MONASH University

SWEET LITTLE LIES:

THE ETHICS OF PLACEBO USE IN CLINICAL PRACTICE

Anton Allen

MA BA(Hons) BSc

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in 2015

Centre for Human Bioethics

School of Philosophical, Historical, and International Studies

Copyright notice

Abstract

This thesis intends to establish whether it is permissible for a doctor or nurse to administer a deceptive placebo to a patient if it is believed that this treatment presents that patient's best or only chance at some therapeutic benefit. To do this I first establish that there is good reason to believe that placebo treatments can have a real therapeutic effect, and that deceptive placebos are currently being used widely in clinical practice around the world. I take a Kantian approach to answer the question of whether a deceptive placebo can be consistent with respect for patient autonomy. I argue that the deception involved in a placebo treatment is not carried out to interfere with or obstruct the function of the will, but merely to make possible a means to the patient's ends that that would otherwise be unavailable. It is therefore possible for a deceptive placebo to administered in a way that is consistent with a Kantian respect for rational agency. Moreover, this analysis highlights the features that separate morally permissible cases of placebo deception from morally impermissible cases, and from morally impermissible cases of paternalistic deception more generally. I examine different standards of informed consent and argue that, as by most workable standards a patient can give valid informed consent to a treatment without knowing its exact nature or mechanism of action, that a deceptive placebo treatment can meet these standards in at least some cases. Lastly I examine the argument that the deception involved in a placebo treatment constitutes a betrayal of the trust between practitioner and patient. To do this I investigate the connection between honesty and trustworthiness, deception and betrayal. I argue that, while outright lies are indeed detrimental to the trust relationship between practitioner and patient, the withholding of that information necessary to facilitate placebo treatment need not be. I therefore conclude that, where a placebo treatment is believed to present a patient's best or only chance at

some therapeutic benefit, it is permissible for a doctor or nurse to withhold information on that treatment's nature in order to facilitate the placebo effect.

Declaration

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Acknowledgements

My heartfelt thanks to my wife Magdalena, without whose unfailing love and support this work would have been impossible. Many thanks are also due to my supervisors Justin Oakley and Michael Selgelid for the essential advice and assistance they have provided me with over these past few years. I would also like to acknowledge the financial support Monash University has offered me through the Australian Postgraduate Award scheme.

This thesis was written on LibreOffice, an open-source office suite developed and maintained by a community of volunteers around the world. It is available from www.libreoffice.org. Software is better when it's free.

Contents

1. Introduction

Our remedies oft in ourselves do lie

– Shakespeare, *All's Well That Ends Well*

Between 1995 and 1998, orthopaedic surgeon (and team physician for the Houston Rockets) J Bruce Moseley led a team carrying out a placebo-controlled trial to investigate the efficacy of arthroscopic surgery used to treat osteoarthritis of the knee.[1] At the time, this procedure was being performed more than 650,000 times a year in the US, and it is still the most common orthopaedic surgery procedure in the country, and one of the most common worldwide.[2] The trial measured the effect on pain and joint mobility of two different kinds of arthroscopic surgery, compared to the effect of a sham surgical procedure carried out by the same surgeon. The results were devastating: arthroscopic surgery was no better than a placebo. The researchers concluded that “the billions of dollars spent on such procedures annually might be put to better use.”[3]

One of the 180 participants who signed up for Moseley's trial was Sylvester Colligan, a 76 year-old World War II veteran. Colligan's knee had been giving him pain for five years when he was diagnosed with osteoarthritis and referred to Moseley. Speaking to the *New York Times* after the conclusion of the trial, Colligan describes the result of the procedure:

> “The surgery was two years ago and the knee never has bothered me since,” he says. “It's just like my other knee now. I give a whole lot of credit to Dr. Moseley. Whenever I see him on the TV during a basketball game, I call the wife in and say, 'Hey, there's the doctor that fixed my

1 Moseley et al (2002)
2 Cullen et al (2009)
3 Moseley et al (2002) p87

knee!'"[4]

A remarkable result, you might think, for a surgical procedure no better than a placebo. But Colligan didn't receive the real surgical procedure: he was assigned to the placebo group in Moseley's trial. Colligan knows this, and he knows that the procedure he received consisted only of a three shallow incisions around his kneecap. But he also knows that he can now walk and mow his lawn without pain. Nor would it have made much difference if he had received the real surgical intervention: participants in the placebo group recorded the same level of improvement in pain and knee function as was seen in both of the surgical intervention groups. In fact, at some points during the two-year follow-up, objectively-measured knee function was significantly worse in the one of the surgical intervention groups than in the placebo group.[5]

Moseley's team did not conclude that arthroscopic surgery is totally ineffective, only that it is no more effective than a placebo. The lack of any significant difference in the improvement recorded by the three groups in this trial could be described as a failure of the surgical procedure, but it could also be called a victory for the placebo effect. Indeed, the researchers concluded that the study had "shown the great potential for a placebo effect with surgery" and advised that "health care researchers should not underestimate the placebo effect, regardless of its mechanism."[6]

Unfortunately, Moseley's study did not include a no-treatment control group, so the improvement recorded by the placebo and experimental groups in the trial cannot be compared with the natural history of the condition. However, the researchers do note that in numerous previous studies (all of which lack placebo controls) arthroscopic procedures

4 Talbot M (2000)
5 Moseley et al (2002) p85
6 Moseley et al (2002) p87

reported substantially better outcomes than no-treatment controls. Suppose for a moment that this gave us good reason to believe that both the surgical procedure and the placebo procedure carried out in this trial are likely to produce significantly better results than no treatment at all for a patient with this kind of arthritis. What would this mean for the treatment of patients with this condition in clinical practice? A minimally invasive placebo procedure, which could be carried out using local anaesthetic, is likely to be just as effective as a more invasive (and more expensive) procedure that requires general anaesthesia. It seems that, by any measure, the placebo treatment is the better option. But what if the placebo treatment only works if the patient doesn't know it's a placebo? Would a doctor be justified in deceiving her patient about the nature of the placebo procedure, if this was necessary to facilitate the best therapeutic outcome?

In this thesis I argue that in some circumstances it is permissible for a doctor or nurse to administer a deceptive placebo to a patient. To be specific, if it is believed that a placebo treatment presents that patient's best, or only, chance at some therapeutic benefit, then a doctor or nurse may deceive her patient by withholding only that information necessary to promote a therapeutic placebo effect.

What is a placebo?

A placebo treatment can take may forms. As well as a the familiar sugar pill, and sham surgical procedures of the kind described above, a placebo could be a saline injection, an inert cream, an active medication at a sub-therapeutic dose, an active medication with no effect on the condition being treated, or any number of other therapies. The key factor is

that a placebo is a therapy that produces a non-specific effect, the placebo effect, rather than having a specific action for the condition being treated. For this thesis I will adopt the definition that Shapiro & Shapiro use in their book *The Powerful Placebo*:

> A placebo is any therapy (or that component of any therapy) that is intentionally or knowingly used for its nonspecific, psychological, or psychophysiological, therapeutic effect, or that is used for a presumed specific therapeutic effect on a patient, symptom, or illness, but is without specific activity for the condition being treated.[7]

A placebo treatment, then, is defined not by its nature but by the effect it is intended to have – the placebo effect. In the next chapter I will examine the placebo effect in more detail, but for now I will define it as the non-specific effect that can result from a placebo treatment.

Given that placebo treatments can have this effect on patients, then, it may seem odd to refer to them as “inert”. But when I describe a placebo as “inert”, I don’t mean that it is totally without activity: as I will argue later, we have good reason to believe that an inert placebo treatment can have a real physical effect on a patient. Rather, “inert” in this sense means that a treatment is without a specific therapeutic activity, and without any specific side-effects, so that the only effect such a treatment would be expected to have is through a non-specific placebo effect. A sugar pill, then, is an inert placebo, in the sense that the only activity it would be expected to have is the placebo effect. By contrast, an antibiotic administered for a viral infection is a placebo, because the specific activity of an an antibiotic would not be expected to have any therapeutic effect for this particular condition, and the treatment is administered only to promote a placebo effect. But an antibiotic is not inert, because it is expected to have a specific effect on the patient – although in this case

7 Shapiro and Shapiro (1997) p41

the antibiotic effect might be more properly classified as a side-effect rather than a therapeutic effect.

The example of Sylvester Colligan's placebo knee surgery has some important differences to the ideal placebo cases that I will consider for the majority of this thesis. Firstly, as Colligan was part of a clinical trial, he was informed about, and consented to, the possibility that he would receive a placebo intervention. But in this thesis I will discuss deceptive placebo treatments: it is generally thought that placebo treatments won't work patients know they are placebos, and I argue in chapter three that even asking patients to consent to the possibility of being given an undisclosed placebo treatment is a bad idea. So for the majority of this thesis I will assume that a patient must be deceived about the inert nature of a placebo treatment for that placebo treatment to be effective. However, in chapter seven I examine the possibility of non-deceptive placebo treatments, and discuss the implications this possibility may have for the conclusions I reach in the intervening chapters.

The other unusual point about Sylvester Colligan's placebo intervention is that it carried with it some specific side-effects: the pain and the scarring caused by the incisions, as well as the side-effects of the general anaesthetic. A sugar pill or saline injection, on the other hand, would be expected to have no such side-effects. For the sake of simplicity, I will assume that the placebo treatments I discuss here are inert, and carry no risk of side effects. However, in chapter seven I will also examine the possibility of adverse reactions to inert placebo treatments through the nocebo effect, and discuss the implications this could have for my conclusions.

For the majority of this thesis, then, I will assume that a placebo is an inert treatment that would be ineffective if its nature was disclosed, and which carries no risk of side-effects. I will not, however, merely assume that a placebo treatment can offer a real therapeutic benefit to patients. Rather, I will argue in the next chapter that we have good reason for thinking that this might be the case.

The ethics of placebo use in clinical practice

In chapter two I will discuss the historical use of placebos in medicine. I will also review current literature on placebos, as well as theories of the placebo effect and the possible mechanisms of action underlying this effect. I conclude that placebos and the placebo effect have always been a part of medicine, and will be a part of medicine for the foreseeable future. Moreover, I find that the possible effects of a placebo treatment are not limited to pleasing a patient or making them think that they have seen some improvement, but that we have good reason to believe that a placebo treatment can also lead to an actual therapeutic benefit for patients.

If it is possible for a therapeutic placebo treatment to provide a real therapeutic benefit for patients, then it must also be possible for that treatment to present the best, or only, chance at achieving this therapeutic benefit. This situation might arise because, as in the previous example of osteoarthritis of the knee, the active interventions available carry more serious side-effects than a placebo treatment, but are not expected to provide any greater therapeutic benefit than a placebo. Or it might be that case that there simply are no active treatments available, so the choice is between a placebo treatment or nothing at all.

These are the cases in which I argue that a doctor or nurse may permissibly administer a deceptive placebo treatment.

In chapter three I review the current ethical guidelines concerning the use of deceptive placebo treatments in clinical practice. Interestingly, while the use of placebo controls in clinical trials is the subject of strict legislation worldwide, guidelines on placebo use in clinical practice are few and far between. So far, only the American Medical Association (AMA), the British Medical Association (BMA), and the German Bundesärztekammer (BÄK) have issued advice on clinical placebo use. To make things worse, this advice is inconsistent: the AMA and the BMA uniformly condemn deceptive placebo use in clinical practice, while the BÄK recommends that deceptive placebos be used in some circumstances. The AMA and the BMA policies, however, share three main points that I believe constitute the main ethical arguments against deceptive placebo use in clinical practice:

- Deceptive placebo treatments are contrary to respect for patient autonomy.
- Deceptive placebo treatments do not meet the standards of informed consent.
- Deceptive placebo treatments undermine patient trust, putting the doctor-patient relationship at risk.

