be any consequences from an insurance stand-point (e.g., if this information was provided to insurance companies, could they deny patients from certain, more expensive or brand-name medications or could they deny patients certain classes of medications all together?). We are still years away from understanding the placebo effect in its totality, and even farther away from really utilizing or relying on these predictive methods for real world clinical applications. However, all of these implications are important to think about as they might significantly impact how we run clinical trials and how we practice medicine.

6.2.2. *Study 2*

The results of this study are important for both practical and clinical reasons. Given that semantic language properties correlated with questionnaires and brain parameters, it may be possible to use language as a tool either in combination with these methods or as a reasonable substitute for when they aren't feasible for monetary or time-related reasons (such as fMRI). For example, the cost of an hour of scan time can be upwards of $500, and analysis of imaging data, while semi-automated depending upon the stage, still relies on trained human eyes to visually identify and correct quality control mistakes, as well as troubleshoot and ultimately make decisions about the data as it's analyzed. The time commitment involved is another important factor – in addition to training a tech or students to run the scanner, one must also wait for preprocessing and quality control of the images, procedures which can take hours to weeks to finish; furthermore, some analyses may also take days to run and complete depending upon computational resources and infrastructure. Monetary and time constraints like these simply aren't realistic for most clinical decisions regarding pain, as patients expect to have a treatment regimen by the end of a visit and don't want to pay additional out-of-pocket medical costs. These constraints are also

certainly not sensible in clinical trials that are trying to recruit as many participants as possible in the shortest amount of time. In contrast, the analyses presented here can all eventually be easily automated, and other than participant compensation or transcription fees, the methods for capturing the data are essentially free. Furthermore, there is no longer an issue of time. After the initial corpus dictionaries have been developed and the scripts which perform LSA have been tested and agreed upon, all that is left to do is to upload the interviews and press a button to start the preprocessing and analysis, a process that can take seconds to minutes. Moreover, since physicians and clinical coordinators must already take patient medical histories as part of their jobs, a short interview could easily be worked into this daily practice almost seamlessly. While more investigation is needed to be able to say which questions in **Study 2**'s interview were the most important (e.g., which sections had the most words that were semantically similar to "awareness", "force", "stigma", and "magnify"), we might be able to reduce the interviews to one- or two-questions that would be under 5 minutes, something convenient from both a physician and an RCT perspective.

The ethical implications of using language to detect illness or propensity to respond to certain drugs are currently unknown and under-researched. A simple Google search for "ethics of semantic language analysis" provides zero meaningful hits; in contrast, "ethics of neuroscience" immediately produces numerous results about neuroethics, social and legal issues regarding using neuroimaging data in courtrooms, and ethical implications about being able to "read minds". Thus it is still too early to be able to gage the social and ethical impact of being able to predict future behaviors or diseases from linguistic analyses. However, one must wonder about issues of privacy and consent. For example, would it be possible to analyze the social media posts (from Twitter or Facebook) of potential clinical trial participants to determine placebo response propensity prior to randomization? While these data sources are public, is it ethical for coordinators to collect this information? Additionally, since language represents our internal private thoughts, which we think we can hide through certain words or avoiding certain topics of conversation, how "informed" is informed consent – participants may agree to an interview and might vaguely understand what the interview is used for, but do they really understand the extent to which language might inform researchers about their biology or future actions?

Language analyses also raise questions about the importance of subjectivity and may create ethical dilemmas when participants' reported experience does not match data obtained with other outcomes. For example, we already know from **Study 3** that memories of previous pain are inaccurate and mismatched with the actual pain or analgesia experienced in the moment. In the interviews used in **Study 2,** participants randomized to a treatment group were asked about the study medication – did they think they received active treatment or placebo and why? What if these retrospective assessments revealed by the interview do not match the daily pain rating data and in turn do not match the permutation stratification used? Which is the more appropriate or correct "responder" – the person reporting it worked (even if their pain intensity and mood remained the same for the duration of the trial) or the person who said it didn't work (whose pain intensity decreased by over 20% during one of the treatment periods)? Additionally, what if semantic language analyses predict response propensity, correlate with pain relief and personality scores, but do not match the opinion or report of the patient? How will we interpret these results? Are our metrics flawed (e.g., is intensity not the best metric to capture placebo analgesia or meaningful response) or is the patient lying, and if the latter, is it ethically permissible to question someone's pain experience, which at the end of the day is ultimately subjective and unknowable? Will language analyses and interviews create new questions and philosophical debates along these lines? Or instead, will such analyses finally be able to bridge methodological gaps between outcome measures, creating links between perceived disability, quality of life, and pain intensity where there were none before to better explain such mismatches? Only time will be able to tell.

6.2.3. *Study 3*

Our findings have important implications for understanding and treating pain in clinical settings. Given the reliance on self-reported numerical ratings of pain to influence the type and duration of treatment in chronic pain patients, our results emphasize that these retrospective measures are inaccurate and often at least 1-2 points higher on a VAS scale than what patients actually experienced. Not only was this bias present in the vast majority of patients, but additionally, the magnitude of the discrepancy corresponded to thresholds often utilized in determining clinically meaningful interventions (~20% reduction in pain intensity). These results suggest that using alternative measures of pain intensity

(such as daily ratings) might provide additional and important information when evaluating and caring for pain patients. Additionally, we show that patients' memories of pain, while flawed, are correlated to their average actual pain. Moreover, we also show that their memories are influenced not only by a well-established cognitive bias present in the majority of people but also by a distinct difference in hippocampal anatomy that is not influenced by other confounds, indicating that this bias is neurologically determined. These findings thus underscore the ethical obligation of physicians and clinical researchers to trust their patients' and participants' recalled pain experiences while understanding that they might be slightly higher than experienced. Our results also highlight a more philosophical question, beckoning investigators and clinicians alike to ask themselves just *who* are we trying to treat or trying to study – the person experiencing the pain *or* the person who will look back on this pain and evaluate it? Ethically speaking, do we try to make spontaneous pain lower (even if the memory of it will still be higher) or do we try to make the memory of the pain lower (in turn possibly causing people more experienced pain)?

Critically, our model combining hippocampus morphometry and pain ratings was able to accurately predict the intensity of retrospective recalled pain, representing a potential tool for future clinical use, such as pre-selecting patients with high self-report accuracy for participating in RCTs in order to minimize bias in efficacy measurements. Furthermore, better understanding the phenomenon of over-estimated pain recall in chronic pain populations might impact decision-making in other clinical settings, from critical events like treatment commencement or dosage amount, to more simple applications such as evaluating inclusion and exclusion criteria based on self-reported pain-ratings.

6.3 Potential limitations of the current studies

6.3.1. *Study 1:*

In this study, we show that there are anatomical, psychological, and functional brain connections that predispose a subset of CBP patients to respond to placebo, and furthermore, that there are some functional biomarkers related to pain and sensation which appear more state-like and others related to personality that seem more state-like. Moreover, we present various multi-parameter brain, questionnaire, and combination models that are able to explain large amounts of variance in placebo response and predict this response with relatively high levels of accuracy. However, there were quite a

few things that we didn't do or didn't look at that might be seen as potential limitations to the trial or analyses. First, due to a large number of screen failures, we ended up losing a lot of potential participants and thus our numbers were not large enough to set aside a validation data set to directly test the models built. Therefore, we do not yet know how well they will perform in a separate cohort. Because we mixed responders from treatment 1 and treatment 2, we do not know the effects of time and learning on behavior, nor do we understand what qualities make someone consistently respond to placebo, as we were only interested in general predictors and thus did not directly test stability of response in time. Importantly and related, we have no idea how generalizable our results are outside of our study – are these models specific to CBP or will they also apply to other chronic pain conditions, and are they specific to kind of ritual or treatment used (a pill) or will they also apply to additional routes of treatment administration or kinds of treatment (injections or patches)? Since we didn't test generalizability, it is also unknown how much the placebo effects seen here are able to be manipulated by changing aspects of the study (like color of the pills, number of scans, or wearing white lab coats).

Additionally, while we included a proper no-treatment arm as an important and useful control, we could have randomized at least twice as many people into this arm so that our numbers of responders and non-responders would have been better matched to the numbers of potential responders and non-responders who didn't receive treatment. Additionally, the no-treatment arm could have served as a validation set if would have later provided them with placebo pills and followed them for at least two weeks (something we did not do here). Another issue related to the no-treatment group was that the research team was aware that they were not receiving treatment, meaning that we could have inadvertently introduced bias prior to analysis in how we interacted with these individuals; for these reasons, if I were to do it again, I would have had an independent assessor not involved in any other aspects of the study administer medications and collect adverse events so that the rest of the study staff could have remained blinded. I also think this would have help made the environment better controlled in general, with certain people having certain roles, thus increasing the consistency of the study activities between subjects.

Fortunately, many of the issues mentioned above will be addressed either in future analyses of the data or in a second phase of data collection (explained below in future directions).

6.3.2. *Study 2:*

In this study, we used exit interviews to study the language of CBP patients to test if how they talked about their pain was related to whether or not they responded to a placebo while in the trial. We found that the semantic similarity of 11 words with all participants' interviews significantly differentiated those who responded from those who did not, 4 of which explained almost 70% of variance in the response. 6 factors created from words of interest were also related to response, correlating with previously seen psychological and biological markers of propensity and identifying additional neural signatures of response from a scan that occurred 6 weeks prior to the interview. While these results suggest that semantic language analysis can be a powerful tool, as this was an initial investigation of this technique, our study had limitations (some logistical and others more philosophical).

First, it's important to reiterate that our results are inconclusive in that we completed the interviews at the end of the study. While we had good reason to do this (see **Chapter 2**), this means that we don't know if our findings represent predictive elements of language (that would also be present to some extent at the start of the study before treatment commencement) *or* if they are consequences of the placebo effect (that only happen after treatment and response). Another limitation is that we only studied language cross-sectionally, meaning that we don't know if these language parameters are binary (i.e., someone either has and displays these properties or does not) or if they are instead dynamic and fluctuate as a result of the amount of pain or analgesia someone is experiencing. While the factors did not correlate to the amount of pain participants reported at the final visit, this doesn't mean that they didn't change throughout the study as a function of the placebo effect (or lack thereof). Additionally, ldue to our relatively small sample size and limited clinical population, we do not know how generalizable our results are, since we did not set aside some of the data for validation purposes (like **Study 1**, here again the no treatment group would have been immensely powerful if we gave them placebo pills after study completion). Further research will be needed to test the interview and semantic language factors on new datasets, including placebo response under different rituals or conditions, as well as in different pain populations (such as OA or fibromyalgia patients). Likewise, we would need to investigate whether the results would change as a function of native language usage or cultural upbringing, in addition to pain

condition or trial contexts. This latter limitation may not be as much of a concern, however, given that a recent study has shown that semantic meaning is stored in specific neural patterns that are relatively consistent across individuals despite differences in experiences, brain activity, cultural identity, and education [268]. However, some research suggests that pain language in particular may show sensitivity to these dimension; previous work has shown that there are clear gender differences in pain language in a recalled pain narrative task [175, 318], and that emotional discourses show complex differences across cultures [319].

Second, many of the analyses were initially driven by relatively subjective decisions. Although choices of the 60 words were determined apriori from the literature and results of **Study 1**, at the end of the day, they were subjectively deemed as important out of nearly 80,000 possible words; not only might these decisions be biased (causing the resulting models to possibly overestimate the amount of variance explained or accuracy measured), but there may have been stronger words that would have better stratified responders from non-responders that were never considered. Instead, we could have approached the data with a more "black-box" methodology where we either employed machine learning to find patterns in the data without restricted inputs *or* taken the brain biomarkers (functional and anatomical) and used them to search the semantic space for language differences. The two hypothesis-free methods (LSA and "blackbox" approach) that were attempted in this study did not provide statistically robust differentiation between groups. This means that we will likely need to better hone our data-driven approaches in the future, as well as perhaps implement additional preprocessing and cleaning steps to further reduce the remaining sources of noise in our data, due to things like heterogeneity of the population, cultural differences, over-use of colloquialisms, or large variance in responses and experiences. Additionally, we might also consider using a different corpus or combination of corpora instead of TASA in future LSA attempts; for instance, we might be able to create our own corpus consisting of pain-specific journal articles, books, ethnographies, and/or medical records as a starting point (Lascaratou), or use a dictionary that accounts for polysemy (such as Wordnet) in order to better understand our findings (without this, we can't determine if the word "back" means body part, backwards, or previous in time, resulting in vague or confusing interpretations)[320].