I will return to each of these claims in chapters four, five, and six, respectively.

Also in chapter three I review the current literature on the prevalence of, and attitudes to, clinical placebo use around the world. Despite the variation in ethical guidelines (or lack thereof) between countries, I find that doctors and nurses all over the world appear to be

using placebos with some enthusiasm. Moreover, the attitudes of physicians and patients in various countries towards placebo treatments seem to be fairly consistently positive. I conclude that the ethical question of clinical placebo use is far from hypothetical: it appears that this is a decision being made every day by doctors and nurses around the world.

In chapter four I turn to the first of the main arguments against clinical placebo use, and examine the question of whether the deception involved in administering a placebo treatment is a violation of patient autonomy. I investigate this question from a Kantian standpoint, on the grounds that a Kantian moral framework, with its strict rules on truthfulness and the supreme value it places on rational agency, constitutes the toughest possible test for a course of action that involves deception. If a deceptive placebo treatment passes this test, I argue, then it could also be expected to pass the less rigorous tests presented by other moral frameworks.

Upon examination of two competing accounts of Kant's categorical imperative, I find that the features peculiar to placebo deception mean that usual Kantian arguments against deceptive maxims do not apply. Firstly, I argue that a maxim of deceptive placebo use can pass a traditional interpretation of Kant's Categorial Imperative test, as outlined in the work of Christine Korsgaard. Upon examination of Barbara Herman's interpretation of the same test, however, it appears that the Kantian prohibition on interfering with the will of rational agents presents a larger obstacle. But an analysis of the effect of placebo deception on an agent's will shows why this particular kind of deception might not violate this rule. Moreover this analysis highlights the features that separate morally permissible cases of placebo deception from morally impermissible ones, and from cases of paternalistic

deception more generally.

The resulting Kantian defence of clinical placebo use is specific to cases where a placebo treatment is administered for its therapeutic effect, and where the deception involved is limited to withholding information on the inert nature of the treatment in order to facilitate this effect. Other cases, for example where a patient is lied to, where a placebo treatment is used to mollify a difficult patient, or where deception is used to give a patient a paternalistic “nudge” towards consenting to an active treatment, are impermissible. This distinction is possible because placebo deception, in the ideal case, is being carried out not only in order to help an agent achieve an end of their own, but it in order to facilitate a means towards that end – the placebo effect – that cannot be willed, and would be unavailable to that agent in the absence of deception. The act of deception involved in an ideal placebo treatment does not pervert the functioning of an agent's will, it does not interfere with or influence the agent's process of choosing ends or means, it just makes available to that process a means that would otherwise be out of reach. It is therefore consistent with respect for patient autonomy.

In chapter five I argue that it is possible for a deceptive placebo treatment to meet the standards of informed consent in at least some cases. Examining various theories and standards of informed consent, I find that the aims and likely risks of a treatment constitute information without which valid informed consent could not be given. However, the same cannot be said of information on the exact nature or mechanism of action of a treatment – in at least some cases patients can give valid informed consent to a treatment without knowing its exact nature or mechanism of action. This is true, I argue, whether that treatment is an active treatment or a placebo treatment. Therefore, withholding information

on the inert nature of a placebo treatment does not make informed consent impossible. Moreover, any rule that requires disclosing the inert nature of all placebo treatments effectively sets the informed consent standard for placebo treatments higher than for active treatments.

Whether or not the level of disclosure necessary to facilitate a therapeutic placebo effect meets the informational needs of any individual patient should be decided on a case-by-case basis, a criterion known as the “subjective standard” of informed consent. While adhering to this standard of consent requires a level of familiarity between clinician and patient that is impractical in many healthcare situations, this level of familiarity is appropriate in most situations where a deceptive placebo treatment should be considered an option. This is because this level of familiarity allows a clinician to judge not only the information that it is appropriate to disclose, but also to more accurately judge the likely risks and benefits of a placebo treatment.

In chapter six I examine the argument that the deception involved in a placebo treatment constitutes a betrayal of the trust between practitioner and patient. The idea that honesty and trustworthiness are closely linked, and that deception constitutes a failure of trust, is strongly intuitive. However, upon examination of several philosophical accounts of trust I find that the relationship between honesty, trust and trustworthiness is not something that has been examined in much depth. I therefore propose my own theory of the relationship between trust and deception.

Drawing on Karen Jones' theory of rich trustworthiness I argue that part of being trustworthy is communicating this trustworthiness to those who can trust you. This means

that the judgement of another's trustworthiness often boils down to judging the honesty of their communications of that trustworthiness. Anything that affects the judgement of another's honesty is therefore likely to affect the judgement of their trustworthiness. So if past acts of deception mean we judge someone as less than honest, then, it may also mean that we judge them as less than trustworthy. Furthermore, insofar as these acts of deception constitute a failure to communicate trustworthiness, they therefore constitute a failure to be trustworthy, in the rich sense.

According to this account of trust it is not clear that lies and deception in a medical relationship constitute a betrayal of patient trust. However it is possible that lies and deception constitute a failure of the practitioner to be trustworthy, and that they will lead to a loss of trust in the relationship. This is particularly true of outright lies, however, and I argue that it need not be the case where deception is limited to the withholding of information specific to a placebo treatment.

Lastly, in chapter seven I examine the implications of the two assumptions that I mentioned earlier: that placebo treatments do not carry the risk of side-effects, and that placebos are necessarily deceptive. I argue that both assumptions might not hold in a real-world case of placebo use. Firstly, a placebo treatment has the potential to lead to adverse effects through the nocebo effect, where the factors that are usually responsible for a beneficial placebo response instead produce an adverse reaction. And secondly some recent studies have indicated that a placebo treatment may have a beneficial result even when patients are informed about the treatment's inert nature. However, I argue that the possibility of a nocebo effect is not a problem that is specific to placebo treatments, as these effects can also occur in active treatments. The risk of nocebo side-effects therefore

does not constitute a reason to avoid administering a placebo treatment where that is believed to be the best or only treatment available. And while the prospect of effective non-deceptive placebo treatments has the potential to overturn many of the arguments I have made so far, at present there is simply not enough evidence to support the conclusion that non-deceptive placebos can be therapeutically effective. I therefore conclude that neither of these possibilities should alter my overall conclusion that it is permissible for a doctor or nurse to administer a deceptive placebo to a patient, if it is believed that this treatment presents that patient's best, or only, chance at some therapeutic benefit.

References

Cullen KA, Hall MJ, Golosinskiy A (2009) “Ambulatory surgery in the United States, 2006” *National Health Statistics Reports* 11: 1-25

Moseley JB et al (2002) “A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee” *New England Journal of Medicine* 347(2): 81-88

Shapiro and Shapiro (1997), *The Powerful Placebo*, Baltimore: Johns Hopkins University Press

Talbot M (2000), “The Placebo Prescription” *New York Times* January 9

2. History and theory of the placebo effect

Mentiris ut medicus [lie like a doctor]

– Medieval proverb

Writing on the possibility of exploiting placebo treatments for therapeutic benefit, Colin Cheyne argues that we encounter two constraints: an epistemic constraint and an ethical constraint.[1] The epistemic constraint is that we cannot employ placebo treatments unless we know (at least to some extent) what they are and how they work, whereas the ethical constraint is that we cannot employ placebo treatments if they violate accepted ethical norms. In this chapter I hope to go some way towards addressing the epistemic constraint – to give a brief account of what a placebo treatment is, why we have reason to believe that placebo treatments can work in at least some cases, and how placebo treatments might work in such cases. To do this I will briefly discuss the history of placebos and their place in medicine in the past. Then I will – again, briefly – discuss current theories on the placebo effect, and its mechanism of action. Although both of these topics are beyond the scope of this thesis, I hope to discuss them in sufficient depth to give some context to my later discussion of the ethics of placebo use in contemporary clinical practice. The two main points I wish to establish in this discussion are as follows: firstly, that the placebo effect is a feature of medicine that has been present for the whole of medical history. And secondly, that the effect of a placebo treatment is not limited to pleasing a patient or making them think that they have improved, but that we have good reason to believe that a placebo treatment can also lead to an actual therapeutic benefit for patients.

1 Cheyne (2005) p178

A brief history of the placebo effect

In many ways, the history of medicine is a history of the placebo effect. Possibly the first person to make this argument was a physician named WR Houston, who in 1937 gave an address to the American College of Physicians, which was published in a 1938 article called "The Doctor Himself as a Therapeutic Agent".[2] In this address Houston points out that, as was well known by that time, in the whole history of Western medicine, from Hippocrates up till about the middle of the 19th century, there exists only a handful of drugs that we now think would be effective at treating the conditions for which they were used:

> Historians sentimentalize the practical values of ancient medicine. One scans the pages of Hippocrates in vain for any treatments of specific value. The pages of medical history read like the log of an old-fashioned ocean voyage, in which it is noted that on such a day a whale spouted, on such a day a flying fish was sighted, or a bit of driftwood, but in which no mention is made of the huge prevailing fact that what was constantly seen day by day, almost to the exclusion of other sights, was the unending green waste of water.[3]

Fresh fruit for scurvy, quinine for malaria, digitalis for cardiac conditions, willow bark and opium for pain: these are more or less the entirety of the effective agents to be found in Western medical history prior to the modern era.[4] Despite pre-modern doctors' knowledge of, and contribution to, the sciences of anatomy, physiology, and pathology, this handful of drugs comprised their entire active pharmacopoeia. The rest of the arsenal of medicines used by pre-modern doctors was, by 1937, known to consist entirely of substances with no specific activity for the conditions which they were used to treat. This means that the overwhelming majority of the thousands and thousands of different pills and potions, oils and lotions, clysters, poultices, unguents, drops, draughts and vapours used by doctors for

2 Houston WR (1938)
3 Ibid. p1417
4 Shapiro and Shapiro (1997) p60-71

all these years were placebos.

The definition of “placebo” I'd like to adopt in this work is that used by Shapiro and Shaipro in their book *The Powerful Placebo*:

> A placebo is any therapy (or that component of any therapy) that is intentionally or knowingly used for its nonspecific, psychological, or psychophysiological, therapeutic effect, or that is used for a presumed specific therapeutic effect on a patient, symptom, or illness, but is without specific activity for the condition being treated.[5]

As well as sugar pills, saline injections, sham surgeries and the like, this definition includes treatments that are administered in the belief that they have some specific activity for the condition being treated, but which later turn out to have no specific activity for that condition. Pre-modern doctors' belief in the active nature of their treatments, then, does not change the fact that they were actually administering placebos.

In this work I'll be using the terms “placebo effect” and “placebo response” interchangeably. Some have argued for a distinction between these two terms, with “placebo effect” referring to any improvement in an individual or group that has taken a placebo, and “placebo response” referring only to the specific change caused by a placebo treatment – that is, the “placebo effect” minus any spontaneous remission, regression to the mean, biased reporting and so on.[6] However, using the terms this way means that a good portion of what would be called the “placebo effect” could be observed in an individual or group that has been given no treatment at all, which I think is stretching the definition of “placebo effect” a bit. Accordingly, then, I will be using both “placebo effect” and “placebo response” to refer to the non-specific psychological or psychophysiological

5 Ibid. p41
6 Benedetti F (2009) p31

effect that can result from a placebo treatment.

The non-active nature of the vast majority of the drugs in a pre-modern doctor's bag was common knowledge amongst doctors by 1937. However, another fact about pre-modern doctors that Houston points out is common knowledge to just about everybody: by and large, these physicians were not as ineffective as the nature of the treatments available to them might lead you to expect. Pre-modern doctors were not scruffy mountebanks existing on the fringes of society, peddling quackery to the desperate or the unwary – to the contrary, they were important and respected members of their communities. That pre-modern doctors managed to reach and maintain this level of respect and status while employing a pharmacopoeia that consisted almost entirely of placebos is remarkable enough, but it is particularly surprising when you consider that the placebos employed by these physicians were not as innocuous as the sugar pills or saline injections that that term brings to mind today. Nor were these placebos limited to the weird and wonderful, but probably fairly harmless, agents like ground Egyptian mummy, usnea (moss scraped from the skull of a hanged criminal), and unicorn's horn. Placebos used by pre-modern doctors frequently included ingredients that were actually harmful to the patient, such as mercury, putrid meat, and human excrement,[7] as well as damaging treatments like purging, sweating, blistering and bloodletting.[8] With placebos like these in common use, it’s a wonder that pre-modern patients even survived their treatments, let alone experienced some therapeutic benefit from them.

Even philosophers were not averse to administering the odd placebo or two in pre-modern times. George Berkeley, the great idealist philosopher, dedicated the latter portion of his

7 Honigfeld (1964) p145
8 Shapiro and Shapiro (1997) p18

life to promoting and administering a treatment called tar water, which was water shaken up with pine tar. His last work, *Siris,* is a meditation on (amongst other things) metaphysics, science, and the nature of God. But the first half of this work is taken up by a discussion of the preparation of tar water and its usefulness for treating a wide range of conditions. Berkeley apparently found tar water to be quite a successful placebo – he notes that in 1741 “twenty-five fevers in my own family [were] cured by this medicinal water, drunk copiously”.[9] And indeed, if you consider some of the more hair-raising treatments that were routinely employed by physicians in the eighteenth century, you can see why a treatment that basically amounted to rest and adequate hydration might seem like a miraculous panacea. But unless you had the good fortune to be living in Cloyne, an eighteenth-century trip to the doctor’s office would have been more likely to leave you bled and poisoned than rested and hydrated. “By and large,” argues Houston, “the doctors were, as reported by that sane and shrewd observer, Montaigne, a danger to their patients.”[10]

To what, then, can we attribute (in Houston's words) “the honor and emolument bestowed on the physician through the ages”?[11] From all appearances, up until fairly recently any enlightened community would have had better reason to bestow tar and feathers on their physicians than honour and emolument. Houston discounts the idea that the placebos themselves could have had anything to do with the success of pre-modern physicians, instead concluding that doctors themselves, or more specifically the relationship between doctor and patient, was the therapeutic agent by which enough good outcomes were effected to keep doctors in society's good graces. While the doctor-patient relationship is still thought to have immense therapeutic value, the years since Houston’s paper was

9 Berkeley G (1747)
10 Houston (1938) p1418
11 Ibid.

published have seen a growing recognition that there is something about a placebo treatment, a placebo effect or response, that also has its part in explaining how inert treatments (or worse) can nonetheless lead to therapeutic outcomes. In Houston's words again: “Through centuries when doctors were doing more harm than good this dynamic force has sustained the medical profession in the esteem of their clientele, it has inspired their fellow citizens with such faith in its values that they were willing to give economic support to the doctor.”[12]

While the modern physician enjoys a far greater selection of effective treatments and active drugs than her pre-modern counterpart, we would be naïve to think that a modern physician's armamentarium is free of inadvertent placebo treatments. It is far from unheard of for a commonly-used treatment to “fail” a standardised placebo-controlled trial – that is, to be shown to be no more effective than a placebo. And if a treatment is no more effective than a placebo treatment, it is reasonable to conclude that any success it has enjoyed in the past is due to the placebo effect, regardless of the active ingredients it contains. To pick a few common treatments that have recently had their specific activity called into question in this way: Tamiflu, the anti-influenza drug which governments around the world spent billions stockpiling as part of preparations for a possible influenza pandemic;[13] antidepressants, when prescribed for mild to moderate depression;[14] and arthroscopic surgery for osteoarthritis of the knee, the most common orthopaedic surgical procedure carried out in the US, and one of the most common procedures worldwide.[15] All these treatments have been administered to patients by clinicians who believed them to be effective, and by and large they have been effective – at least sufficiently so that they

12 Ibid.
13 Jefferson et al (2009)
14 Kirsch et al (2008)
15 J. Bruce Moseley et al (2002)

continued to be administered to millions of patients worldwide. But if the cited studies are correct, the success of these these treatments can be wholly attributed to the same factor that led to the success of a pre-modern physician's treatments of owl's blood or unicorn's horn: the placebo effect. So while there may be debate about the strength or usefulness of this effect, its place in modern medicine, just like its place in pre-modern medicine, is undeniable.

The placebo effect, then, is a therapeutic measure that could be called upon by the ancient healer and the modern physician alike. It is a thread that has run through the practice of medicine from its infancy, and it may well have been responsible for maintaining the reputation of medical profession for the millennia that passed before there existed knowledge of more than a handful of effective drugs to treat the ailments that plagued the populace. While this history is fascinating, and stands testament to the power of the placebo effect, these inadvertent placebos are not the kind of placebo treatment that I am concerned with in this thesis. A practitioner who administers to a patient a treatment that they believe is active, but which is found later to be a placebo, is not engaging in the kind of deception that I wish to evaluate here. My interest is in cases of deceptive placebo use, where a doctor or nurse knows that a treatment is a placebo, but believes that it nonetheless presents a patient’s best or only chance at some therapeutic benefit. If antidepressants are shown to be no better than a placebo for treating mild to moderate depression, what is a practitioner to do when faced with a patient with this kind of depression? The fact that antidepressants “failed” in standardised controlled trials does not mean that you would do just as well to forgo treatment altogether – while the active treatment available might have been shown to return results not significantly better than a placebo, both of these treatments might still have been shown to be significantly more effective than no treatment at all. But in addition to their therapeutic benefit,

antidepressants can have some unpleasant side-effects for the patient, which is a risk not found in inert treatments. On the other hand however, administering a placebo treatment requires deceiving the patient about the nature of this treatment. Whether or not this deception can be justified is the question I'm interested in examining in the following chapters of this essay, and thus for the rest of this chapter I will focus on deceptive placebos, rather than inadvertent ones.

The kind of placebos I'm interested in here, deceptive placebos, have also played their part in the history of medicine. References to treating patients with bread pills, mint water, and other placebos can be found in the writings of 18th century doctors like William Cullen,[16] John Hunter,[17] Alexander Gordon,[18] and Thomas Trotter.[19] In 1807 Thomas Jefferson wrote in a letter that “One of the most successful physicians I have ever known has assured me that he used more bread pills, drops of coloured water, and powders of hickory ash than of all other medicines put together. It was certainly a pious fraud.”[20] Richard Cabot, a professor of medicine at Harvard, described in a 1902 address to the Academy of Medicine in New York[21] how he was “brought up, as I suppose every physician is, to use placebos”,[22] and how he subsequently employed them in his practice “by the bushel”.[23] Indeed, the term “placebo” is generally only used to refer to deceptive placebos, and has been associated with deception for almost as long as it's been part of the English language.

16 Cullen W (1772) p218-219
17 Hunter J (1788) p69-70
18 Gordon A (1788)
19 Trotter T (1792) p184
20 Ford (ed.) (1898) p427
21 Cabot RC (1903)
22 Ibid. p349
23 Ibid.

Placebo etymology

The term “placebo” originates from the Latin for “I shall please”. It entered the English language by way of Psalm 116 of St Jerome's Latin Vulgate Bible[24], the ninth verse of which reads “Placebo Domino in regione vivorum”, or “I shall be pleasing to the Lord in the land of the living.”[25] In the middle ages this Latin verse was sung by the congregation at funerals as part of the Catholic “Office for the Dead”. At the time, the families of some deceased people would hire professional mourners to attend the funeral, to add their voices to the chorus and bulk out the numbers in the crowd. These professional mourners became known as “placebo singers” after the first verse they'd have to sing. As you might imagine, “placebo singers” were not generally held in high regard: the sincerity of their funeral prayers was highly questionable, and the impression they gave of the popularity of the deceased was flattering to say the least. Thus the word “placebo” came to be a derogatory term meaning a sycophant, flatterer, or parasite.[26] It can be found in this sense in Chaucer's *Canterbury Tales*, where “Placebo” is both the name of January's sycophantic brother in the Merchant's Tale, and part of an admonition of flatterers in the Parson's Tale: “Flatterers are the devil's chaplains, that sing ever ‘Placebo'“.

The word “placebo”, then, signified sycophancy and flattery, and was thus associated not just with pleasing or satisfying another, but with acts of deception in order to do so. One of the earliest recorded uses of the term in medicine is consistent with this sense of the word: The 1811 edition of *Quincy's Lexicon-Medicum* defines “placebo” as “an epithet given to any medicine adopted more to please than to benefit the patient”.[27] The earliest known

24 Aronson J (1999) p716
25 This is now thought to be a mistranslation of the original Hebrew, and in more recent editions of the bible this verse reads “I shall walk with the Lord in the land of the living”.
26 Shapiro AK (1968)
27 Quincy, J (1811)

recorded use of the term “placebo” in a medical context, though, is in the lecture notes of the prominent Scottish physician William Cullen from 1772[28]. Cullen describes treating a paralytic patient with mustard, saying of the treatment “I own that I did not trust much to it, but I gave it because it is necessary to give a medicine, and as what I call a placebo.”[29] A little over a decade later, the 1785 edition of Motherby's *New Medical Dictionary* defines “placebo” as “a commonplace method or medicine”[30], suggesting that both the term and the practice were widespread by that time – although it's impossible to tell whether William Cullen's influence had served to popularise a term he had coined, or whether he had merely adopted a term that was already becoming common in medicine at the time.

The first recorded use of the term “placebo effect” does not come until 1920, when TC Graves writes in *The Lancet* about treating a patient with testicular extract. He notes that such treatments “have been regarded as having a psychotherapeutic value as placebos”, but that in this case "a real psychotherapeutic effect appears to have been produced, especially when one notes the placebo effects of the drugs given prior to admission”.[31] The decades following the publication of this paper saw an increasing recognition of and interest in the placebo effect, which can be seen in the rising numbers of clinical trials employing placebo controls.

Placebo controls in clinical trials

Although the use of placebos in the context of a clinical trial is beyond the scope of this

28 Kerr et al. (2008)
29 Ibid.p90
30 Shapiro (1968) p654
31 Graves, TC (1920)

essay, the adoption of placebo controls in clinical trials is important historically as it signifies the growing awareness in the medical community about the placebo effect – that there is some effect caused by receiving an inert treatment that needed to be controlled for. The first actual use of inert controls in a trial is widely thought to have taken place in Benjamin Franklin and Antoine Lavoisier's trial of Franz Mesmer's “animal magnetism” techniques in 1784,[32] where the effects on patients of supposedly “magnetised” water was compared to that of ordinary water which patients merely thought had been treated. Another high-profile controlled trial took place a decade and a half after this in 1799, when English physician John Haygarth tested the effectiveness of “Perkins tractors”: twin metal rods which were supposed to effect a cure by conducting away a disease-causing electrical fluid in the patient's body. Haygarth compared the effects of this procedure to those of a sham treatment using rods that looked and felt similar to the original, but were made from non-conducting wood.[33] Both of these trials were able to reveal that the treatments under study were nostrums, and that the success these procedures had enjoyed previously were down to the patient's mental state, rather than any active treatment effects.