While quantitative approaches like these are more efficient and perhaps less biased overall then

other methods, they can sometimes lead to “fishing expeditions” where one can look for and often find any answer in the data, even though it might be meaningless. Along these lines, we could have utilized a qualitative approach as a starting point for data analysis to narrow the focus in a more systematic way. For example, we could have coded the data to identify important themes or patterns and let these results inform the future quantitative analyses. This kind of method is rooted within grounded theory, which provides a list of flexible strategies for focusing and expediting data collection and analysis [321]. These strategies include simultaneously collecting and analyzing data (so that each can inform one another), pursuing emergent themes early on in analysis, constructing categories that explain and synthesize social and cognitive processes in the data, and integrating findings into a framework that specifies the causes, conditions, and consequences of the studied phenomenon [321]. Additionally, we really have no idea what language parameters are distinct or important in this participant cohort at baseline. It would be pertinent to first identify how CBP patients differ from healthy pain-free controls in their language parameters and from this information, then investigate how these parameters change as a function of placebo response. Such an analysis could be done with a combination of qualitative and quantitative approaches.

Third, finally, and perhaps most importantly, while language has been used to study various pathologies and personalities prior to us investigating placebo analgesia, the assumption that language can be used to study pain in the first place might need additional consideration. Most feeling-states are difficult to express at baseline, and this is something that is not unique to the feeling of pain as people generally struggle to either translate strong sensations into words (e.g., parental love, death of someone close to them, or experience of an orgasm) or have difficulty using language to adequately express their emotions, thus relying on an overuse of metaphors or recycled clichés. However, given that pain is both a physical and emotional state, the failure of language may be more apparent in the context of pain, and “painful bodies might be uniquely and especially indisposed to acts of communication” [249].

Virginia Woolf once argued that there was “a poverty of language” with regards to pain, commenting that:

> “English, which can express the thought of Hamlet and the tragedy of Lear, has no words for the shiver and the headache…The merest schoolgirl, when she falls in love, has Shakespeare and Keats to speak her mind for her, but let a sufferer try to describe a pain in his head to a doctor and language at one runs dry.” [249]

This deficiency in pain language is due, in part, to the fact that talking about pain has both social and psychological implications and consequences, both of which may constrain the way that people talk about their experiences (e.g., to downplay or exaggerate, depending upon the context). Pain language is, like its subject, profoundly socially alienating. When experiencing a chronic illness, acts of communicating pain may be emotionally painful in and of themselves, and there is always the risk that talking about how someone is suffering may be harmful and profoundly depressing instead of helpful or nurturing [171]. Although some people like telling their pain stories, many often seek solitude and silence instead of talking, therefore making pain a secret; similarly, some cultures actively discourage talking about suffering as complaining is seen as rude or a sign of weakness. In addition, pain words are, for the most part, inherently internally focused since it is such a private and subjective experience. This makes pain language also hard to communicate, since outside of acute pain stories (like burning a finger on the stove, stubbing a toe, or getting papercut), many people have not experienced prolonged pain and thus can only rely on words to guess at what a person is experiencing, even if those words do not give justice to this experience. This lack of previous experience with chronic pain or illness can in turn contribute to lack of empathy from listeners, whether they are family members or physicians. As Elaine Scarry wrote, "To have pain is to have certainty, to hear about pain is to have doubt" [322]. Because pain is un-sharable in this way, language ultimately cannot capture its essence. Moreover, there is no specificity of words to describe pain qualities outside of academic or research endeavors (like the MPQ) and no general consistently between stories, as everyone's experience is different (both medically and culturally). Some patients may describe neuropathic pain as "burning", others as "tingling", and others as "shooting" without consensus, and for every person who speaks of "flames", there are 5 others for whom "it just hurts". This lack of reliable words can in turn cause many physicians or loved ones to look at chronic pain patients as "unreliable narrators" [322], further contributing to unrelieved suffering.

In addition to social barriers and an overall lack of richness or consistency in pain words, chronic pain often either does not always allow for language to occur or stops language altogether. This presents another potential problem when using interviews to study or predict pain behaviors. One way that chronic pain makes pain stories untellable is that they give no space or time for the person suffering to reflect upon their pain; instead the pain is immediate, continuously present, and extremely distracting, which

makes thinking about it or talking about it exhausting and difficult [168, 170]. Additionally, people often feel a disconnection between their mind and their body that is in pain ("what is this thing and why is it betraying me?") [290], and this in turn makes people want to ignore their bodies and not focus on their pain, again repressing language. Furthermore, "physical pain does not simply resist language but actively destroys it" [322]. This means that pain might take away one's ability to communicate in written or typed form (such as in rheumatioid arthritis), or it might be so intense at a given moment that it literally takes one's breath way. When someone's hurt and in incredible pain, one often witnesses language being destroyed; patients may utter a monosyllable or cry, and words are replaced with groans and moans, winces or screams. The more intense or present the pain, the more we resort to physical or basic communication as opposed to verbal or written communication to express our pain (meaning that interviews might not always catch these elements of the pain experience).

These limitations highlight how early we are in the study of language analytics and how much more there is to trouble shoot and consider in study and/or interview designs, especially in regards to the topic of pain. However, this doe not mean that language analyses will be in vain, nor does it mean that the findings presented here are less important. Instead, our results demonstrate the potential utility of language as a tool for quantifying pain and analgesia, as well as predicting future complex behaviors. Moreover, these findings also show the power of combining multiple methods together in one study, where each method informs the other. This is particularly true with language analyses that aim at understanding the meaning of passages or the content of narratives. Identifying content is important, but this by itself does not explain *why* this content is there in the first place; such questions can only be answered by also including data and analyses from biological sciences, psychology, and other social sciences to more fully capture the physiological and social mechanisms underlying semantics of language and the pain experience.

6.3.3. *Study 3:*

In this study**,** we show that pain memory bias in chronic pain patients is determined by specific properties of their daily experience in combination with the shape of their left hippocampus. Because this bias is stable in time and because hippocampal morphology does not change following the development

of chronic pain, we argue that this kind of discrepancy is trait-like and possibly generalizable to other kinds of memories regarding aversive and rewarding stimuli. However, because we lack a healthy control group with similar recall assessments (for example, of previous monetary compensation), we cannot definitively say that this bias is not somehow affected by the presence of chronic pain. As discussed in this thesis, the lived experience of being in long-term, persistent pain not only affects the functional circuitry and anatomy of the brain but also results in accompanying changes in perception, cognition, and behavior. It is possible that having chronic pain further increases the direction or magnitude of pain discrepancy compared to that of otherwise healthy individuals, for example. A previous study by Tasmuth and colleagues [323] found that individuals who developed chronic pain after breast cancer surgery ended up reported significantly higher memories of post-operative pain than individuals who remained pain free, begging the question of whether increased post-surgical pain can lead to chronicity *or* whether instead chronic pain results in an increased tendency to overestimate past painful experiences. Another study showed that while most chronic pain participants displayed an overestimated pain memory, the extent of this bias was dependent on the condition, with cervical and low back pain patients being more accurate in their recall than individuals with chronic headaches or abdominal pain [207]. Therefore, the idea that chronic pain might still contribute uniquely to memory distortion is something we need to keep in mind.

Another possible limitation in **Study 3** was that we did not consider how motivation might play a role in the memory bias seen in our study, something that should be considered. Incentives for individuals with chronic pain may be different than those of pain-free participants, especially in the context of a clinical trial involving treatment for pain, and therefore differences in motivations may influence reported memory of pain. While the present paper only presents data from the initial first two weeks of a clinical trial, many of these participants continued in the study and went on to be randomized into a no-treatment, placebo treatment, and active treatment group (**Study 1**). Although the study was blinded, a subset of participants may have, through hope or therapeutic misconception, believed that reporting higher pain than what they actually experienced might lead to an increased probability of being treated. Likewise, since we collected recalled pain measurements for the duration of the study, some participants may have desired to please the researcher or may have worried that their time in the study or future studies might

be negatively impacted if they didn't report analgesia, opting to report lower pain than what they actually experienced later in the study (as an example, see [324]). The finding that the extent and direction of memory bias was stable around a year after the study suggests that neither of these things happened. However, we cannot rule-out the possibility that in addition to recall bias, there may also have been a response or report bias that interacted with participants' memories to influence the magnitude or direction of discrepancy. We use the terms *response* and *report* here to stress that while a participant's recall of their pain may already be unconsciously biased (from factors like peak pain, average pain, and ending mood), this memory may go through additional downstream, more conscious filters before actually being reported to study staff. Because of this, discrepancy scores could be a mix of multiple kinds of biases, an idea that we did not directly test here but could explore in the future.

6.4 Future directions and concluding remarks

We plan on further exploring and expanding upon the results of all three studies in the near future. For **Study 1**, neuroimaging data will next be studied longitudinally to explore how the networks identified change in time with response to placebo. At this time, we will also be able to further differentiate types of participants, moving beyond responder and non-responders to also include early responders (treatment 1), late responders (treatment 2), and consistent responders (those whose pain decreases or remains down for both treatment periods). This additional differentiation will allow aid in potentially identifying additional anatomical and functional mechanisms behind response reliability and stability. Moreover, **Study 1** was part of the first of two phases in a grant. In phase 2**,** which is currently recruiting 140 CBP patients, we will be able to directly test some of the findings reported here. All participants in this second phase will complete an abridged version of the questionnaire battery and finish an MRI scanning session at the beginning of the study. We will then extract the relevant scores and neuroimaging data to input into our models and calculate a likelihood of placebo propensity; based on this calculation, we will then assign participants into a responder versus non-responder group and follow them longitudinally to test the validity of the phase 1 models. Additionally, we will also be investigating the interaction between placebo effects and active treatment effects; once stratified based on response likelihood, patients will be randomized into either a placebo or active treatment group and followed for 6

weeks. The idea is to see if responders who receive active treatment will have a response equal to or greater than those who receive and respond to placebo, and if greater, to investigate the mechanisms behind this interaction (e.g., are they additive or nonlinear effects).

The results of **Study 2** will be tested for validity in the form of a shorter interview at the beginning of phase 2; more specifically, at least the 4-word model will be assessed in identifying responders from non-responders before randomization. This will help address one of the remaining questions unanswered in this thesis – namely, are the semantic similarities to the interviews due to predictive qualities of people's thoughts and languages (causal) or instead do these similarities arise as a result of having responded to placebo (consequential). We will also be tracking people's language use in time by having them describe various pictures for 2 minutes at three time points during the study (before, during, and after treatment). Through this task, we are aiming to capture language that is not related to pain (and less biased by questions asked), as well as potential changes to language that might be able to pick up changes in pain (such as analgesia from treatment). These changes were not investigated here and will hopefully provide new and useful information that will also show that language can be used as a surrogate for other clinical markers.

The results from **Study 3** have already influenced numerous studies in the lab and we currently have 3 clinical trials that are also collecting retrospective pain reports and daily pain ratings for the duration of the studies. With this information, we will be able to see if the extent of memory bias differs depending on the pain cohort (e.g., CBP versus SBP or chronic pelvic pain), and we will be able to test whether our hippocampus+pain rating model validates in these separate cohorts as well. We will also be able to test whether memory bias plays a role in the occurrence or magnitude of placebo response and/or influences the content and semantic relationships in patients narratives.