The first use of the term “placebo” to refer to the inert control used in a trial came in 1854,[34] when American physician Austin Flint tested the effectiveness of a rheumatism treatment against that of a placebo:

> The recorded cases were treated throughout the whole course of the disease with only palliative measures... [b]ut to secure the moral effect of a remedy given specially for the disease, the patients were placed on the use of a placebo which consisted, in nearly all of the cases, of the tincture of quassia, very largely diluted. This was given regularly, and became well known in my

32 Franklin, B (1785)
33 Haygarth J (1801)
34 Flint A (1863)

> wards as the placeboic remedy for rheumatism. The favourable progress of the cases was such as to secure for the remedy generally the entire confidence of the patients.[35]

Despite the success of these seminal controlled trials in exposing quackery, and in highlighting the effect that the patient's state of mind can have on a treatment, placebo treatments remained a disreputable secret in medicine. Inert controls were employed in trials only sporadically before the mid-20th century, and these inert treatments were rarely, if ever, referred to as “placebos” in the literature.

While a smattering clinical trials at the beginning of the 20th century employed inert controls,[36] the 1930s saw the incidence of controlled trials increase markedly. The 1930s also saw the regular adoption of the term “placebo” to refer to the inert control used in a clinical trial[37], an indication that researchers realised that the placebo effect was one of the factors that needed to be controlled for. Likewise, a 1938 placebo-controlled trial of cold vaccines[38] returned a negative result for the vaccines, but a positive result for the placebo effect. While the vaccine arm of the trial did not show any significant improvement over that of the control arm, the improvement shown by both groups was comparable to that shown by the participants in previous uncontrolled trials of the vaccines, where the interventions had been judged to be a success. This led the authors to report that

> one of the most significant aspects of this study is the great reduction in the number of colds which the members of the control groups reported during the experimental period. In fact these results were as good as many of those reported in uncontrolled studies which recommended the use of cold vaccines.[39]

Also in the 1930s, an American pharmacologist named Harry Gold led a team carrying out

35 Ibid. p21
36 Shapiro and Shapiro (1997) pp137-141
37 Evans W, Hoyle C (1933)
38 Diehl HS et al (1938)
39 Ibid.p.1172

a 5-year controlled trial of xanthines for angina.[40] As this trial progressed, the researchers realised that the physicians who were tasked with administering the drug and questioning the patient about its effect on their pain behaved differently depending on whether they had administered the active treatment or the placebo treatment. Realising this could bias the study, they began to "blind" nature of the treatment not only to the participants, but also to the researchers – the first use of double-blinding in a placebo-controlled trial. Following the conclusion of this trial in 1937, all the clinical studies in Gold's laboratory were carried out using a double-blind placebo-controlled methodology.[41] Gold remarked in a 1968 interview that "the placebo and the double-blind were companions after the 1937 study."[42] Gold went on to become an important advocate for recognition of the placebo effect and the need for placebo controls in clinical trials: in 1946 he convened a conference entitled "The use of placebos in therapy", where he argued that the effect of placebo treatments had to be accounted for and controlled in the clinical trial of any new medication.[43]

Shapiro and Shapiro argue that the increasing use of placebo-controlled trials in the 1930s and 1940s helped give legitimacy to the deceptive placebo, finding a place in respectable medicine for a treatment which had previously been viewed with derision, as something "given by unethical clinicians to satisfy or fool a patient".[44] The advent of the placebo-controlled trial meant that respectable medical researchers could now be seen employing placebo treatments, and writing about the results in august medical journals. However, despite these clinical observations of the placebo effect in action, the placebo effect as a phenomenon in medicine did not receive widespread interest or recognition until 1955, when the American anaesthetist and Harvard professor Henry Beecher published his

40 Gold H, Kwit NT, Otto H (1937)
41 Shapiro and Shapiro (1997) p142-143
42 Shapiro and Shapiro (1997) p143
43 Gold H (1946)
44 Shapiro and Shapiro (1997) p144

landmark paper “The Powerful Placebo”.

Henry Beecher and “The Powerful Placebo”

Henry Beecher became interested in the placebo effect while he was serving in Allied field hospitals during World War II. Morphine was frequently in short supply, and on one occasion supplies ran out entirely while Beecher was preparing to operate on a soldier with terrible injuries. Without morphine the operation would be extremely painful, and Beecher worried that it might even cause the patient to undergo a fatal cardiovascular shock. Lacking any other options, a nurse injected the soldier with a syringe full of saline. To everybody's surprise, the patient reacted to this placebo shot as though he had been given active pain medication, and the operation was able to continue. The patient appeared to experience significant pain relief, and did not go into cardiovascular shock. For the remainder of his tour, Beecher repeated this trick whenever supplies of morphine ran out. He observed that it worked – not every time, but frequently enough that a saline shot was definitely preferable to offering no treatment to a patient in pain.[45]

Beecher pursued his newfound interest in the therapeutic possibilities of placebos when he returned to his career after the war ended. His investigations culminated in his 1955 paper “The Powerful Placebo”[46], a meta-analysis of 15 studies where a placebo treatment was compared to an active treatment, for a range of conditions. Beecher concluded that:

> It is evident that placebos have a high degree of therapeutic effectiveness in treating subjective responses, decided improvement, interpreted under the unknowns technique [blinding] as a real

45 Evans, D (2003) p1
46 Beecher, HK (1955)

> therapeutic effect, being produced in 35.2 ± 2.2% of cases. This is shown in over 1,000 patients in 15 studies covering a wide variety of areas: wound pain, the pain of angina pectoris, headache, nausea, phenomena related to cough and to drug-induced mood changes, anxiety and tension, and finally the common cold.[47]

The influence of Beecher's paper cannot be overstated. It quantified[48] the placebo effect, and cemented it as scientific fact. It is still cited in nearly every scientific paper written on the placebo effect.[49] This work was also the grounds on which Beecher, Harry Gold, and other medical reformers successfully campaigned to have placebo arms included in clinical trials, and for the double-blind randomised controlled trial to become the standard test for new medications. In his paper, Beecher argues

> it should be apparent that “clinical impression” is hardly a dependable source of information without essential safeguards of the double unknowns technique [double-blinding], the use of placebos also as unknowns, randomization of administration, the use of correlated data (all agents studied in the same patients), and mathematical validation of any supposed differences.[50]

It is thus due in large part to Beecher's work that the randomised, double blind, placebo-controlled trial was established as the gold standard for medical research. Interestingly, for such a landmark study, Beecher's paper is seriously flawed. The flaws of Beecher's study were raised in a 1997 paper by Kienle and Keine,[51] where they point out that, among other problems, 13 out of of the 15 studies which Beecher includes in his meta-analysis do not have no-treatment controls. This means that, for those 13 studies, the improvement Beecher attributes to the placebo effect might well have occurred due to the natural

47 Ibid. p1606
48 However, Beecher's conclusion that 35% of patients experienced some therapeutic benefit from a placebo treatment is often misinterpreted to mean that 35% of the therapeutic effect of an active treatment is due to the placebo effect.
49 Evans (2003) p.4
50 Beecher, HK (1955) p1605
51 Kienle, G and Kiene, H (1997)

progression of the condition, the patient's own healing processes, or any number of other factors.

The only way to isolate the placebo effect as the cause of an observed change in a group is to compare the group that receives a placebo treatment with a group that receives no treatment at all. For example, a controlled trial of a cold remedy might show that an active treatment is about as effective as a placebo treatment, with both arms of the trial experiencing complete recovery in around seven days. It would however be inaccurate to attribute this recovery to the placebo effect, as seven days seems like a fairly reasonable time for an untreated cold to clear up all by itself. On the other hand, if this trial had a third arm, a control group of patients who received no treatment at all, then the recovery of patients in the active treatment and placebo groups can be compared with the natural progression of the disease. Any statistically significant difference in the recovery time between the no-treatment group and the placebo group could then be attributed to the placebo effect. Unfortunately, all but two of the studies in Beecher's analysis fail to include a no-treatment control arm, so the change in condition of the placebo group in the majority of the trials he studied cannot be attributed to the placebo effect. Ironically, Beecher's study falls victim to the same *post hoc ergo propter hoc* fallacy that he identified in those clinical trials that lack a placebo control.

Ted Kaptchuk[52] argues that the skewed results in Beecher's paper were not merely the result of sloppy research. Rather, he posits that Beecher's goal in publishing “The Powerful Placebo” was not to promote understanding of the placebo effect, but instead to demonstrate the necessity of adopting randomised placebo-controlled trials in medical research. In order to do this, Beecher may have deliberately fudged his results in order to

52 Kaptchuk T (1998)

achieve a conclusion that would be more useful to proponents of randomised controlled trials. After all, the greater the power of the placebo effect, the more pressing the need to control for it in clinical trials, and the easier it would be for medical reformers to convince their colleagues to change their ways. This could explain why Beecher ignored factors that he must have been aware of, like regression to the mean and the natural progression of conditions, and also why he left out of his analysis some well-designed studies that included both placebo and no-treatment arms, but didn't support his conclusion.[53] But while Beecher was undoubtedly correct that the widespread adoption of randomised controlled trials was necessary in order to impose scientific rigour on post-war medical research, and while his paper was instrumental in achieving this, the flaws in his research mean that Beecher's oft-cited conclusion that 35% of patients receive a real therapeutic effect from a placebo treatment looks to be groundless.

Is the placebo effect "real"?

Is there any basis, then, on which to argue that placebo treatments are capable of effecting a real therapeutic change in a patient? The story I have told so far, of the essential part that the placebo effect has played in the history of medicine, will do little to sway those sceptical about the placebo effect. Placebo critics can argue that it is more likely that the placebos administered by physicians and researchers over the centuries have merely served to trick patients into thinking that they felt better, rather than creating a real therapeutic benefit for them. The mere impression of therapeutic benefit could explain how placebo treatments sustained the medical profession though its pre-modern years, without invoking some kind of spooky psychophysiological response to inert treatments.

53 Ibid. p1724

This criticism is given further credence by the fact that the placebo effect is most strongly associated with the relief of pain, depression, and other conditions where the measures used are subjective rather than directly observable.

On the one hand it is tempting to dismiss this criticism as relying on a false distinction – broadly speaking, a distinction between “getting better” and merely “feeling better”. What is the difference between a patient actually feeling less pain following a placebo treatment, and that patient merely feeling like they feel less pain? One might similarly ask where the distinction lies between feeling less depressed and actually being less depressed. But here we can see a difference: while the purpose of an analgesic is solely to reduce the feeling of pain in a patient, the purpose of an antidepressant is presumably to address the underlying cause of the feelings of depression. So the question becomes whether a placebo treatment for depression merely makes the patient think they feel better, or whether it can actually result in, say, increased serotonin levels in the patient's brain – an actual therapeutic benefit.