Finally, more investigation is needed into the extent to which the results of all 3 studies influence one another. As one example, we know based on previous research and **Study 3** that self-report measures are highly subjective, "colored by incentives, decision biases and heuristics, and cultural display rules" [282]. While we only studied the bias related to reported memory of pain intensity, our study heavily relied on self-report in the questionnaire battery provided to participants, as well as in the exit interview. Does memory bias play a role in how additional questionnaires were answered or in the

semantic language properties of patient narratives, and if so, can we detect this and relate these back to hippocampal morphometry? As another example, we show in **Study 1** that different outcome measures produce different results when measuring the placebo response (with intensity measures and numeric scales correlated to brain parameters of response but measures associated with pain qualities or affective components showing no relationship). Research findings from other labs also echo these results: it has been previously shown that reductions in pain-related brain processing was unrelated to individual differences in self-reported placebo analgesia in healthy subjects [282, 325], and that expected and recalled placebo efficacy were significantly correlated whereas concurrent (momentary/experienced) efficacy did not correlate with either of these measures [326]. The magnitude of the placebo effect has also been shown to significantly differ depending on the type of self-report measurement used [327]. Therefore, it would be prudent for us to test the relationship between self-reported pain relief in the interview and those in the various PROs. Investigations like these might not only better elucidate mechanisms of placebo response but may also show that different outcome measures reflect different components of the placebo effect, a possibility which has been proposed but not yet identified. Finally, given the amount of data we will have with phase 1 and phase 2 combined (in addition to all we know from previous studies in the lab), we might be able to create a multidimensional space of placebo response propensity. Through combining the brain anatomy, personality traits, psychological states, stable and unstable resting state connectivity, semantic properties, and memory biases associated with response propensity *with* other data related to healthy history, chronic pain type, duration of pain (etc), we might be able to map each person uniquely into this hyperspace to better understand how all of these factors interact with one another on an individual level and on a group level.

What remains clear to me after completing this dissertation is that the study of pain transcends the boundaries of a single field, as does the study of the placebo response – both phenomena rely on diverse and complex biological, psychological, social, and contextual mechanisms. Therefore, we need to break down current academic borders and form interdisciplinary collaborations across many fields, including but not limited to neuroscience, psychology, computer science, linguistics, anthropology, and biomedical ethics. It is only through fostering these unique cross-disciplinary relationships that we will be able to better tackle these research problems, abolish chronic pain, and harness the power of the mind.

REFERENCES

1. Jackson, M., *Pain: The science and culture of why we hurt*. 2002, UK: Bloomsbury Publishing Plc.
2. Merskey, H., *Logic, truth and language in concepts of pain.* Qual Life Res, 1994. **3 Suppl 1**: p. S69-76.
3. Descartes R, C.E., Lessing J. Rosenwald Collection, John Davis Batchelder Collection *Renati Des Cartes Specimina philosophiae, sev Dissertatio de methodo recte regendae rationis & vertitatis in scientiis investigandae: Dioptrice et Meteora. Amstelodami,: Apud L. Elzevirium.* 1644: Library of Congress.
4. Nolte, J., *The Human Brain: An Introduction to Its Functional Anatomy*. 6th ed. 2009, Philadlephia, PA, USA: Mosby Elsevier.
5. Felten, D.L., & Shetty, Anil N., *Netter's Atlas of Neuroscience*. 2nd ed. 2010, Philadelphia, PA, USA: Saunders Elsevier.
6. Cramer, G.D., & Darby, Susan A. , *Clinical Anatomy of the Spine, Spinal Cord, and NAS*. 3rd ed. 2014, St. Louis, MO USA: Elsevier Mosby.
7. Mayer, D.J. and D.D. Price, *Central nervous system mechanisms of analgesia.* Pain, 1976. **2**(4): p. 379-404.
8. Fields, H.L. and M.M. Heinricher, *Anatomy and physiology of a nociceptive modulatory system.* Philos Trans R Soc Lond B Biol Sci, 1985. **308**(1136): p. 361-74.
9. Heinricher, M.M., S. McGaraughty, and D.A. Farr, *The role of excitatory amino acid transmission within the rostral ventromedial medulla in the antinociceptive actions of systemically administered morphine.* Pain, 1999. **81**(1-2): p. 57-65.
10. Ossipov, M.H., G.O. Dussor, and F. Porreca, *Central modulation of pain.* J Clin Invest, 2010. **120**(11): p. 3779-87.
11. Levine, J.D., N.C. Gordon, and H.L. Fields, *The mechanism of placebo analgesia.* Lancet, 1978. **2**(8091): p. 654-657.
12. Levine, J.D., et al., *The narcotic antagonist naloxone enhances clinical pain.* Nature, 1978. **272**(5656): p. 826-7.
13. Head, H., *Studies in Neurology*. 1920, London: Hoddler & Stoughton.
14. Apkarian, A.V., et al., *Human brain mechanisms of pain perception and regulation in health and disease.* Eur J Pain, 2005. **9**(4): p. 463-84.
15. Geha, P.Y. and A.V. Apkarian, *Brain imaging findings in neuropathic pain.* Curr Pain Headache Rep, 2005. **9**(3): p. 184-8.
16. Buchel, C., et al., *Dissociable neural responses related to pain intensity, stimulus intensity, and stimulus awareness within the anterior cingulate cortex: a parametric single-trial laser functional magnetic resonance imaging study.* J Neurosci, 2002. **22**(3): p. 970-6.
17. Craig, A.D., *A new view of pain as a homeostatic emotion.* Trends Neurosci, 2003. **26**(6): p. 303-7.

18. Hofbauer, R.K., et al., *Cortical representation of the sensory dimension of pain.* J Neurophysiol, 2001. **86**(1): p. 402-11.
19. Schnitzler, A. and M. Ploner, *Neurophysiology and functional neuroanatomy of pain perception.* J Clin Neurophysiol, 2000. **17**(6): p. 592-603.
20. Bushnell, M.C., et al., *Pain perception: is there a role for primary somatosensory cortex?* Proc Natl Acad Sci U S A, 1999. **96**(14): p. 7705-9.
21. Tracey, I. and P.W. Mantyh, *The cerebral signature for pain perception and its modulation.* Neuron, 2007. **55**(3): p. 377-91.
22. Mac, L.P., *Psychosomatic disease and the visceral brain; recent developments bearing on the Papez theory of emotion.* Psychosom Med, 1949. **11**(6): p. 338-53.
23. Ehrsson, H.H., et al., *Threatening a rubber hand that you feel is yours elicits a cortical anxiety response.* Proc Natl Acad Sci U S A, 2007. **104**(23): p. 9828-33.
24. Tsakiris, M., et al., *Neural signatures of body ownership: a sensory network for bodily self-consciousness.* Cereb Cortex, 2007. **17**(10): p. 2235-44.
25. Chua, K.S., et al., *Pain and loss of function in head and neck cancer survivors.* J Pain Symptom Manage, 1999. **18**(3): p. 193-202.
26. Zhang JM, A.J., *Cytokines, inflmmation, and pain.* Int Anesthesiol Clin, 2007. **45**: p. 27-37.
27. Baliki, M.N. and A.V. Apkarian, *Nociception, Pain, Negative Moods, and Behavior Selection.* Neuron, 2015. **87**(3): p. 474-91.
28. Phillips, K. and D.J. Clauw, *Central pain mechanisms in chronic pain states--maybe it is all in their head.* Best Pract Res Clin Rheumatol, 2011. **25**(2): p. 141-54.
29. Elman, I. and D. Borsook, *Common Brain Mechanisms of Chronic Pain and Addiction.* Neuron, 2016. **89**(1): p. 11-36.
30. Apkarian, A.V., M.N. Baliki, and P.Y. Geha, *Towards a theory of chronic pain.* Prog Neurobiol, 2009. **87**(2): p. 81-97.
31. Murray, C.J. and A.D. Lopez, *Measuring the global burden of disease.* N Engl J Med, 2013. **369**(5): p. 448-57.
32. Sullivan, M.D. and C.Q. Howe, *Opioid therapy for chronic pain in the United States: promises and perils.* Pain, 2013. **154 Suppl 1**: p. S94-100.
33. van Tulder, M.W., et al., *Nonsteroidal anti-inflammatory drugs for low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group.* Spine (Phila Pa 1976), 2000. **25**(19): p. 2501-13.
34. Breivik, H., et al., *Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment.* Eur J Pain, 2006. **10**(4): p. 287-333.
35. Johannes CB, L.T., Zhou X, Johnston JA, Dworking RH, *The prevalence of chornic pain in the United States adults: results of an Internet-based survey.* Journal of Pain, 2010. **11**: p. 1230-1239.
36. Dworkin, R.H., J. Katz, and M.J. Gitlin, *Placebo response in clinical trials of depression and its implications for research on chronic neuropathic pain.* Neurology, 2005. **65**(12 Suppl 4): p. S7-19.
37. Andersson, G.B., Frymoyer, J.W., *The Epidemiology of Spinal Disorders*. 1997, Philadelphia, PA: Lippincott-Raven Publishers.

38. Frymoyer, J.W., *Back pain and sciatica.* N Engl J Med, 1988. **318**(5): p. 291-300.
39. Hart, L.G., Deyo, R.A., Cherkin, D.C>, *Physician office visits for low back pain. Frequency, clinical evaluation, and treatment patterns from a U.S. national survey.* Spine (Phila Pa 1976), 1995. **20**(1): p. 11-9.
40. Guo, H.R., et al., *Back pain prevalence in US industry and estimates of lost workdays.* Am J Public Health, 1999. **89**(7): p. 1029-35.
41. Carey, T.S., et al., *The outcomes and costs of care for acute low back pain among patients seen by primary care practitioners, chiropractors, and orthopedic surgeons. The North Carolina Back Pain Project.* N Engl J Med, 1995. **333**(14): p. 913-7.
42. Baliki, M.N., et al., *Corticostriatal functional connectivity predicts transition to chronic back pain.* Nat Neurosci, 2012. **15**(8): p. 1117-9.
43. Freburger, J.K., et al., *The rising prevalence of chronic low back pain.* Arch Intern Med, 2009. **169**(3): p. 251-8.
44. Ehrlich, G.E., *Low back pain.* Bull World Health Organ, 2003. **81**(9): p. 671-6.
45. Vachon-Presseau, E., et al., *Corticolimbic anatomical characteristics predetermine risk for chronic pain.* Brain, 2016. **139**(Pt 7): p. 1958-70.
46. Mutso, A.A., et al., *Reorganization of hippocampal functional connectivity with transition to chronic back pain.* J Neurophysiol, 2014. **111**(5): p. 1065-76.
47. Baliki, M.N., et al., *Functional reorganization of the default mode network across chronic pain conditions.* PLoS One, 2014. **9**(9): p. e106133.
48. Hashmi, J.A., et al., *Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits.* Brain, 2013. **136**(Pt 9): p. 2751-68.
49. Apkarian, A.V., et al., *Chronic back pain is associated with decreased prefrontal and thalamic gray matter density.* J Neurosci, 2004. **24**(46): p. 10410-5.
50. Mutso, A.A., et al., *Abnormalities in hippocampal functioning with persistent pain.* J Neurosci, 2012. **32**(17): p. 5747-56.
51. Vachon-Presseau, E., et al., *The stress model of chronic pain: evidence from basal cortisol and hippocampal structure and function in humans.* Brain, 2013. **136**(Pt 3): p. 815-27.
52. Mansour, A.R., et al., *Brain white matter structural properties predict transition to chronic pain.* Pain, 2013. **154**(10): p. 2160-8.
53. Farmer, M.A., M.N. Baliki, and A.V. Apkarian, *A dynamic network perspective of chronic pain.* Neurosci Lett, 2012. **520**(2): p. 197-203.
54. Mansour, A.R., et al., *Chronic pain: the role of learning and brain plasticity.* Restor Neurol Neurosci, 2014. **32**(1): p. 129-39.
55. Vachon-Presseau, E., et al., *The Emotional Brain as a Predictor and Amplifier of Chronic Pain.* J Dent Res, 2016. **95**(6): p. 605-12.
56. Geha, P.Y., et al., *The brain in chronic CRPS pain: abnormal gray-white matter interactions in emotional and autonomic regions.* Neuron, 2008. **60**(4): p. 570-81.
57. Parks, E.L., et al., *Brain activity for chronic knee osteoarthritis: dissociating evoked pain from spontaneous pain.* Eur J Pain, 2011. **15**(8): p. 843 e1-14.
58. Rodriguez-Raecke, R., et al., *Brain gray matter decrease in chronic pain is the consequence and not the cause of pain.* J Neurosci, 2009. **29**(44): p. 13746-50.

59. Gracely, R.H., et al., *Pain catastrophizing and neural responses to pain among persons with fibromyalgia.* Brain, 2004. **127**(Pt 4): p. 835-43.
60. Luerding, R., et al., *Working memory performance is correlated with local brain morphology in the medial frontal and anterior cingulate cortex in fibromyalgia patients: structural correlates of pain-cognition interaction.* Brain, 2008. **131**(Pt 12): p. 3222-31.
61. Schweinhardt, P., et al., *Increased gray matter density in young women with chronic vulvar pain.* Pain, 2008. **140**(3): p. 411-9.
62. Farmer, M.A., et al., *Brain functional and anatomical changes in chronic prostatitis/chronic pelvic pain syndrome.* J Urol, 2011. **186**(1): p. 117-24.
63. Farmer, M.A., et al., *Brain White Matter Abnormalities in Female Interstitial Cystitis/Bladder Pain Syndrome: A MAPP Network Neuroimaging Study.* J Urol, 2015. **194**(1): p. 118-26.
64. Mayer, E.A., et al., *Differences in brain responses to visceral pain between patients with irritable bowel syndrome and ulcerative colitis.* Pain, 2005. **115**(3): p. 398-409.
65. Kim, H.J. and S.Y. Seo, *Recurrent emotion-triggered headache following primary headache associated with sexual activity.* J Neurol Sci, 2008. **273**(1-2): p. 142-3.
66. Baliki, M.N., A.T. Baria, and A.V. Apkarian, *The cortical rhythms of chronic back pain.* J Neurosci, 2011. **31**(39): p. 13981-90.
67. Baliki, M.N., et al., *Brain morphological signatures for chronic pain.* PLoS One, 2011. **6**(10): p. e26010.
68. Gwilym, S.E., et al., *Thalamic atrophy associated with painful osteoarthritis of the hip is reversible after arthroplasty: a longitudinal voxel-based morphometric study.* Arthritis Rheum, 2010. **62**(10): p. 2930-40.
69. Obermann, M., et al., *Gray matter changes related to chronic posttraumatic headache.* Neurology, 2009. **73**(12): p. 978-83.
70. Gore, M., et al., *The burden of chronic low back pain: clinical comorbidities, treatment patterns, and health care costs in usual care settings.* Spine (Phila Pa 1976), 2012. **37**(11): p. E668-77.
71. Shmagel, A., R. Foley, and H. Ibrahim, *Epidemiology of Chronic Low Back Pain in US Adults: Data From the 2009-2010 National Health and Nutrition Examination Survey.* Arthritis Care Res (Hoboken), 2016. **68**(11): p. 1688-1694.
72. Stubbs, B., et al., *The epidemiology of back pain and its relationship with depression, psychosis, anxiety, sleep disturbances, and stress sensitivity: Data from 43 low- and middle-income countries.* Gen Hosp Psychiatry, 2016. **43**: p. 63-70.
73. Borsook, D., et al., *Reward deficiency and anti-reward in pain chronification.* Neurosci Biobehav Rev, 2016. **68**: p. 282-97.
74. Dick, B.D. and S. Rashiq, *Disruption of attention and working memory traces in individuals with chronic pain.* Anesth Analg, 2007. **104**(5): p. 1223-9, tables of contents.
75. Veldhuijzen, D.S., et al., *Pain and attention: attentional disruption or distraction?* J Pain, 2006. **7**(1): p. 11-20.

76. Veldhuijzen, D.S., et al., *Effect of chronic nonmalignant pain on highway driving performance.* Pain, 2006. **122**(1-2): p. 28-35.
77. Ling, J., et al., *Short-term prospective memory deficits in chronic back pain patients.* Psychosom Med, 2007. **69**(2): p. 144-8.
78. Baliki, M.N., et al., *Predicting value of pain and analgesia: nucleus accumbens response to noxious stimuli changes in the presence of chronic pain.* Neuron, 2010. **66**(1): p. 149-60.
79. Baliki, M.N., et al., *Parceling human accumbens into putative core and shell dissociates encoding of values for reward and pain.* J Neurosci, 2013. **33**(41): p. 16383-93.
80. Verdejo-Garcia, A., et al., *Executive function and decision-making in women with fibromyalgia.* Arch Clin Neuropsychol, 2009. **24**(1): p. 113-22.
81. Walteros, C., et al., *Altered associative learning and emotional decision making in fibromyalgia.* J Psychosom Res, 2011. **70**(3): p. 294-301.
82. Apkarian, A.V., et al., *Chronic pain patients are impaired on an emotional decision-making task.* Pain, 2004. **108**(1-2): p. 129-36.
83. Berger, S.E., et al., *Risky monetary behavior in chronic back pain is associated with altered modular connectivity of the nucleus accumbens.* BMC Res Notes, 2014. **7**: p. 739.
84. Neumann, S., et al., *Inflammatory pain hypersensitivity mediated by phenotypic switch in myelinated primary sensory neurons.* Nature, 1996. **384**(6607): p. 360-4.
85. Woolf, C.J. and M.W. Salter, *Neuronal plasticity: increasing the gain in pain.* Science, 2000. **288**(5472): p. 1765-9.
86. Baliki, M.N., et al., *Resting-sate functional reorganization of the rat limbic system following neuropathic injury.* Sci Rep, 2014. **4**: p. 6186.
87. Han, J.S., W. Li, and V. Neugebauer, *Critical role of calcitonin gene-related peptide 1 receptors in the amygdala in synaptic plasticity and pain behavior.* J Neurosci, 2005. **25**(46): p. 10717-28.
88. Chang, P.C., et al., *Role of nucleus accumbens in neuropathic pain: linked multi-scale evidence in the rat transitioning to neuropathic pain.* Pain, 2014. **155**(6): p. 1128-39.
89. Ren, W., et al., *The indirect pathway of the nucleus accumbens shell amplifies neuropathic pain.* Nat Neurosci, 2016. **19**(2): p. 220-2.
90. Li, X.Y., et al., *Alleviating neuropathic pain hypersensitivity by inhibiting PKMzeta in the anterior cingulate cortex.* Science, 2010. **330**(6009): p. 1400-4.
91. Apkarian, A.V., et al., *Role of adult hippocampal neurogenesis in persistent pain.* Pain, 2016. **157**(2): p. 418-28.
92. Ji, G., et al., *Cognitive impairment in pain through amygdala-driven prefrontal cortical deactivation.* J Neurosci, 2010. **30**(15): p. 5451-64.
93. Pais-Vieira, M., et al., *Inflammatory pain disrupts the orbitofrontal neuronal activity and risk-assessment performance in a rodent decision-making task.* Pain, 2012. **153**(8): p. 1625-35.
94. Pais-Vieira, M., D. Lima, and V. Galhardo, *Sustained attention deficits in rats with chronic inflammatory pain.* Neurosci Lett, 2009. **463**(1): p. 98-102.

95. Pais-Vieira, M., et al., *Cognitive impairment of prefrontal-dependent decision-making in rats after the onset of chronic pain.* Neuroscience, 2009. **161**(3): p. 671-9.
96. Ren, W., et al., *Peripheral Nerve Injury Leads to Working Memory Deficits and Dysfunciton of the Hippocampus by Upregulation of TNF-a in Rodents.* Neuropsychopharmacology, 2011. **26**(979-992).
97. Keller, P.H., et al., *How Health Professionals Conceptualize and Represent Placebo Treatment in Clinical Trials and How Their Patients Understand It: Impact on Validity of Informed Consent.* PLoS One, 2016. **11**(5): p. e0155940.
98. de Craen, A.J., et al., *Placebos and placebo effects in medicine: historical overview.* J R Soc Med, 1999. **92**(10): p. 511-5.
99. Wager, T.D. and L.Y. Atlas, *The neuroscience of placebo effects: connecting context, learning and health.* Nat Rev Neurosci, 2015. **16**(7): p. 403-18.
100. Ochoa, J.L., *Chronic pains associated with positive and negative sensory, motor, and vaso-motor manifestations: CPSMV (RSD; CRPS?). Heterogeneous somatic versus psychopathologic origins.* Contemporary Neurology, 1997. **2**: p. 1-20.
101. Annoni, M., *Highlights from the 2013 Science of Placebo thematic workshop.* Ecancermedicalscience, 2013. **7**: p. 346.
102. Benedetti, F., *The importance of considering the effects of perceived group assignment in placebo-controlled trials.* Eval Health Prof, 2005. **28**(1): p. 5-6.
103. Benedetti, F., *Mechanisms of placebo and placebo-related effects across diseases and treatments.* Annu Rev Pharmacol Toxicol, 2008. **48**: p. 33-60.
104. Hyland, M.E., *Using the placebo response in clinical practice.* Clinical Medicine, 2003. **3**: p. 347-350.
105. Tracey, I., *Getting the pain you expect: mechanisms of placebo, nocebo and reappraisal effects in humans.* Nat Med, 2010. **16**(11): p. 1277-83.
106. Finniss, D.G., et al., *Biological, clinical, and ethical advances of placebo effects.* Lancet, 2010. **375**(9715): p. 686-95.
107. Haake, M., et al., *German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups.* Arch Intern Med, 2007. **167**(17): p. 1892-8.
108. Vase, L., et al., *Factors contributing to large analgesic effects in placebo mechanism studies conducted between 2002 and 2007.* Pain, 2009. **145**(1-2): p. 36-44.
109. Petrovic, P., et al., *Placebo and opioid analgesia-- imaging a shared neuronal network.* Science, 2002. **295**(5560): p. 1737-1740.
110. Eippert, F., et al., *Activation of the opioidergic descending pain control system underlies placebo analgesia.* Neuron, 2009. **63**(4): p. 533-43.
111. Benedetti, F., C. Arduino, and M. Amanzio, *Somatotopic activation of opioid systems by target-directed expectations of analgesia.* J Neurosci, 1999. **19**(9): p. 3639-48.
112. Eippert, F., et al., *Direct evidence for spinal cord involvement in placebo analgesia.* Science, 2009. **326**(5951): p. 404.
113. Scott, D.J., et al., *Individual differences in reward responding explain placebo-induced expectations and effects.* Neuron, 2007. **55**(2): p. 325-36.