This question, whether placebo treatments can effect a real therapeutic benefit for patients, has important consequences for the ethical discussion I wish to have later in this essay. A deceptive placebo treatment always involves deception about the inert nature of the treatment, but if placebo treatments can offer no actual therapeutic benefit to a patient, then a placebo treatment will also involve deception about the purpose and likely effects of the treatment. For example, a clinician can offer a patient a placebo shot of saline, and tell them honestly “this shot is to help with your pain”, as saline shots have been shown to have an analgesic effect. In this way the patient is informed about the purpose and likely effect of the treatment, and the deception involved is limited to concealing the treatment's

inert nature. But if it is not possible for a placebo antidepressant to lead to an actual therapeutic result for a depressed patient, then offering that patient a placebo antidepressant will require a greater degree of deception. “This pill will help with your condition” is a lie, and as I will argue in later chapters, I do not believe that lying to patients in order to promote a placebo response can be justified. On the other hand, telling the patient honestly “this pill won't affect your underlying condition, but it might make you feel better” is unlikely to inspire the kind of confidence upon which the placebo effect is thought to rely.

Arguments against the therapeutic benefit of placebo treatments often point to the fact that the placebo effect is most commonly associated with conditions where the outcomes are measured subjectively, such as pain and depression, which is exactly what you'd expect if placebo treatments are only able to make patients think they feel better, rather than effecting some real therapeutic benefit. However, significant placebo effects have been observed in studies where the outcomes are measured objectively by the researchers, rather than reported subjectively by patients. For example, placebos have been shown to increase dopamine levels in patients with Parkinson's disease,[54] reduce the size of duodenal ulcers,[55] reduce facial swelling following dental surgery,[56] and lower blood pressure in patients with hypertension.[57] Even in the case of placebo analgesia, researchers have used brain imaging techniques to show that subjective reports of placebo pain reduction are accompanied by changes in several regions of the brain involved in pain transmission.[58] However, observations like these were called into question by a high-profile study released by Hróbjartsson & Gøtzsche in 2001, which strongly

54 Lidstone SC et al. (2010)
55 De Craen AJM et al (1999)
56 Ho et al (1988)
57 Preston RA et al (2000)
58 Benedetti F (2009) p37

criticised the therapeutic potential of placebo treatments.[59]

Following Kienle and Keine's criticisms of Beecher's “Powerful Placebo” meta-analysis, Hróbjartsson & Gøtzsche carried out a far more extensive meta-analysis of 114 clinical trials, all of which included a no-treatment arm as well as placebo and active treatment arms. Their conclusion that the placebo was “powerless” was widely reported – even making the front page of the New York Times[60]. However, while the media jumped at the opportunity to report that the placebo effect was a myth, the actual conclusions drawn in the paper are a little more subtle. Hróbjartsson & Gøtzsche divided the studies they analysed into four distinct groups: whether the outcomes were measured subjectively (reported by the participants) or objectively (measured the by researchers), and whether those outcomes were binary (for example, dead or alive) or continuous (for example, the degree of pain). The researchers concluded that for binary outcomes there was no statistically significant placebo effect for either subjective or objective measures. For studies with continuous outcomes, however, there was a statistically significant placebo effect for subjective measures, but not for objective measures. They conclude that placebo treatments “had possible small benefits in studies with continuous subjective outcomes and for the treatment of pain”,[61] but that “[o]utside the setting of clinical trials, there is no justification for the use of placebos.”[62]

Although undoubtedly more rigorous than Beecher's original meta-analysis, Hróbjartsson & Gøtzsche's study has also seen its methods called into question. Wampold et al.[63] criticised Hróbjartsson & Gøtzsche's analysis on several grounds. They argued that many

59 Hróbjartsson A, Gøtzsche P (2001)
60 *New York Times* May 24, 2001
61 Hróbjartsson A, Gøtzsche P (2001) p1594
62 Ibid.
63 Wampold BE et al. (2005)

of the trials included in the meta-analysis were not well designed to detect a placebo effect. Some trials were unblinded or poorly blinded, something which would be expected to have a detrimental effect on a patient's placebo response[64]. Conversely, the placebo treatments in some trials were delivered surreptitiously, which again is not thought to be a good way to provoke a placebo effect[65]. And some trials simply studied conditions where a placebo effect would not be expected. One trial, for example, measured the analgesic effect of injections of glucose or sucrose against a placebo injection of sterile water on neonates.[66] However, there's no theory of the placebo effect that would predict that an injection of sterile water would produce a placebo response in a newborn baby.[67] When these studies were removed from the data set, Wampold et al re-did the meta-analysis and found that the placebo effect present in studies with objective measures was comparable to that found in studies with subjective measures, and in both cases the placebo effect was nearly as large as the treatment[68] effect.[69] The placebo effect, they concluded, is "robust."[70]

The debate between the researchers continues, with Hróbjartsson & Gøtzsche accusing Wampold et al of "powerful spin" in a 2007 article,[71] and Wampold et al responding in the same issue that this accusation was itself an attempt at spin on Hróbjartsson & Gøtzsche's part.[72] So what are we to make of all of this? With scientific debate proceeding at such a level, and no end in sight, clinicians could be forgiven for throwing their hands up in despair and walking away. But I would suggest that, due to the long history of placebo use

64 Ibid. p842
65 Ibid.
66 Wampold et al. (2007) p382
67 Ibid. p383
68 The placebo effect is measured as the difference in outcomes between the placebo group and the no-treatment group, while the treatment effect is measured as the difference in outcomes between the active treatment group and the placebo group. If the placebo effect is roughly the same size as the treatment effect, this suggests that around half the effect of taking an active medication is due to the placebo effect.
69 Wampold BE et al (2005) p845
70 Ibid. p850
71 Hróbjartsson A, Gøtzsche P (2007) p377
72 Wampold BE et al. (2007) p381ff

in medicine and in clinical trials, the view that the placebos can have a real therapeutic effect is the orthodox position, and that the burden of proof should lie on those who wish to overturn this view. Absent any decisive evidence to the contrary, then, we should at least accept that a therapeutic placebo effect is a possibility for some conditions. So I would conclude that, at present, if a clinician reads that a significant placebo effect has been observed in trials on patients with a certain condition, it would be reasonable for them to conclude that placebo treatments could actually have an effect on this condition, rather than writing off the results as bunkum or sloppiness.

It is not my place here to comment on the acceptable level of evidence required for a medical professional to believe in a certain treatment, but merely to argue that the requirement should not be greater for placebo treatments than for others. If, say, a peer-reviewed article in the *BMJ* constitutes good reason for a clinician to believe that a certain active treatment presents the best or only chance at achieving some therapeutic benefit for a patient with a certain condition, then this same standard of evidence should also constitute good reason for her to believe that a placebo treatment could provide some therapeutic benefit for that patient.

Placebos and Siberian Hamsters

Further evidence for the existence of a therapeutic placebo effect in humans is provided by the observation of a phenomenon similar to the placebo effect in other mammalian species.[73] One case of this in particular is interesting, because it possibly addresses a concern that lies behind much of the scepticism about the placebo effect: that it just

73 Evans D (2003) p99

doesn't seem to make sense that the human body has some self-healing capacity which is seemingly only activated by an inert treatment like a sugar pill. Why, if we have the capacity to do this, wouldn't it just happen by itself whenever it was needed? An answer to this question was suggested by a psychologist named Nicholas Humphery, and according to his theory it all comes down to the cost of mounting an immune response.[74]

The immune system is a very resource-intensive set of tissues and organs, and mounting a sustained immune response against an infection is a very costly activity for the human (or animal) body to undertake. Most of the symptoms we feel when suffering from a cold or flu, for example, are not caused by the virus running rampant through our tissues, but are actually the side-effects of the body's immune response. So in the experiment in question,[75] researchers decided to see whether this resource-intensive process could be moderated in response to environmental stimuli. Siberian hamsters were infected with a non-lethal disease under two different lighting conditions, designed to simulate either summer or winter. Those hamsters living in cages where the lighting was set to simulate summer, with long days and short nights, became ill in the normal way, mounted an immune response, and recovered. However, those hamsters living under winter lighting conditions, with long nights and only short periods of light, just tolerated the infection and went about their business as best they could.

Humphery's posited explanation for this observation is that, Siberian winters being what they are, any hamsters that took a few days off to do the hamster equivalent of sitting in bed drinking Lemsip and watching DVDs of Buffy the Vampire Slayer could well end up starving to death. In such conditions the short-term cost of mounting an immune response

74 Humphery N (2002) p255-288
75 Prendergast BJ et al (2002)

against a non-lethal infection could well outweigh the long-term benefits of not being infected. In summer, however, conditions are less harsh, food is more plentiful, and the cost-benefit analysis starts to swing in favour of taking a few days off sick while your immune system does its thing.[76] The lighting conditions in the hamsters' cages, then, can act as a cue the animals to activate or suppress their immune systems.

There is no reason to think that this process is specific to Siberian hamsters. Human beings respond to stressful conditions by producing the stress hormone cortisol, which serves to suppress the immune response.[77] This is why students will often report falling sick as soon as an exam period is finished, or right after handing in a big assignment. As soon as the stressful situation is over cortisol levels in the body drop, which means the immune system is no longer suppressed and it can go about mounting a response against any infection that the person might have picked up during the time of stress. Humphery argues that in people, as well as other mammals, the unconscious is capable of exerting control over the immune system, and it has evolved to assess the potential costs and benefits of mounting an immune response before giving it the go-ahead. The placebo effect comes about because we have come to associate treatments like pills and injections with the kind of therapeutic benefit that might make an immune response less costly than it normally is. So placebo treatments can function in the same way that summer lighting conditions do for Siberian hamsters, that is, as a signal to our unconscious mind indicating that the costs of mounting an immune response may now be outweighed by the potential benefits.

76 Trimmer et al (2013) have produced a mathematical model of this cost-benefit analysis.
77 See for example Webster Marketon JI and Glaser R (2008)

Mechanisms of the placebo effect

The assertion that placebos *can* have some therapeutic effect almost immediately raises the question of *how* they can have some therapeutic effect. How is it that a treatment characterised by its inert nature can nonetheless effect some change on a patient? What is the nature of the association we make between pills or injections and the therapeutic benefit of the drug they supposedly contain? This, too, can be a divisive question. Until fairly recently, there were two competing schools of thought regarding the psychological mechanism underlying the placebo effect: one held that the placebo effect was a result of unconscious conditioning, while the other held that conscious expectation of benefit was the mechanism that lay behind the placebo effect.

The classical conditioning theory of the placebo effect posits that the method of delivery of various treatments – pills, injections, syrups, etc – take the place of the conditional stimuli in a Pavlovian stimulus substitution pair, with the actual drug or treatment being delivered making up the unconditional stimuli.[78] To draw an analogy with Pavlov's famous experiment, the active treatments here are the dog food put in front of Pavlov's dogs, and the methods of their delivery are the sound of the bell ringing that accompanied the food. Over the course of your life, you unconsciously come to associate taking a pill or receiving a shot with the therapeutic benefit caused by the active drugs in these treatments, and so even an inert treatment can produce a conditioned response in your body, just as Pavlov was able to get his dogs to salivate by ringing a bell in the absence of any food.

Gobel et al showed how a classical conditioning-mediated placebo effect can can take

78 Montgomery GH, Kirsch I (1997) p107

place in human beings.[79] Every 12 hours for three days, a group of volunteers was given an immunosuppressive drug In the form of a pill, and a distinctive drink (green coloured strawberry milk with a drop of lavender oil added) to wash it down. Five days later, the participants were given the same drink and a placebo pill. A number of aspects of the volunteers' immune function were subsequently analysed and found to be suppressed, just as if they had taken the immunosuppressive drug.