114. Kong, J., et al., *Functional connectivity of the frontoparietal network predicts cognitive modulation of pain.* Pain, 2013. **154**(3): p. 459-67.
115. Colloca, L. and C. Grillon, *Understanding placebo and nocebo responses for pain management.* Curr Pain Headache Rep, 2014. **18**(6): p. 419.
116. Holmes, R.D., A.K. Tiwari, and J.L. Kennedy, *Mechanisms of the placebo effect in pain and psychiatric disorders.* Pharmacogenomics J, 2016.
117. Hyland, M.E., B. Whalley, and A.W. Geraghty, *Dispositional predictors of placebo responding: a motivational interpretation of flower essence and gratitude therapy.* J Psychosom Res, 2007. **62**(3): p. 331-40.
118. Geers, A.L., et al., *Reconsidering the role of personality in placebo effects: dispositional optimism, situational expectations, and the placebo response.* J Psychosom Res, 2005. **58**(2): p. 121-7.
119. Lou, J.S., et al., *Effect of expectancy and personality on cortical excitability in Parkinson's disease.* Mov Disord, 2013. **28**(9): p. 1257-62.
120. Morton, D.L., et al., *Reproducibility of placebo analgesia: Effect of dispositional optimism.* Pain, 2009. **146**(1-2): p. 194-8.
121. Geers, A.L., I.M. Handley, and A.R. McLarney, *Discerning the role of optimism in persuasion: the valence-enhancement hypothesis.* J Pers Soc Psychol, 2003. **85**(3): p. 554-65.
122. Pecina, M., et al., *Personality trait predictors of placebo analgesia and neurobiological correlates.* Neuropsychopharmacology, 2013. **38**(4): p. 639-46.
123. Colloca, L. and F. Benedetti, *Placebo analgesia induced by social observational learning.* Pain, 2009. **144**(1-2): p. 28-34.
124. Schweinhardt, P., et al., *The anatomy of the mesolimbic reward system: a link between personality and the placebo analgesic response.* J Neurosci, 2009. **29**(15): p. 4882-7.
125. Morton, D.L., et al., *Placebo analgesia as a case of a cognitive style driven by prior expectation.* Brain Res, 2010. **1359**: p. 137-41.
126. Shapiro, A.K., & Shapiro, Elaine, *The powerful placebo : from ancient priest to modern physician*. 1997, Baltimore: Johns Hopkins University Press.
127. Vase, L., J.L. Riley, 3rd, and D.D. Price, *A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia.* Pain, 2002. **99**(3): p. 443-52.
128. Vase, L., et al., *The contributions of suggestion, desire, and expectation to placebo effects in irritable bowel syndrome patients. An empirical investigation.* Pain, 2003. **105**(1-2): p. 17-25.
129. Zubieta, J.K., et al., *Belief or Need? Accounting for individual variations in the neurochemistry of the placebo effect.* Brain Behav Immun, 2006. **20**(1): p. 15-26.
130. Kaptchuk, T.J., et al., *Do "placebo responders" exist?* Contemp Clin Trials, 2008. **29**(4): p. 587-95.
131. Horing, B., et al., *Prediction of placebo responses: a systematic review of the literature.* Front Psychol, 2014. **5**: p. 1079.
132. Jaksic, N., B. Aukst-Margetic, and M. Jakovljevic, *Does personality play a relevant role in the placebo effect?* Psychiatr Danub, 2013. **25**(1): p. 17-23.

133. Enck, P., F. Benedetti, and M. Schedlowski, *New insights into the placebo and nocebo responses.* Neuron, 2008. **59**(2): p. 195-206.
134. Enck, P., et al., *The placebo response in medicine: minimize, maximize or personalize?* Nat Rev Drug Discov, 2013. **12**(3): p. 191-204.
135. Enck, P. and S. Klosterhalfen, *The placebo response in clinical trials-the current state of play.* Complement Ther Med, 2013. **21**(2): p. 98-101.
136. Hyland, M.E., G.T. Lewith, and C. Westoby, *Developing a measure of attitudes: the holistic complementary and alternative medicine questionnaire.* Complement Ther Med, 2003. **11**(1): p. 33-8.
137. Dolgin, E., *Fluctuating baseline pain impliated in failure of clinical trials.* Nature Medicine, 2010. **16**(1053).
138. Kong, J., et al., *Are all placebo effects equal? Placebo pills, sham acupuncture, cue conditioning and their association.* PLoS One, 2013. **8**(7): p. e67485.
139. Benedetti, F., *Placebo effects: from the neurobiological paradigm to translational implications.* Neuron, 2014. **84**(3): p. 623-37.
140. Benedetti, F., *Drugs and placebos: what's the difference?: Understanding the molecular basis of the placebo effect could help clinicians to better use it in clinical practice.* EMBO Rep, 2014. **15**(4): p. 329-32.
141. Stone, D.A., et al., *Patient expectations in placebo-controlled randomized clinical trials.* J Eval Clin Pract, 2005. **11**(1): p. 77-84.
142. Schafer, S.M., L. Colloca, and T.D. Wager, *Conditioned placebo analgesia persists when subjects know they are receiving a placebo.* J Pain, 2015. **16**(5): p. 412-20.
143. Stein, N., et al., *White matter integrity of the descending pain modulatory system is associated with interindividual differences in placebo analgesia.* Pain, 2012. **153**(11): p. 2210-7.
144. Benedetti, F., et al., *Neurobiological mechanisms of the placebo effect.* J Neurosci, 2005. **25**(45): p. 10390-402.
145. Kaptchuk, T.J., et al., *Sham device v inert pill: randomised controlled trial of two placebo treatments.* BMJ, 2006. **332**(7538): p. 391-7.
146. Hashmi, J.A., et al., *Lidocaine patch (5%) is no more potent than placebo in treating chronic back pain when tested in a randomised double blind placebo controlled brain imaging study.* Mol Pain, 2012. **8**: p. 29.
147. Hashmi, J.A., et al., *Brain networks predicting placebo analgesia in a clinical trial for chronic back pain.* Pain, 2012. **153**(12): p. 2393-402.
148. Hashmi, J.A., et al., *Functional network architecture predicts psychologically mediated analgesia related to treatment in chronic knee pain patients.* J Neurosci, 2014. **34**(11): p. 3924-36.
149. Tétreault, P., et al., *Brain connectivity predicts placebo response across chronic pain clinical trials.* PLoS Biol, 2016. **Submitted**.
150. Colloca, L., et al., *Placebo analgesia: psychological and neurobiological mechanisms.* Pain, 2013. **154**(4): p. 511-4.
151. Benedetti, F. and M. Amanzio, *The placebo response: how words and rituals change the patient's brain.* Patient Educ Couns, 2011. **84**(3): p. 413-9.

152. Whalley, B., M.E. Hyland, and I. Kirsch, *Consistency of the placebo effect.* J Psychosom Res, 2008. **64**(5): p. 537-41.
153. Vachon-Presseau, E., et al., *Corticolimbic anatomical characteristics predetermine risk for chronic pain.* Brain, 2016. **In Press**.
154. Dieppe, P., *Trial designs and exploration of the placebo response.* Complement Ther Med, 2013. **21**(2): p. 105-8.
155. Kaptchuk, T.J. and F.G. Miller, *Placebo Effects in Medicine.* N Engl J Med, 2015. **373**(1): p. 8-9.
156. Tuttle, A.H., et al., *Increasing placebo responses over time in U.S. clinical trials of neuropathic pain.* Pain, 2015. **156**(12): p. 2616-26.
157. Rutherford, B.R. and S.P. Roose, *A model of placebo response in antidepressant clinical trials.* Am J Psychiatry, 2013. **170**(7): p. 723-33.
158. Sigman, M. and G.A. Cecchi, *Global organization of the Wordnet lexicon.* Proc Natl Acad Sci U S A, 2002. **99**(3): p. 1742-7.
159. Ireland, M.E., et al., *Language style matching predicts relationship initiation and stability.* Psychol Sci, 2011. **22**(1): p. 39-44.
160. Pennebaker, J.W. and L.D. Stone, *Words of wisdom: language use over the life span.* J Pers Soc Psychol, 2003. **85**(2): p. 291-301.
161. Pennebaker, J.W., M.R. Mehl, and K.G. Niederhoffer, *Psychological aspects of natural language. use: our words, our selves.* Annu Rev Psychol, 2003. **54**: p. 547-77.
162. Pennebaker, J.W. and L.A. King, *Linguistic styles: language use as an individual difference.* J Pers Soc Psychol, 1999. **77**(6): p. 1296-312.
163. Toivonen, R., et al., *Networks of emotion concepts.* PLoS One, 2012. **7**(1): p. e28883.
164. Schwartz, H.A., et al., *Personality, gender, and age in the language of social media: the open-vocabulary approach.* PLoS One, 2013. **8**(9): p. e73791.
165. Mota, N.B., et al., *Graph analysis of dream reports is especially informative about psychosis.* Sci Rep, 2014. **4**: p. 3691.
166. Bedi, G., et al., *A window into the intoxicated mind? Speech as an index of psychoactive drug effects.* Neuropsychopharmacology, 2014. **39**(10): p. 2340-8.
167. Hoffman, L.M., et al., *Concurrent and construct validity of oral language measures with school-age children with specific language impairment.* J Speech Lang Hear Res, 2011. **54**(6): p. 1597-608.
168. Kleinman, A., *Illness Narratives: Suffering, Healing, and the Human Condition*. 1988, USA: Basic Books, Inc. 284.
169. Heshusius, L., *Inside Chronic Pain: An Intimate and Critical Account*. 2009, New York: Cornell University Press.
170. Frank, A.W., *The Wounded Storyteller: Body, Illness, and Ethics*. 1995, Chicago: University of Chicago Press. 253.
171. Osborn, M.S., J.A. , *The personal experience of chronic benign lower back pain: An interpretative phenomenological analysis.* British Journal of Health Psychology, 1998. **3**(1).
172. Crowley-Matoka, M., et al., *Problems of quality and equity in pain management: exploring the role of biomedical culture.* Pain Med, 2009. **10**(7): p. 1312-24.

173. Crowley-Matoka, M., True, Gala, *No One Wants To Be the Candy Man: Ambivalent Medicalization and Clinician Subjectivity in Pain Management.* Cultural Anthropology, 2012. **27**(4): p. 689-712.
174. Halliday, M.A.K., *On the grammar of pain*, in *Functions of Language*. 1998, John Benjamins Publishing Co. p. 1-32.
175. Lascaratou, C., *The Language of Pain*. 2007: John Benjamins Publishing Company. 238.
176. Lin F., H.D., Jin Y., Tao Y., Jiang Z., *Mapping the central structure core in social representation of pain.* Journal of Applied Social Psychology, 2013. **43**: p. 1931-1945.
177. Immordino-Yang, M.H. and V. Singh, *Hippocampal contributions to the processing of social emotions.* Hum Brain Mapp, 2013. **34**(4): p. 945-55.
178. Saxbe, D.E., et al., *The embodiment of emotion: language use during the feeling of social emotions predicts cortical somatosensory activity.* Soc Cogn Affect Neurosci, 2013. **8**(7): p. 806-12.
179. de Boer, F., *Evaluating the Comparability of Two Measures of Lexical Diversity*. 2014.
180. Liegeois, F.J., et al., *Pediatric traumatic brain injury: language outcomes and their relationship to the arcuate fasciculus.* Brain Lang, 2013. **127**(3): p. 388-98.
181. Campbell, R.S. and J.W. Pennebaker, *The secret life of pronouns: flexibility in writing style and physical health.* Psychol Sci, 2003. **14**(1): p. 60-5.
182. Pennebaker, J.W., *Your use of pronouns reveals your personality.* Harv Bus Rev, 2011. **89**(12): p. 32-3.
183. Landauer, T.K., Dumais, S.T., *A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction, and Representation of Knowledge.* Psychological Review, 1997. **104**(2): p. 211-240.
184. Landauer, T.K., Foltz, P.W., Laham, D., *An introduction to latent semantic analysis.* Discource Processes, 1998. **25**(2-3): p. 259-284.
185. Evans, D.A. *Mapping vocabularies using "latent semantics"*. in *Carnegie Mellon University Research Showcase @ CMU*. 1991.
186. Bhagwant, *Latent Semantic Analysis - LSA - Tutorial*, in *Wordpress*. 2011.
187. Kontostathis A, P.W., *A framework for understanding latent semantic indexing (LSI) performance.* Information Processing and Management: an International Journal, 2006. **42**(1): p. 56-73.
188. Petersen, M.K., *Latent semantics of action verbs reflect phonetic parameters of intensity and emotional content.* PLoS One, 2015. **10**(4): p. e0121575.
189. Sole RV, C.-M.B., Valverde S, Steels L, *Language Networks: Their Structure, Function, and Evolution.* Wiley Periodicals, Inc, 2010. **15**(6).
190. Salovey, P. and A.F. Smith, *Memory for the Experience of Physical Pain*, in *Memory for Everyday and Emotional Events*, N.L. Stein, et al., Editors. 1997, Lawrence Erlbaum Associates: Mahwah, New Jersey. p. 295-314.
191. Eich, E., et al., *Memory for pain: relation between past and present pain intensity.* Pain, 1985. **23**(4): p. 375-80.
192. Norvell, K.T., F. Gaston-Johansson, and G. Fridh, *Remembrance of labor pain: how valid are retrospective pain measurements?* Pain, 1987. **31**(1): p. 77-86.