By a classical conditioning model of the placebo effect, then, placebo treatments can be effective because they “piggy-back” on a patient's past experiences with effective treatment. Although this process could provide a good explanation for the placebo effect in the 20th century, it seems unlikely that classical conditioning could account for the effectiveness of the placebos used by pre-modern doctors. For how can you come to unconsciously associate treatment methods with their active therapeutic effects when barely any of the treatment methods available actually *have* any active therapeutic effects? For these pre-modern placebos, then, a conscious expectation of benefit seems like a more likely mechanism of action.

An expectancy model of the placebo effect results from studies suggesting that a person's response to a stimulus is affected by their expectations of the result of that stimulus.[80] Thus, if a person expects a certain outcome to result from taking a pill, this expectation makes it more likely for that outcome to occur – whether or not there is an active ingredient in the pill to promote that outcome. Whereas a conditioning model of the placebo response relies on past experience to influence the unconscious mind, conscious expectation can result from any number of factors: testimony from friends, family, or other trusted people,

79 Goebel MU et al (2002)
80 See for example Kirsch I (1999)

the way a treatment is presented, the attitude of the clinician themselves, or even advertising can influence a patient's expectations of a treatment.

Zubieta et al showed how a placebo treatment administered with an expectation of pain relief can have an actual analgesic effect on patients.[81] They used PET scans to show that a placebo treatment, along with the expectation of pain relief, led to the release of endogenous opioid neurotransmitters (endorphins) in the brain. This change in the brain was accompanied by a reduction in the intensity of the pain reported by the trial participants.

An expectancy model of the placebo effect could possibly explain many of the oddities observed in placebo research, such as the way large placebo pills are more effective than small ones,[82] or the way coloured placebo pills produce different effects in countries where those colours have different meanings. For example, blue placebo pills are more effective tranquilisers than red placebo pills everywhere except for Italy, where the colour blue is strongly associated with the Forza Azzurri, the Italian national football team.[83] Such cultural differences could easily affect the expectations a person has when taking a pill, but it is difficult to see how it could affect the conditioning that results from past experiences with medication.

There has been lively debate in the past as to which of these two mechanisms provides the better explanation for the placebo effect, with the likes of Voudoiris et al[84] barracking for classical conditioning, and Irving Kirsch,[85] among others, arguing that expectancy is the

81 Zubieta JK et al. (2005)
82 Buckalew LW, Ross S (1981)
83 Moerman D (2002) p49
84 Voudouris et al (1990)
85 Kirsch I (1991)

superior model. However, this dispute has largely been resolved,[86] and the prevailing opinion now is that classical conditioning and expectancy are not mutually exclusive, but that either or both can be involved in a given placebo response. Indeed, expectancy and conditioning are now thought to be only two of a number of culprits that could be responsible for a placebo response, with other mechanisms such as Daniel Moerman's meaning response[87] also possibly playing their part. What we know as the placebo effect looks like it could actually be a number of effects resulting from a number of psychological and neurobiological mechanisms. According to a 2010 review[88], the psychological mechanisms thought to contribute to the placebo effect include "expectation, conditioning, learning, memory, motivation, somatic focus, reward, anxiety reduction and meaning."[89]

Conclusion

I hope that this chapter has gone some way towards answering the basic questions of what a placebo treatment consists of, why we have reason to believe that placebo treatments can work in at least some cases, and how placebo treatments might work in such cases: what Colin Cheyne referred to as the "epistemic constraint" faced by those who are considering employing a placebo treatment.[90] Granted, there is lively and ongoing debate about the effectiveness of placebo treatments, but open debate is at the heart of science, and with it, medicine. Although it is possible that some spectacular study will come out in the future showing that placebo treatments are actually devoid of any therapeutic efficacy beyond making patients think they feel better, the same is also true of

86 For a good review see Stewart-Williams S and Podd J(2004)
87 Moerman D (2002)
88 Finniss DG et al (2010)
89 Ibid. p687
90 Cheyne (2005) p178

almost any active medication currently in use. I therefore conclude that, at present, if a clinician reads that a placebo treatment has been shown to have a therapeutic effect in patients with a certain condition, then she should be no more sceptical of this report than she would of a similar report involving an active medication.

In the following chapters, my focus turns to the ethical constraint faced by placebo treatments: if a clinician does believe that a deceptive placebo treatment presents a patient's best or only chance at achieving some therapeutic benefit, is it permissible for her to administer that deceptive placebo treatment? In order to assess this, I will first examine the information we have on placebo use in clinical practice around the world, and the ethical guidelines on placebo use that currently exist. Then I turn to a number of theoretical problems posed by deceptive placebo use in clinical practice.

References

Aronson J (1999) “Please, please me” *British Medical Journal* 318:716

Beecher, HK (1955) “The Powerful Placebo” *Journal of the American Medical Association* 159(17):1602-1606

Benedetti F (2009) *Placebo Effects: Understanding the Mechanisms in Health and Disease.* Oxford: Oxford University Press

Berkeley, G (1747) *Siris: A Chain of Philosophical Reflexions and Inquiries concerning the Virtues of Tar-Water and divers other Subjects connected together and arising one from another.* Dublin: W Innys

Buckalew LW, Ross S (1981) "Relationship of perceptual characteristics to efficacy of placebos" *Psychological Reports* 49(3): 955–61.

Cabot RC (1903) “The use of truth and falsehood in medicine: an experimental study” *American Medicine* 5: 344-349

Cheyne C (2005) “Exploiting placebo effects for therapeutic benefit” *Health Care Analysis* 13(3): 177-188

Cullen W (1772) *Clinical lectures 1772-3*, in *RCPE Manuscript Cullen* retrieved from http://www.jameslindlibrary.org/cullen-w-1772/

De Craen AJM et al (1999) “Placebo effect in the treatment of duodenal ulcer” *British Journal of Clinical Pharmacology* 48: 853-860

Diehl HS et al (1938) “Cold vaccines. An evaluation based on a controlled study” *Journal of the American Medical Association* 111: 1168-1173

Evans, D (2003) *Placebo: Mind Over Matter in Modern Medicine*, London: HarperCollins

Evans W, Hoyle C (1933) “The comparative value of drugs used in the continuous treatment of angina pectoris” *Quarterly Journal of Medicine* 26:311-38

Finniss DG et al (2010) “Biological, clinical, and ethical advances of placebo effects” *Lancet* 375: 686-95

Flint A (1863) “A contribution toward the natural history of articular rheumatism; consisting of a report of thirteen cases treated solely with palliative measures” A*merican Journal of the Medical Sciences* 46:17-36.

Ford (ed.) (1898) *The Writings of Thomas Jefferson* vol 10, New York: Federal Edition

Franklin, B (1785) *Report of Dr. Benjamin Franklin, and other commissioners, charged by the King of France, with the examination of the animal magnetism, as now practised at Paris.* London: J Johnston

Goebel MU et al. (2002) “Behavioral conditioning of immunosuppression is possible in humans” *Federation of American Societies for Experimental Biology Journal* 16:1869–73

Gold H (1946) “Cornell conferences on therapy: the use of placebos in therapy” *New York State Journal of Medicine* 46: 1718-27

Gold H, Kwit NT, Otto H (1937) “The xanthines (theobromine and aminophylline) in the treatment of cardiac pain” *Journal of the American Medical Association* 108: 2173-79

Gordon A (1788) *The practice of physick*. Unpublished manuscript retrieved from http://www.jameslindlibrary.org/gordon-a-1788

Graves, TC (1920) “Commentary on a case of hystero-epilepsy with delayed puberty: treated with testicular extract” *Lancet* 196(5075): 1134-1135

Haygarth J (1801) *Of the imagination, as a cause and as a cure of disorders of the body: Exemplified by fictitious tractors and epidemical convulsions.* Bath: R. Cruttwell

Ho et al (1988) “Reduction of post-operative swelling by a placebo effect” *Journal of Psychosomatic Research* 32:197–205

Honigfeld G (1964) “Non-specific factors in treatment: I. Review of placebo reactions and placebo reactors” *Diseases of the Nervous System* 25: 145-156

Houston, WR (1938) “The Doctor Himself as a Therapeutic Agent” *Annals of Internal*

Medicine 11(8): 1416-1425

Hróbjartsson A, Gøtzsche P (2001) “Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment” *New England Journal of Medicine* 344(21): 1594-1603

Hróbjartsson A, Gøtzsche P (2007) “Powerful Spin in the Conclusion of Wampold et al.’s Re-Analysis of Placebo Versus No-Treatment Trials Despite Similar Results as in Original Review”, *Journal of Clinical Psychology* 63(4): 373-377

Humphery N (2002) *The Mind Made Flesh*, Oxford: Oxford University Press

Hunter J (1788) *A treatise on the venereal disease.* London: G Nicol and J Johnson

Jefferson et al (2009), “Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis” *British Medical Journal* 339: b5106

Kaptchuk T (1998) “Powerful placebo: the dark side of the randomised controlled trial” *Lancet* 351: 1722–25

Kerr et al. (2008) “William Cullen and a missing mind-body link in the early history of placebos” *Journal of the Royal Society of Medicine* 101(2): 89-92

Kienle, G and Kiene, H (1997) “The Powerful Placebo Effect: Fact or Fiction?” *Journal of Clinical Epidemiology,* 50(12): 1311-1318

Kirsch, I. (1991). “The placebo effect as a conditioned response: Failures of the 'litmus test.' *Behavioral and Brain Sciences* 14: 200-204.

Kirsch I (ed) (1999) *How Expectancies Shape Experience.* Washington DC: American Psychological Association.

Kirsch et al (2008), “Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration” *PLoS Medicine* 5(2):e45.

Lidstone SC et al. (2010) “Effects of expectation on placebo-induced dopamine release in Parkinson disease” *Archives of General Psychiatry* 67(8): 857-865

Moerman D (2002) *Medicine, Meaning, and the Placebo Effect* Cambridge: Cambridge University Press

Montgomery GH, Kirsch I (1997) “Classical conditioning and the placebo effect” *Pain* 72: 107-113

Moseley J. Bruce et al (2002), “A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee” *New England Journal of Medicine* 347: 81-88

Prendergast BJ et al (2002) “Photorefractoriness of immune function in male Siberian hamsters (Phodopus sungorus)” *Journal of Neuroendocrinology* 14(4): 318-29

Preston RA et al (2000) “Placebo-Associated Blood Pressure Response and Adverse Effects in the Treatment of Hypertension: Observations From a Department of Veterans Affairs Cooperative Study” *Archives of Internal Medicine* 160(10): 1449-1454

Quincy, J (1811) *Quincy's Lexicon-Medicum: A New Medical Dictionary.* Revised by R Hooper. London

Shapiro AK (1968) "Semantics of the placebo". *Psychiatric Quarterly,* 42(4): 653-695

Shapiro AK, Shapiro E (1997), *The Powerful Placebo*, Baltimore: Johns Hopkins University Press

Stewart-Williams S and Podd J (2004) “The Placebo Effect: Dissolving the Expectancy Versus Conditioning Debate” *Psychological Bulletin* 130(2):324-40.

Trimmer et al (2013) “Understanding the placebo effect from an evolutionary perspective” *Evolution and Human Behavior* 34(1): 8-15

Trotter T (1792) *Observations on the scurvy*. London: Longman

Voudouris NJ et al (1990) “The role of conditioning and verbal expectancy in the placebo response” *Pain* 43: 121-128

Wampold BE et al (2005) “The Placebo Is Powerful: Estimating Placebo Effects in Medicine and Psychotherapy From Randomized Clinical Trials” J*ournal of Clinical*

Psychology 61(7): 835-854

Wampold BE et al. (2007) “The Story of Placebo Effects in Medicine: Evidence in Context” *Journal of Clinical Psychology* 63(4): 379-390

Webster Marketon JI and Glaser R “Stress Hormones and Immune Function” *Cell Immunology* 252: 16-26

Zubieta JK et al. (2005) "Placebo Effects Mediated by Endogenous Opioid Activity on μ-Opioid Receptors" *Journal of Neuroscience* 25(34): 7754-7762

3. The current use and regulation of placebos in clinical practice

Nothing works better

– Universal Placebos (www.placebo.com.au) advertising slogan

Introduction

Before I begin my own ethical analysis of placebo use in clinical practice, I will briefly survey the existing ethical guidelines that govern clinical placebo use around the world. In doing so, I identify three salient ethical issues that are raised in the two national policies – that of the US and the UK – that that prohibit deceptive placebo use. These are that deceptive placebos are a violation of patient autonomy, that deceptive placebos are inconsistent with informed consent or shared decision-making, and that deceptive placebos are a violation of patient trust. I will examine each of these issues at length in the subsequent chapters.

In this chapter I will also undertake a brief review of the literature concerning the prevalence of, and attitudes to, placebo use in clinical practice around the world. From this review I conclude that deceptive placebos are used extensively in hospitals and clinics around the world, and thus that the questions I examine in this essay are far from purely theoretical.

Ethical guidelines on clinical placebo use around the world

The World Medical Association's Declaration of Helsinki contains strict guidelines for the use of placebos in clinical trials, but the WMA has no specific advice regarding the use of placebos in clinical practice. Some national medical associations have announced guidelines on clinical placebo use, but others are silent on this topic, and the ethical advice offered to doctors and nurses regarding the use of placebos in clinical practice varies from country to country.

Placebo prohibition: The United States and the United Kingdom

In the US, the American Medical Association adopted an ethics policy in 2006 that categorically prohibits the deceptive use of placebos in clinical practice:

> A placebo is a substance provided to a patient that the physician believes has no specific pharmacological effect upon the condition being treated. In the clinical setting, the use of a placebo without the patient's knowledge may undermine trust, compromise the patient-physician relationship, and result in medical harm to the patient.
>
> Physicians may use placebos for diagnosis or treatment only if the patient is informed of and agrees to its use. A placebo may still be effective if the patient knows it will be used but cannot identify it and does not know the precise timing of its use. A physician should enlist the patient's cooperation by explaining that a better understanding of the medical condition could be achieved by evaluating the effects of different medications, including the placebo. The physician need neither identify the placebo nor seek specific consent before its administration. In this way, the physician respects the patient's autonomy and fosters a trusting relationship, while the patient still may benefit from the placebo effect.
>
> A placebo must not be given merely to mollify a difficult patient, because doing so serves the

> convenience of the physician more than it promotes the patient's welfare. Physicians can avoid using a placebo, yet produce a placebo-like effect through the skillful use of reassurance and encouragement. In this way, the physician builds respect and trust, promotes the patient-physician relationship, and improves health outcomes.[1]

This policy was based on an earlier report by the organisation, which argued that deceptive placebo use "conflicts with contemporary notions of patient autonomy and the practice of shared decision making"[2] and that clinicians who use deceptive placebos risk undermining their patients' trust.

Medical authorities in the UK have also recently adopted an official position on the use of placebos in clinical practice. This came about due to a 2010 government report on homeopathy, in which the House of Commons Science and Technology Committee declared that "homeopathy is a placebo treatment"[3] and should therefore be subject to the NHS policy on placebos. This recommendation turned out to be the cause of some embarrassment, because the NHS had to tell the committee that it was currently impossible to bring homeopathic treatments under the umbrella of their placebo policy, as they didn't have one. "The Government should have a policy on prescribing placebos"[4] thundered the press release accompanying the committee's report. The British Medical Association was happy to suggest one, and the 2012 edition of the BMA's ethics and law handbook *Medical Ethics Today* contains guidelines on placebos that closely match those of their American counterparts:

> There are fundamental problems associated with doctors prescribing placebos within a clinical setting and, in the BMA's view, the unacknowledged use of placebos for patients with capacity is

1 AMA (2014) Opinion 8.083 - Placebo Use in Clinical Practice
2 Bostick et al p59
3 House of Commons Science and Technology Committee (2010) (2) p32
4 House of Commons Science and Technology Committee (2010) (1)

> unethical. Prescribing and administering a placebo must entail some degree of patient deception because to maximise the placebo effect, a patient needs to believe that the 'dummy' treatment administered is real. Deceiving a patient, even where the doctor is acting in his or her best interests, obviously undermines trust and risks damaging the doctor–patient relationship. Depriving a patient of the opportunity to exercise informed choice does not respect patient autonomy and runs contrary to the concordant model of prescribing and the principles of shared decision making.[5]

Comparing the two policies, I would argue that the points they share constitute the three main ethical claims about the use of deceptive placebo treatments in clinical practice:

- Deceptive placebo treatments are contrary to respect for patient autonomy.
- Deceptive placebo treatments do not meet the standards of informed consent.
- Deceptive placebo treatments undermine patient trust, putting the doctor-patient relationship at risk.

In the following chapters I will examine each of these claims in detail. In the next chapter, I will argue that deceptive placebo treatments can in fact be consistent with a Kantian concept of respect for rational agency. In chapter five, I will argue that deceptive placebo treatments can indeed meet the standards of informed consent, at least to the same level that active treatments regularly do. And in chapter six, I will argue that deceptive placebo treatments can be administered in a way that is concordant with patient trust, and s should not put the doctor-patient relationship at risk.

Another point raised in the AMA policy is that a placebo treatment may result in medical harm to the patient. This is an interesting claim, because it is difficult to see how an inert

5 BMA (2012) p565

treatment might cause somebody harm. However, there are two reasons that I can think of that a placebo treatment might be considered potentially harmful to a patient. Firstly, where a placebo treatment is administered in the place of an active treatment that could have provided a greater therapeutic benefit. In this case, the placebo doesn't actually harm the patient, but the treatment is not the best that can be offered, and so the patient is left worse off than they might be. However, this problem is not inherent to placebo treatments – whenever there is more than one treatment option for any given condition, there is always the chance that the one picked will be sub-optimal (assuming, that is, that the treatments can't be administered together). All a clinician can do is administer or recommend whichever treatment that they believe is best for the patient. This is true whether that treatment is active or a placebo, which is why my ideal placebo case is one where a clinician believes that a placebo treatment presents a patient's best, or only, chance at some therapeutic benefit. If this is the case, then the risk of this kind of “medical harm” – or sub-optimal benefit – presented by a placebo treatment is no greater than that presented by an active treatment in the same circumstances.

For example, say a physician is faced with the choice of administering a patient an inert placebo treatment, or an active medication that she believes will be no more effective than the placebo, and which carries the risk of some unpleasant side effects. The physician believes that the inert placebo will have the same positive therapeutic effect as the active medication, but without the negative side-effects. If she administers the placebo and she is wrong, then the patient will miss out on the therapeutic benefit the active treatment would've provided, and be worse off than he might be. But if the physician is correct that the placebo will have a therapeutic effect, and opts for the active medication anyway, the patient will likewise be worse off than he might be – he will receive the same benefit he

would have received from the inert placebo treatment, but with the added risk of side effects.

I would argue that any situation where a clinician has good reason to believe that a placebo treatment might be a better option than an active treatment must be similar to the above, in that there is some drawback to the active treatment that makes the placebo an attractive alternative. The possibility of this drawback must be weighed up against the possibility that the placebo treatment is ineffective, yes, but either of these outcomes has the potential to leave the patient worse off than if the other treatment had been selected. In general, then, where there is a decision to be made between a placebo treatment and an active one it is simply incorrect to claim that the placebo alone is at risk of leaving the patient worse off than they might be.

There is, however, another phenomenon that the risk of medical harm from a placebo treatment might be referring to – the nocebo effect. The nocebo effect might be referred to as the placebo effect's “evil twin”, and comes about when a patient's negative expectations manifest in negative outcomes, which can lead a treatment to be less effective that it should be, or perhaps even harmful to the patient. This effect can occur with active treatments or with completely inert ones. Just as expectations of benefit can cause an inert treatment to have a beneficial outcome, expectations of harm can cause an inert treatment to have a harmful outcome.

The possibility of nocebo effects is not something I consider for the majority of this thesis. For the sake of simplicity, I make the assumption that placebo treatments carry no risk of side-effects. This is because my aim here is to ascertain whether or not a deceptive

placebo treatment can be ethically justified in cases where it presents the best chance of achieving some therapeutic benefit. Granted, the possibility of a nocebo effect could have some bearing the question of whether or not a placebo treatment actually *does* present the best chance of achieving some therapeutic benefit in a given case, but this is a question for a doctor, not an ethicist. So for the next few chapters I will simply assume that a placebo treatment carries with it no risk of side-effects. However, in chapter seven I take a closer look at this assumption, and assess how the possibility of the nocebo effect could make a difference to the conclusions I draw in the previous chapters.

Another interesting point of the AMA policy is that, while it prohibits the use of deceptive placebos in clinical practice, and states that placebos may only be used with patient consent, the standard of consent the policy appeals to means that this does not rule out undisclosed placebo treatments altogether. The policy notes that placebo treatments "may still be effective if the patient knows it will be used but cannot identify it and does not know the precise timing of its use", and furthermore that a physician "need neither identify the placebo nor seek specific consent before its administration."[6] This implies that a doctor or nurse concealing the nature of a placebo treatment from a patient might not be considered to be engaging in deception, providing that the patient has previously consented to possibly being given a placebo treatment at some unspecified time.

This claim speaks to the second assumption I make for the majority of this thesis: that a placebo treatment requires deception in order to be effective. I also examine this assumption in chapter seven, where I assess the implications that the possibility of effective open-label, or overt, placebo treatments has for the conclusions I draw in the intervening chapters. However, the AMA policy suggests that total transparency may not

6 AMA (2014) Opinion 8.083 - Placebo Use in Clinical Practice

be the only way to avoid deception, and that covert placebo treatments might be administered in a non-deceptive manner by informing patients about, and having them consent to, the possibility of receiving a covert placebo treatments in future. Whether or not this procedure could meet the standards of informed consent is a question I will examine in chapter five, but here I would like to examine the claim that this could be an effective way to promote a placebo response.

If a healthcare institution reveals to a patient that there is a chance they will be administered a covert placebo treatment at some unspecified time in the future, and asks them to consent to this possibility, this creates a situation very like that of a clinical trial, where participants are informed of the chance that they will be assigned to the control group and receive a covert placebo in place of the experimental treatment. By all accounts this consent procedure does not prevent the placebo effect from occurring in clinical trials, so it seems unlikely that it would do so in the context of clinical practice. While it might not rule out the possibility of a placebo effect, however, several studies have indicated that such a consent procedure might not be the most effective way to promote a clinical placebo response.