193. Lowe, N.K. and J.E. Roberts, *The convergence between in-labor report and postpartum recall of parturition pain.* Res Nurs Health, 1988. **11**(1): p. 11-21.
194. Algom, D. and S. Lubel, *Psychophysics in the field: perception and memory for labor pain.* Percept Psychophys, 1994. **55**(2): p. 133-41.
195. Babel, P., *Memory of pain induced by physical exercise.* Memory, 2015: p. 1-12.
196. Smith, W.B. and M.A. Safer, *Effects of present pain level on recall of chronic pain and medication use.* Pain, 1993. **55**(3): p. 355-61.
197. Kent, G., *Memory of dental pain.* Pain, 1985. **21**(2): p. 187-94.
198. Rocha, E.M., T.A. Marche, and C.L. von Baeyer, *Anxiety influences children's memory for procedural pain.* Pain Res Manag, 2009. **14**(3): p. 233-7.
199. Feine, J.S., et al., *Memories of chronic pain and perceptions of relief.* Pain, 1998. **77**(2): p. 137-41.
200. Linton, S.J. and L. Melin, *The accuracy of remembering chronic pain.* Pain, 1982. **13**(3): p. 281-5.
201. Redelmeier, D.A. and D. Kahneman, *Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures.* Pain, 1996. **66**(1): p. 3-8.
202. Chajut, E., et al., *In pain thou shalt bring forth children: the peak-and-end rule in recall of labor pain.* Psychol Sci, 2014. **25**(12): p. 2266-71.
203. Kahneman, D., *Thinking, fast and slow*. 1st ed. 2011, New York: Farrar, Straus and Giroux. 499 p.
204. Redelmeier, D.A., J. Katz, and D. Kahneman, *Memories of colonoscopy: a randomized trial.* Pain, 2003. **104**(1-2): p. 187-94.
205. Liu, X., et al., *Overgeneral autobiographical memory in patients with chronic pain.* Pain Med, 2014. **15**(3): p. 432-9.
206. Bryant, R.A., *Memory for pain and affect in chronic pain patients.* Pain, 1993. **54**(3): p. 347-51.
207. Jamison, R.N., T. Sbrocco, and W.C. Parris, *The influence of physical and psychosocial factors on accuracy of memory for pain in chronic pain patients.* Pain, 1989. **37**(3): p. 289-94.
208. Lefebvre, J.C. and F.J. Keefe, *The effect of neuroticism on the recall of persistent low-back pain and perceived activity interference.* J Pain, 2013. **14**(9): p. 948-56.
209. Sohl, S.J. and F. Friedberg, *Memory for fatigue in chronic fatigue syndrome: relationships to fatigue variability, catastrophizing, and negative affect.* Behav Med, 2008. **34**(1): p. 29-38.
210. Pashko, S., *Shifting between our two self-identities can cause the placebo effect and response shift.* The Journal of Transpersonal Psychology, 2013. **45**(1).
211. Gendreau, M., M.R. Hufford, and A.A. Stone, *Measuring clinical pain in chronic widespread pain: selected methodological issues.* Best Pract Res Clin Rheumatol, 2003. **17**(4): p. 575-92.
212. Price, D.D., et al., *An analysis of factors that contribute to the magnitude of placebo analgesia in an experimental paradigm.* Pain, 1999. **83**(2): p. 147-56.
213. Austin, P.C. and E.W. Steyerberg, *Predictive accuracy of risk factors and markers: a simulation study of the effect of novel markers on different*

performance measures for logistic regression models. Stat Med, 2013. **32**(4): p. 661-72.
214. Cramer, J., *The Origins of Logistic Regression*, in *Logit Models from Economics and Other Fields*. 2002, Cambridge University Press.
215. Christian, B.G., Tom, *Algorithms to Live By*. 2016, New York, NY, USA: Henry Holt and Company, LLC. 351.
216. Tong, S.K., Daphne, *Support Vector Machine Active Learning with Applicaitons to Text Classification.* Journal of Machine Learning Research, 2001: p. 45-66.
217. Steyerberg, E.W., Yvonne Vergouwe, *Towards better clinical prediction models: seven steps for development and an ABCD for validation.* Eur Heart J, 2014. **35**(29): p. 1925-1931.
218. Kohavi, R., *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection.* International Joint Conference on Artificial Intelligence, 1995.
219. Dubitzky, W., Granzow, M., Berrar, D.P., *Fundamentals of Data Mining in Genomics and Proteomics*. 2007: Springer Science + Business Media, LLC.
220. Petre, B., et al., *Smoking increases risk of pain chronification through shared corticostriatal circuitry.* Hum Brain Mapp, 2015. **36**(2): p. 683-94.
221. Harden, R.N., et al., *Medication Quantification Scale Version III: update in medication classes and revised detriment weights by survey of American Pain Society Physicians.* J Pain, 2005. **6**(6): p. 364-71.
222. (NCCIH), N.C.f.C.a.I.H. *NCCIH Clinical Research Toolbox*. [cited 2016 May 6th]; Available from: https://nccih.nih.gov/grants/toolbox.
223. Melzack, R., *The short-form McGill Pain Questionnaire.* Pain, 1987. **30**(2): p. 191-7.
224. Galer, B.S. and M.P. Jensen, *Development and preliminary validation of a pain measure specific to neuropathic pain: the Neuropathic Pain Scale.* Neurology, 1997. **48**(2): p. 332-8.
225. Freynhagen, R., et al., *painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain.* Curr Med Res Opin, 2006. **22**(10): p. 1911-20.
226. McCracken, L.M., K.E. Vowles, and C. Eccleston, *Acceptance of chronic pain: component analysis and a revised assessment method.* Pain, 2004. **107**(1-2): p. 159-66.
227. McCracken, L.M., et al., *Acceptance and change in the context of chronic pain.* Pain, 2004. **109**(1-2): p. 4-7.
228. Jensen, M.P., et al., *The Chronic Pain Coping Inventory: development and preliminary validation.* Pain, 1995. **60**(2): p. 203-16.
229. Romano, J.M., M.P. Jensen, and J.A. Turner, *The Chronic Pain Coping Inventory-42: reliability and validity.* Pain, 2003. **104**(1-2): p. 65-73.
230. Sullivan, M.J.L.B., S.R.; Pivik, J. , *The Pain Catastrophizing Scale: Development and validation.* Psychological Assessment, 1995. **7**(4): p. 524-532.
231. McCracken, L.M. and L. Dhingra, *A short version of the Pain Anxiety Symptoms Scale (PASS-20): preliminary development and validity.* Pain Res Manag, 2002. **7**(1): p. 45-50.

232. Ruscheweyh, R., et al., *Validation of the pain sensitivity questionnaire in chronic pain patients.* Pain, 2012. **153**(6): p. 1210-8.
233. Younger, J., et al., *Development of the Stanford Expectations of Treatment Scale (SETS): a tool for measuring patient outcome expectancy in clinical trials.* Clin Trials, 2012. **9**(6): p. 767-76.
234. Horne, R., et al., *The perceived sensitivity to medicines (PSM) scale: an evaluation of validity and reliability.* Br J Health Psychol, 2013. **18**(1): p. 18-30.
235. Wallston, K.A., M.J. Stein, and C.A. Smith, *Form C of the MHLC scales: a condition-specific measure of locus of control.* J Pers Assess, 1994. **63**(3): p. 534-53.
236. Mehling, W.E., et al., *Self-reported interoceptive awareness in primary care patients with past or current low back pain.* J Pain Res, 2013. **6**: p. 403-18.
237. Mehling, W.E., et al., *The Multidimensional Assessment of Interoceptive Awareness (MAIA).* PLoS One, 2012. **7**(11): p. e48230.
238. Gross, J.J. and O.P. John, *Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being.* J Pers Soc Psychol, 2003. **85**(2): p. 348-62.
239. Olafsson, R.P., et al., *Self reported attentional control with the Attentional Control Scale: factor structure and relationship with symptoms of anxiety and depression.* J Anxiety Disord, 2011. **25**(6): p. 777-82.
240. Barry, T.J.H., D.; Lenaert, B.; Debeer, E.; Griffith, J.W., *The eACS: Attentional control in the presence of emotion.* Personality and Individual Differences, 2013. **55**: p. 777-782.
241. Baer, R.A.S., G.T; Hoplins ,J; Krietemeyer, J; Toney, L. , *Using Self-Report Assessment to Explore Facets of Mindfulness.* Assessment, 2006. **13**(1): p. 27-45.
242. Kotov, R.I.B., S.B.; Watson, D.B. . *Multidimensional Iowa Suggestibility Scale (MISS) Brief Manual*. 2004; Available from: https://medicine.stonybrookmedicine/system/files/MISSBriefManual.pdf.
243. Goldberg, L.R., *The development of markers for the Big-Five factor structure.* Psychological Assessment, 1992. **4**: p. 26-42.
244. Costa, P.T., & McCrae, R.R., *Revised NEO Personality Inventory (NEO-PIR) and NEO Five Factor Inventory (NEO-FFI) professional manual*. 1992, Psychological Assessment Resources: Odessa, FL.
245. Scheier, M.F.C., Charles S,; Bridges, M. W. , *Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test.* Journal of Personality and Social Psychology, 1994. **67**(6): p. 1063-1078.
246. De Baets, S.B., M. , *Development of the loss aversion questionnaire*, V.B. School, Editor. 2012.
247. Beck, A.T., Steer, R.A., *Manual for the Beck Depression Inventory*. 1993, Psychological Corporation: San Antonio, TX.
248. Watson, D., L.A. Clark, and A. Tellegen, *Development and validation of brief measures of positive and negative affect: the PANAS scales.* J Pers Soc Psychol, 1988. **54**(6): p. 1063-70.

249. Bourke, J., *The Story of Pain: From Prayer to Painkillers*. 2014, UK: Oxford University Press.
250. Tetreault, P., et al., *Brain Connectivity Predicts Placebo Response across Chronic Pain Clinical Trials.* PLoS Biol, 2016. **14**(10): p. e1002570.
251. Nasreddine, Z.S., et al., *The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment.* J Am Geriatr Soc, 2005. **53**(4): p. 695-9.
252. Von Korff, M., M.P. Jensen, and P. Karoly, *Assessing global pain severity by self-report in clinical and health services research.* Spine (Phila Pa 1976), 2000. **25**(24): p. 3140-51.
253. Jensen, M.P., et al., *Assessment of pain quality in chronic neuropathic and nociceptive pain clinical trials with the Neuropathic Pain Scale.* J Pain, 2005. **6**(2): p. 98-106.
254. Nuzzo, R., *How scientists fool themselves - and how they can stop.* Nature, 2015. **526**(7572): p. 182-5.
255. MacCoun, R. and S. Perlmutter, *Blind analysis: Hide results to seek the truth.* Nature, 2015. **526**(7572): p. 187-9.
256. Feinberg, D.A., et al., *Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging.* PLoS One, 2010. **5**(12): p. e15710.
257. Power, J.D., et al., *Methods to detect, characterize, and remove motion artifact in resting state fMRI.* Neuroimage, 2014. **84**: p. 320-41.
258. Power, J.D., et al., *Functional network organization of the human brain.* Neuron, 2011. **72**(4): p. 665-78.
259. Rubinov, M. and O. Sporns, *Complex network measures of brain connectivity: uses and interpretations.* Neuroimage, 2010. **52**(3): p. 1059-69.
260. Rubinov, M. and O. Sporns, *Weight-conserving characterization of complex functional brain networks.* Neuroimage, 2011. **56**(4): p. 2068-79.
261. Zalesky, A., A. Fornito, and E.T. Bullmore, *Network-based statistic: identifying differences in brain networks.* Neuroimage, 2010. **53**(4): p. 1197-207.
262. Patenaude, B., et al., *A Bayesian model of shape and appearance for subcortical brain segmentation.* Neuroimage, 2011. **56**(3): p. 907-22.
263. van Erp, T.G., et al., *Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium.* Mol Psychiatry, 2016. **21**(4): p. 585.
264. Pigott, T.D., *A Review of Methods for Missing Data.* Educational Research and Evaluation, 2001. **7**(4): p. 353-383.
265. Steger, M., *An Illustration of Issues in Factor Extraction and Identification of Dimensionality in Psychological Assessment Data.* Journal of Personality Assessment, 2006. **86**(3): p. 263-272.
266. Hess, C.W., Ritchie, K.P., Landry, R.G. , *The Type-Token Ration and Vocabulary Performance.* Psychological Reports, 1984. **55**: p. 51-57.
267. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R., *Indexing by latent semantic analysis.* Journal of the American Society for Information Science, 1990. **41**: p. 391-407.

268. Huth, A.G., et al., *Natural speech reveals the semantic maps that tile human cerebral cortex.* Nature, 2016. **532**(7600): p. 453-8.
269. Berryman, C., et al., *Evidence for working memory deficits in chronic pain: a systematic review and meta-analysis.* Pain, 2013. **154**(8): p. 1181-96.
270. Dick, B., C. Eccleston, and G. Crombez, *Attentional functioning in fibromyalgia, rheumatoid arthritis, and musculoskeletal pain patients.* Arthritis Rheum, 2002. **47**(6): p. 639-44.
271. Mifflin, K., J. Chorney, and B. Dick, *Attention and Working Memory in Female Adolescents With Chronic Pain and Pain-free Female Adolescents: A Preliminary Pilot Study.* Clin J Pain, 2016. **32**(7): p. 609-16.
272. Grisart, J.M. and M. Van der Linden, *Conscious and automatic uses of memory in chronic pain patients.* Pain, 2001. **94**(3): p. 305-13.
273. Lefebvre, J.C. and F.J. Keefe, *Memory for pain: the relationship of pain catastrophizing to the recall of daily rheumatoid arthritis pain.* Clin J Pain, 2002. **18**(1): p. 56-63.
274. Munoz, M. and R. Esteve, *Reports of memory functioning by patients with chronic pain.* Clin J Pain, 2005. **21**(4): p. 287-91.
275. Noel, M., et al., *Remembering pain after surgery: a longitudinal examination of the role of pain catastrophizing in children's and parents' recall.* Pain, 2015. **156**(5): p. 800-8.
276. Pallegama, R.W., et al., *Influence of Catastrophizing and Personality Traits on Recalled Ratings of Acute Pain Experience in Healthy Young Adults.* Pain Med, 2016.
277. Simon, C.B., et al., *Comparative Associations of Working Memory and Pain Catastrophizing With Chronic Low Back Pain Intensity.* Phys Ther, 2016. **96**(7): p. 1049-56.
278. Russo, S.J., et al., *Neurobiology of resilience.* Nat Neurosci, 2012. **15**(11): p. 1475-84.
279. Drevets, W.C., J.L. Price, and M.L. Furey, *Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression.* Brain Struct Funct, 2008. **213**(1-2): p. 93-118.
280. Okada, N., et al., *Abnormal asymmetries in subcortical brain volume in schizophrenia.* Mol Psychiatry, 2016.
281. Poldrack, R.A., et al., *Long-term neural and physiological phenotyping of a single human.* Nat Commun, 2015. **6**: p. 8885.
282. Koban, L., Ruzic, L. & Wager, T.D. , *Brain Predictors of Individual DIfferences in Placebo Responding*, in *Placebo and Pain*, M.A.F. L. Colloca, & K. Meissner, Editor. 2013, Elsevier/Academic Press.
283. Roy, M., et al., *Representation of aversive prediction errors in the human periaqueductal gray.* Nat Neurosci, 2014. **17**(11): p. 1607-12.
284. Pouli, N., et al., *The experience of living with knee osteoarthritis: exploring illness and treatment beliefs through thematic analysis.* Disabil Rehabil, 2014. **36**(7): p. 600-7.
285. Clarke, S.P., et al., *Personal experience of osteoarthritis and pain questionnaires: mapping items to themes.* Disabil Rehabil, 2014. **36**(2): p. 163-9.

286. Ballantyne, J.C. and M.D. Sullivan, *Intensity of Chronic Pain--The Wrong Metric?* N Engl J Med, 2015. **373**(22): p. 2098-9.
287. Chiesi, F., et al., *The accuracy of the Life Orientation Test-Revised (LOT-R) in measuring dispositional optimism: evidence from item response theory analyses.* J Pers Assess, 2013. **95**(5): p. 523-9.
288. Nakano, K., *Psychometric properties of the Life Orientation Test-Revised in samples of Japanese students.* Psychol Rep, 2004. **94**(3 Pt 1): p. 849-55.
289. Schweinhardt, P., K.M. Sauro, and M.C. Bushnell, *Fibromyalgia: a disorder of the brain?* Neuroscientist, 2008. **14**(5): p. 415-21.
290. Osborn, M.S., J.A., *Living with a body separate from the self. The experience of the body in chronic benign low back pain: An interpretative phenomenological analysis.* Scandinavian Journal of Caring Science, 2006. **20**(2): p. 216-222.
291. van Haselen, R.A., et al., *The Constitutional Type Questionnaire: validation in the patient population of the Royal London Homoeopathic Hospital.* Br Homeopath J, 2001. **90**(3): p. 131-7.
292. Manning, J.R., et al., *Spontaneously reactivated patterns in frontal and temporal lobe predict semantic clustering during memory search.* J Neurosci, 2012. **32**(26): p. 8871-8.
293. Tang, N.K.S., P.M.; Hodges, A.; Wright, K.J.; Hanna, M.; Hester, J. , *Effects of mood on pain responses and pain tolerance: an experimental study in chronic back pain patients.* Pain, 2008. **138**(2): p. 392-401.
294. Gogtay, N., et al., *Dynamic mapping of normal human hippocampal development.* Hippocampus, 2006. **16**(8): p. 664-72.
295. Satpute, A.B., et al., *Human anterior and posterior hippocampus respond distinctly to state and trait anxiety.* Emotion, 2012. **12**(1): p. 58-68.
296. Chandler, M.J., et al., *Three-word recall in normal aging.* J Clin Exp Neuropsychol, 2004. **26**(8): p. 1128-33.
297. Saletin, J.M., et al., *Human Hippocampal Structure: A Novel Biomarker Predicting Mnemonic Vulnerability to, and Recovery from, Sleep Deprivation.* J Neurosci, 2016. **36**(8): p. 2355-63.
298. Johnson, S.L., et al., *Hippocampal shape abnormalities of patients with childhood-onset schizophrenia and their unaffected siblings.* J Am Acad Child Adolesc Psychiatry, 2013. **52**(5): p. 527-536 e2.
299. Strange, B.A., et al., *Functional organization of the hippocampal longitudinal axis.* Nat Rev Neurosci, 2014. **15**(10): p. 655-69.
300. Bannerman, D.M., et al., *Regional dissociations within the hippocampus--memory and anxiety.* Neurosci Biobehav Rev, 2004. **28**(3): p. 273-83.
301. Fanselow, M.S. and H.W. Dong, *Are the dorsal and ventral hippocampus functionally distinct structures?* Neuron, 2010. **65**(1): p. 7-19.
302. Poppenk, J., et al., *Long-axis specialization of the human hippocampus.* Trends Cogn Sci, 2013. **17**(5): p. 230-40.
303. Yarkoni, T., et al., *Large-scale automated synthesis of human functional neuroimaging data.* Nat Methods, 2011. **8**(8): p. 665-70.

304. Voineskos, A.N., et al., *Hippocampal (subfield) volume and shape in relation to cognitive performance across the adult lifespan.* Hum Brain Mapp, 2015. **36**(8): p. 3020-37.
305. Robinson, J.L., N. Salibi, and G. Deshpande, *Functional connectivity of the left and right hippocampi: Evidence for functional lateralization along the long-axis using meta-analytic approaches and ultra-high field functional neuroimaging.* Neuroimage, 2016. **135**: p. 64-78.
306. Fairhurst, M., et al., *An fMRI study exploring the overlap and differences between neural representations of physical and recalled pain.* PLoS One, 2012. **7**(10): p. e48711.
307. Yu, E.C.L., D. A.; Chater, N. *Retrospective Evaluations of Gambling Wins: Evidence for a 'Peak-End' Rule*. in *Proeceedings of the 30th Annual Conference of the Cognitive Science Society*. 2008. Austin, TX: Cognitive Science Society.
308. Kobak, K.A., et al., *Why do clinical trials fail? The problem of measurement error in clinical trials: time to test new paradigms?* J Clin Psychopharmacol, 2007. **27**(1): p. 1-5.
309. Marchant, J., *Strong placebo response thwarts painkiller trials.* Nature News, 2015.
310. Herman, J., *Therapeutic nihilism?* Isr J Med Sci, 1996. **32**(3-4): p. 259-264.
311. Gawande, A., *Overkill*, in *The New Yorker*. 2015.
312. Miller, F.G. and T.J. Kaptchuk, *The power of context: reconceptualizing the placebo effect.* J R Soc Med, 2008. **101**(5): p. 222-5.
313. Tilburt, J.C., et al., *Prescribing "placebo treatments": results of national survey of US internists and rheumatologists.* BMJ, 2008. **337**: p. a1938.
314. Bostick, N.A., et al., *Placebo use in clinical practice: report of the American Medical Association Council on Ethical and Judicial Affairs.* J Clin Ethics, 2008. **19**(1): p. 58-61.
315. Amanzio, M., et al., *Response variability to analgesics: a role for non-specific activation of endogenous opioids.* Pain, 2001. **90**(3): p. 205-15.
316. Kaptchuk, T.J., et al., *Placebos without deception: a randomized controlled trial in irritable bowel syndrome.* PLoS One, 2010. **5**(12): p. e15591.
317. Carvalho, C., et al., *Open-label placebo treatment in chronic low back pain: a randomized controlled trial.* Pain, 2016.
318. Robinson, M.E., et al., *Gender role expectations of pain: relationship to sex differences in pain.* J Pain, 2001. **2**(5): p. 251-7.
319. Pavlenko, A., *Emotions and the body in Russian and English.* Pragmatics and Cognition, 2002.
320. Leleu, T.D., et al., *Application of latent semantic analysis for open-ended responses in a large, epidemiologic study.* BMC Med Res Methodol, 2011. **11**: p. 136.
321. Charmaz, K., *Grounded Theory*, in *Qualitative Psychology: A Practical Guide to Research Methods*, J.A. Smith, Editor. 2003, Sage: London, UK.
322. Scarry, E., *The Body in Pain: The Making and Unmaking of the World*. 2985, New York, USA: Oxford University Press.

323. Tasmuth, T., A.M. Estlanderb, and E. Kalso, *Effect of present pain and mood on the memory of past postoperative pain in women treated surgically for breast cancer.* Pain, 1996. **68**(2-3): p. 343-7.
324. Cushman, D., et al., *Discrepancies in describing pain: is there agreement between numeric rating scale scores and pain reduction percentage reported by patients with musculoskeletal pain after corticosteroid injection?* Pain Med, 2015. **16**(5): p. 870-6.
325. Wager, T.D., et al., *Predicting individual differences in placebo analgesia: contributions of brain activity during anticipation and pain experience.* J Neurosci, 2011. **31**(2): p. 439-52.
326. Jarcho, J.M., et al., *Placebo analgesia: Self-report measures and preliminary evidence of cortical dopamine release associated with placebo response.* Neuroimage Clin, 2016. **10**: p. 107-14.
327. Senkowski, D., M. Hofle, and A.K. Engel, *Crossmodal shaping of pain: a multisensory approach to nociception.* Trends Cogn Sci, 2014. **18**(6): p. 319-27.

APPENDIX

I. Sample of excerpts from exit interviews

i. Pain descriptions

"Physical sensation pretty much is a dull uh, numbing pain at the lower back. It begins there and it's always constant, it's always with me when I wake up in the morning I know it's there. When I go to bed at night I know it's there. If I walk too much it screams, and then to go along with it the neck pain and the shoulder pain acts in conjunction with whatever's going on with the lower back pain."

"The feeling like it was bad enough like I wouldn't want to get off the sofa. If if I got off the sofa it was only to walk 10 feet to the kitchen to get something to eat or drink or to go to bed. But I'd spend all day on the sofa, which isn't that unusual for me because I haven't been working. And the pain, like often it would be like getting stabbed in the back or sawed in the back or stabbed and sawed in the back or like somebody poking you with the screwdriver and then moving the screwdriver around in big circles to twist up the muscles."

ii. Pain's effects on mood

"Yeah I definitely get grumpy, grumpier. One thing that's been very maybe shocking is not the right word, but that I've noticed or that's become obvious from filling out your questionnaires and I think more so from filling out the phone app is how depressed I am. I was in a great mood this weekend, and I think on one of the days I put a +3 for my mood and that to me was like I I was happier than I've been in a long time and it's a +3 out of a -10 to 10. But but I've had depression before a couple years ago. I I hadn't had a real job in a long long time. I've had odd jobs and then like kind of a real job umm that I don't get paid for and so there's depression around that, but like so I I know my baseline is already pretty low but definitely with the back pain I can see how much it does bring me down."

iii. Pain's effects on interpersonal relationships

"It definitely like it makes me feel like I stick out more even though like I work with disability so you know it's actually given me a larger respect for my clients. I feel like I have a partial disability. It's just never been proven and…there are hoops to go through... and it's kind of like mental illness in a sense. It's not visible enough for people to be mindful, but it's there you know? And it's like I sometimes don't even want to discuss my pain because it's like my mom will ask me like almost every day are you in pain? I'm like mom, I'm like yes but it's, it's also like such an annoying topic. It's like not changing so it's like what am I supposed to say like yeah, I'm in pain but I will say what's improved is no one asks me, besides you guys, what's your scale of pain today? I hated that question after my car wreck and my back. It's like what is at stake? I don't know like it's not a 1, it's definitely not a 2, it feels like I don't have to go to the ER so I kinda gauge right, I can handle it. And I think a lot of people ask me, they're like why don't you take like more stronger prescriptions? I'm in a fog. Like for my job I drive all over Illinois. I need to have focus. I hate being in a fog and the pain is still there. Like that's numbing my brain, it's not numbing my pain. If it localized, if the medications just focused on my pain and I

still had some clarity, I'd take it all day any day. But yeah, I view myself differently internally but externally I feel like a lot of people think everything is fine and that's a hard thing, that's a hard thing to deal with because I'm also in charge of a lot so I'm running meetings, I'm managing like four people right now and umm six agencies across Illinois. So when you walk into a room you have to have a certain presence. You know you're, you're literally training people how to deal with disability or you're doing a talk for research, you know you're presenting a poster board, sometimes you have to stand there for two hours and standing is like torture sometimes but I need to get through this poster board session. And I think the biggest fear not fear even, it's just more of like embarrassment of like asking for help. That.. I'm the worst at doing that because I want to do it on my own. You know? I don't want to ask someone for a chair. No, I'm going to stand just like everybody else and I think that's pride more than anything, I have a lot of it, but it's, it does make me feel different, you know? And it's just communicating with people I barely know. Then I don't even want to get into how I got injured. It's such a saga."

iv. Treatment responses

"Recently the pills that you gave me, I don't know what kind of pill is that but the pain is going down at last. And I even talked to my doctor, my family doctor, if he can approve it, whatever, for me because that pill worked. The pill worked great for me. I'm satisfied, I'm satisfied because I feel like a new man, you know, but I was (inaudible 20:25) about three weeks ago or four weeks ago, 60 years old you can't expect a lot, you know, but I feel new because of the study. I took that pill and still, I don't have no pain ... Medication is not in the pharmacy yet? For example, this is gonna be the last visit, if I need a medication or something like that, can I call you?"

v. Medical history influencing expectations

"Umm I know I discussed it when I first signed up for this and in some of the like beginning interviews and stuff and maybe on some of the questionnaires. I had Guillain-Barre when I was 14. I was paralyzed uh mostly paralyzed in legs and arms and lungs and what's the other one? So like nervous system and respiratory system. And I know what it's like thinking you'll never walk again. So like I'm sure if back surgery were an option I would think about what if something goes wrong and I could be that person again. I'm afraid of that happening again. Yeah, and then also you know it's, there's just a general adage that once you have back problems you'll always have back problems and so I assume it's something I'll deal with forever no matter what, you know?"

II. 272-node parcellation and coordinates used for Study 2

node id	x	y	z	module id
1	-25	-98	-12	1
2	27	-97	-13	1
3	24	32	-18	0
4	-56	-45	-24	0
5	8	41	-24	4
6	-21	-22	-20	0
7	17	-28	-17	0
8	-37	-29	-26	0
9	65	-24	-19	8
10	52	-34	-27	0
11	55	-31	-17	8
12	34	38	-12	4
13	-7	-52	61	0
14	-14	-18	40	0
15	0	-15	47	12
16	10	-2	45	0
17	-7	-21	65	12
18	-7	-33	72	12
19	13	-33	75	12
20	-54	-23	43	12
21	29	-17	71	12
22	10	-46	73	12
23	-23	-30	72	12
24	-40	-19	54	12
25	29	-39	59	12
26	50	-20	42	12
27	-38	-27	69	12
28	20	-29	60	12
29	44	-8	57	12
30	-29	-43	61	12
31	10	-17	74	12
32	22	-42	69	12
33	-45	-32	47	12
34	-21	-31	61	12
35	-13	-17	75	12
36	42	-20	55	12
37	-38	-15	69	12
38	-16	-46	73	12
39	2	-28	60	12
40	3	-17	58	12
41	38	-17	45	12
42	-49	-11	35	12
43	36	-9	14	0
44	51	-6	32	12

45	-53	-10	24	12
46	66	-8	25	12
47	-3	2	53	12
48	54	-28	34	15
49	19	-8	64	12
50	-16	-5	71	12
51	-10	-2	42	0
52	37	1	-4	0
53	13	-1	70	12
54	7	8	51	12
55	-45	0	9	15
56	49	8	-1	15
57	-34	3	4	0
58	-51	8	-2	15
59	-5	18	34	15
60	36	10	1	15
61	32	-26	13	0
62	65	-33	20	15
63	58	-16	7	15
64	-38	-33	17	0
65	-60	-25	14	15
66	-49	-26	5	0
67	43	-23	20	12
68	-50	-34	26	15
69	-53	-22	23	15
70	-55	-9	12	15
71	56	-5	13	12
72	59	-17	29	15
73	-30	-27	12	0
74	-41	-75	26	0
75	6	67	-4	4
76	8	48	-15	4
77	-13	-40	1	0
78	-18	63	-9	0
79	-46	-61	21	4
80	43	-72	28	0
81	-44	12	-34	4
82	46	16	-30	4
83	-68	-23	-16	4
84	-58	-26	-15	4
85	27	16	-17	0
86	-44	-65	35	4
87	-39	-75	44	4
88	-7	-55	27	4
89	6	-59	35	4
90	-11	-56	16	0

91	-3	-49	13	4
92	8	-48	31	4
93	15	-63	26	0
94	-2	-37	44	4
95	11	-54	17	0
96	52	-59	36	4
97	23	33	48	4
98	-10	39	52	4
99	-16	29	53	4
100	-35	20	51	4
101	22	39	39	4
102	13	55	38	4
103	-10	55	39	4
104	-20	45	39	4
105	6	54	16	4
106	6	64	22	4
107	-7	51	-1	4
108	9	54	3	4
109	-3	44	-9	4
110	8	42	-5	4
111	-11	45	8	4
112	-2	38	36	0
113	-3	42	16	4
114	-20	64	19	4
115	-8	48	23	4
116	65	-12	-19	4
117	-56	-13	-10	4
118	-58	-30	-4	4
119	65	-31	-9	8
120	-68	-41	-5	4
121	13	30	59	4
122	12	36	20	0
123	52	-2	-16	4
124	-26	-40	-8	0
125	27	-37	-13	0
126	-34	-38	-16	0
127	28	-77	-32	0
128	52	7	-30	4
129	-53	3	-27	4
130	47	-50	29	4
131	-49	-42	1	0
132	-31	19	-19	4
133	-2	-35	31	4
134	-7	-71	42	4
135	11	-66	42	4
136	4	-48	51	0

137	-46	31	-13	4
138	-10	11	67	0
139	49	35	-12	4
140	8	-91	-7	1
141	17	-91	-14	1
142	-12	-95	-13	1
143	18	-47	-10	1
144	40	-72	14	1
145	8	-72	11	1
146	-8	-81	7	1
147	-28	-79	19	1
148	20	-66	2	1
149	-24	-91	19	1
150	27	-59	-9	1
151	-15	-72	-8	1
152	-18	-68	5	1
153	43	-78	-12	1
154	-47	-76	-10	1
155	-14	-91	31	1
156	15	-87	37	1
157	29	-77	25	1
158	20	-86	-2	1
159	15	-77	31	1
160	-16	-52	-1	1
161	42	-66	-8	1
162	24	-87	24	1
163	6	-72	24	1
164	-42	-74	0	1
165	26	-79	-16	1
166	-16	-77	34	1
167	-3	-81	21	1
168	-40	-88	-6	1
169	37	-84	13	1
170	6	-81	6	1
171	-26	-90	3	1
172	-33	-79	-13	1
173	37	-81	1	1
174	-44	2	46	0
175	48	25	27	0
176	-47	11	23	0
177	-53	-49	43	8
178	-23	11	64	0
179	58	-53	-14	0
180	24	45	-15	42
181	34	54	-13	42
182	-21	41	-20	0

183	-18	-76	-24	0
184	17	-80	-34	0
185	35	-67	-34	0
186	47	10	33	0
187	-41	6	33	0
188	-42	38	21	42
189	38	43	15	42
190	49	-42	45	8
191	-28	-58	48	0
192	44	-53	47	8
193	32	14	56	0
194	37	-65	40	4
195	-42	-55	45	8
196	40	18	40	0
197	-34	55	4	42
198	-42	45	-2	42
199	33	-53	44	0
200	43	49	-2	42
201	-42	25	30	42
202	-3	26	44	15
203	11	-39	50	0
204	55	-45	37	8
205	42	0	47	0
206	31	33	26	0
207	48	22	10	0
208	-35	20	0	0
209	36	22	3	0
210	37	32	-2	0
211	34	16	-8	0
212	-11	26	25	0
213	-1	15	44	15
214	-28	52	21	42
215	0	30	27	15
216	5	23	37	15
217	10	22	27	0
218	31	56	14	42
219	26	50	27	42
220	-39	51	17	42
221	2	-24	30	4
222	6	-24	0	0
223	-2	-13	12	0
224	-10	-18	7	0
225	12	-17	8	0
226	-5	-28	-4	0
227	-22	7	-5	0
228	-15	4	8	0

229	31	-14	2	0
230	23	10	1	0
231	29	1	4	0
232	-31	-11	0	0
233	15	5	7	0
234	9	-4	6	0
235	54	-43	22	0
236	-56	-50	10	0
237	-55	-40	14	0
238	52	-33	8	0
239	51	-29	-4	0
240	56	-46	11	0
241	53	33	1	0
242	-49	25	-1	4
243	-16	-65	-20	0
244	-32	-55	-25	0
245	22	-58	-23	0
246	1	-62	-18	0
247	33	-12	-34	0
248	-31	-10	-36	0
249	49	-3	-38	4
250	-50	-7	-39	4
251	10	-62	61	0
252	-52	-63	5	0
253	-47	-51	-21	0
254	46	-47	-17	0
255	47	-30	49	12
256	22	-65	48	1
257	46	-59	4	0
258	25	-58	60	1
259	-33	-46	47	0
260	-27	-71	37	1
261	-32	-1	54	0
262	-42	-60	-9	1
263	-17	-59	64	0
264	29	-5	54	0
265	-10	14	-2	20
266	10	14	-2	20
267	-22	-2	-22	20
268	26	-2	-22	20
269	-24	-14	-18	20
270	26	-14	-18	20
271	-28	-34	-6	20
272	30	-34	-6	20

Coordinates for resting state networks are provided (including 264 Power coordinates and an additional 8 for limbic regions as described in **Study 2**). Module IDs utilized in the paper are also provided: 0 = other;

1 = visual; 4 = DMN; 8 = parietal; 12 = sensorimotor; 15 = saliency network; 20 = limbic regions; 42 = frontal regions.