In a meta-analysis of studies of placebo analgesia carried out by Vase et al, it was found that the placebo effect reported in studies that specifically investigated placebo analgesia was significantly higher than that reported in the placebo arms of clinical trials of analgesics.[7] Vase et al's posited explanation for this discrepancy is that the participants in nearly all of the analgesic trials were informed of the possibility that they would be receiving a placebo, while the participants in nearly all the placebo studies were not. The clinical trial participants' knowledge that they might receive a placebo treatment, they

7 Vase L et al (2002)

argue, could serve to lower their expectations of pain relief, and therefore lessen the placebo effect. A study by Kirsch and Weixel also found that the magnitude of the placebo effect observed in participants given a deceptive placebo was significantly larger than that observed in participants given a placebo under double-blind conditions, where they were informed that a placebo was a possibility.[8] These researchers likewise cited lowered expectations as the likely culprit for the enfeebled placebo effect. Studies by Pollo et al[9] and Greers et al[10] have also reported that the magnitude of placebo effects is reduced by the knowledge that receiving a placebo treatment is a possibility.

If the cited studies are correct, then, the AMA's recommended procedure for employing covert placebo treatments is not the most effective way to promote therapeutic placebo effects. The bad news does not stop there, however. The effects of lowered patient expectation are not limited to placebo treatments – lowered expectations could also have a detrimental effect on the effectiveness of any active treatments administered to these patients.

It is generally thought that the same mechanisms that lie behind the placebo effect – expectancy, conditioning, and the like – are also responsible for part of the therapeutic effect of nearly every active treatment.[11] Benedetti refers to the effect these factors have on active treatments as “placebo-like effects”.[12] In fact, the idea that these placebo-like effects operate in active treatments in exactly the way they do in placebo treatments is built into the design of randomised controlled trials, where the efficacy of an intervention is measured by subtracting the outcome of the placebo group from that of the intervention

8 Kirsch I, Weixel LJ (1988)
9 Pollo A et al (2001)
10 Greers A et al (2006)
11 See for example Colloca et al (2004)
12 Benedetti (2009) p36

group – the assumption being that the magnitude of the change caused by these factors (as well as by bias and other statistical anomalies) is the same in both groups, so the difference between the two can be attributed to the specific activity of the treatment. The proportion of a treatment's effectiveness that can be attributed to the placebo effect will vary from treatment to treatment, however in some cases it is estimated to comprise up to 50% of the clinical effect of a drug.[13]

These placebo-like effects in active treatments could be expected to be sensitive to exactly the same factors as are the placebo effects of inert treatments. The upshot of this is that anything that lowers a patient's expectation of effective treatment, and therefore lowers the expected magnitude of the placebo effect, will hurt active treatments as well as placebos. If asking a patient to consent to a possible future covert placebo treatment can lower their expectations, then, the placebo-like effects that make up part of the therapeutic benefit of active treatments could be affected by this just as the effect of placebo treatments are.[14] There is some experimental evidence for this: a review of meta-analyses carried out by Sinyor et al concluded that antidepressants in clinical trials with no placebo arm had significantly higher efficacy than the ones in trials that included a placebo arm.[15] Similar results have been reported in trials of migraine medication[16] and schizophrenia.[17]

Asking a patient to consent to a possible future covert placebo treatment, then, presents an (arguably) non-deceptive and consent-friendly way to put placebo treatments on the table. However, by doing this you run the risk of reducing the effectiveness of not only those placebo treatments, but any active treatments that might be administered as well.

13 Kirsch I and Sapirstein G (1998)
14 Note however that this should not affect the outcome of clinical trials: if the placebo group and the treatment group are being affected equally, then the difference between the two should remain the same.
15 Sinyor M et al (2010)
16 Diener HC et al (1999)
17 Mallinckrodt CH et al (2010)

This disclosure has the potential to weaken any future course of treatment, all to enable a treatment option that, at best, can provide minor benefits. The risk of less favourable treatment outcomes far outweighs the potential benefit presented by the placebo treatments this act of consent enables, so this cannot be said to be in patients' best interests. I therefore disagree that this practice can be touted as a more ethically acceptable replacement for deceptive placebo treatments.

Placebo permission: Germany

The German medical association, the Bundesärztekammer (BÄK), took a markedly different position to that of the AMA and the BMA with the release of their 2010 report on placebo use in clinical practice.[18] After three years reviewing the literature on placebo treatments, the report's authors concluded that placebos can sometimes be more effective than standard drug treatments, that they can reduce unpleasant side effects, and that they can result in lower health costs. The BÄK now recommend that placebos be used in clinical practice, when the following conditions are met:

- The patient has a minor condition and expresses a desire for treatment.
- No other treatment proven to be effective is available.
- A placebo treatment seems likely to succeed.[19]

They also recommend that doctors receive training in the use of placebos, and that placebo use is added to the curricula of medical schools in the country. To top it off, the

18 Jütte R et al (2010)
19 Ibid. p1418

BÄK calls for internationally recognised guidelines governing the use of placebos in clincal practice.

So how exactly does the BÄK recommend that a doctor goes about prescribing a placebo? As far as I can tell (the report is in German) this is not something that is examined in any great detail. However, in the press conference accompanying the release of the report Robert Jütte, the lead author, suggested that doctors should avoid lying, instead telling their patients that they are prescribing “a medicine that is not typically administered, but that it might nevertheless help them”.[20] This raises the question as to whether a placebo treatment could be honestly described as “medicine” (although this could be a mistranslation), but the general advice is clear enough: describe the treatment in terms vague enough to avoid lying, but promising enough to promote the patients expectations of a positive outcome. This is close to the advice that Lichtenberg et al give on how a doctor should describe a placebo treatment to a patient, their suggestion being: “I would like to offer you a pill which I believe can help lessen your suffering. I do not know exactly how it works. I have other pills to offer whose mechanism is clearer, but I am not sure that they will work better for you, and they may also entail more serious side effects.”[21] Whether or not statements like these can meet the standards of informed consent is something I examine in chapter five.

It is interesting that the BÄK recommendations limit placebo use to minor conditions. What might the rationale for this be? Well, one might argue, a placebo treatment could at best be expected to provide a minor benefit, so minor conditions are the only ones suitable to be treated in such a way. However, this is a little strange. In a case where a placebo

20 Kupferschmidt, K (2011) p364
21 Lichtenberg et al (2004) p552

treatment is the best or only chance at achieving even a minor therapeutic benefit, surely this minor benefit is better than nothing? Recall the case that sparked Henry Beecher's interest in placebo treatments in the first place: a soldier with terrible injuries is being prepared for surgery and, lacking any morphine or other pain relief options, a nurse injects the soldier with a syringe full of saline. Surely this soldier could be in no way described as having a “minor condition” However, in a situation where a placebo treatment is the only available analgesic, it is definitely a better option than nothing at all.

Granted, the above is an extreme situation, one that is far removed from the comfortable GP's offices of modern-day Germany, and it might be argued that the BÄK guidelines were never intended to cover such eventualities. But it is not too difficult to imagine a similar scenario that might take place in even a well-resourced clinic in non-emergency conditions. For example, a patient might present with a non-minor ailment for which the only active treatment is a drug that has recently been shown to be no better than a placebo, and which can have some unpleasant side-effects. If the clinician believes that a placebo treatment has a good chance of providing some therapeutic benefit without the attendant risk of side-effects, why should the severity of the patient's condition be a reason for taking this option off the table?

I can think of two reasons why administering a placebo treatment to a patient with a serious condition might be thought to be more problematic than administering one to a patient with a minor condition. Firstly, a placebo treatment might prevent a patient from seeking out more effective means of treatment. Had he not been given a bottle of sugar pills by doctor A, a patient might have gone to see doctor B, who could've offered him a more effective treatment to help with his condition. Where this patient has a minor ailment

that is going to clear up by itself in a week or two, this is not going to do him much harm. But where this patient has a serious condition, this situation starts to to look a little more dire.

The idea that a placebo treatment could dissuade a patient from seeking out a more effective treatment, however, presupposes that more effective means of treatment are available for this particular condition. If this is the case, then a placebo would not be the best treatment available, and when doctor A decided that it was, she was lacking in knowledge, or made an error of judgement, or both. In any case, she has prescribed the wrong treatment. Presumably this is something could happen whether or not that wrong treatment is a placebo.

The second reason that prescribing a placebo treatment to a patient with a serious condition might be thought to be problematic is that, if this patient discovers that they have been given a placebo, they may lose faith in that doctor, or even in medicine more widely. I examine this argument more closely in chapter six, but for now it is enough to note that a loss of trust could indeed have serious consequences. If a patient no longer trusts their doctor they may not seek care when they should, they may not share as much information as they might otherwise, and they may not follow medical advice as closely as they otherwise would. For someone with a serious condition, then, the loss of trust may cause far more harm in the long run than the minor benefit they might initially gain from a placebo treatment. In cases where a patient has a minor condition that's likely to clear up in a week or two, it might be argued, this might not be such a pressing concern.

However, I would strongly disagree with assertion that gambling with the trust of patients is

more acceptable if those patients have minor conditions. While it may be true that patients with serious conditions are more immediately vulnerable to harm from the kind of practical consequences the loss of trust in a medical relationship can have, loss of trust is not a short-term affair. The fact that a patient is presenting with a minor condition now is no reason to suppose that they will not develop a more serious one in future. And if they do develop a more serious condition in future, any mistrust they've developed about the medical profession will likely still be with them.

Whether a placebo treatment might be a violation of patient trust, and why this might be, are questions I examine at greater length in chapter six. But it is enough here to note that, if a placebo does violate patient trust, this is a problem for administering placebos to patients with minor conditions as well as those with more serious ailments. I would therefore suggest that, in a situation where a doctor or nurse believes that a placebo is the best or only treatment available, the question of whether or not this treatment is permissible should not hinge on the severity of the patient's condition.

The three policies I have discussed above constitute all the advice I can find from medical associations around the world on placebo use in clinical practice. As far as I can tell, the Australian Medical Association, the Canadian Medical Association, and the New Zealand Medical Council have issued no guidelines on this topic. Regardless of this, the available evidence suggests that clinical use of placebos is fairly prevalent worldwide.

Placebo use in clinical practice around the world

While the ethical guidelines concerning placebo use vary from country to country, surveys of healthcare professionals carried out around the world indicate that the use of placebos in clinical practice is highly (and consistently) prevalent. In a 2009 survey of New Zealand doctors, around three-quarters reported having used a placebo treatment, with around half having used one in the previous year.[22] In a Danish study, 86% of clinicians reported using a placebo treatment at least once, with 48% reporting the use of placebos more than ten times within a year.[23] In Israel, 60% of a clinicians responding to a survey reported using placebos, with over than 30% doing so more than once a month.[24] A 2008 study of US internists and rheumatologists reported that nearly half of the 679 respondents used placebo treatments "at least two to three times a month",[25] while a survey of Swiss GPs and paediatricians indicated that 72% had used a placebo treatment, with 5% reporting employing placebo treatments "daily".[26] In Germany, 76% of GPs surveyed reported using a placebo treatment in the previous 12 months[27], while a 2013 study indicated that around three quarters of general practitioners in the UK use placebo treatments once a week.[28] Finally, 75% of nurses surveyed in an Iranian teaching hospital reported using a placebo treatment at least once within the last year.[29]

In 2010, Fässler et al conducted a systematic review of 22 studies of clinical placebo use amongst health care professionals[30] The heterogeneity of the data set meant that this was

22 Holt S, Gilbey A (2009)
23 Hróbjartsson A, Norup M (2003)
24 Nitzan U, Litchenberg P (2004)
25 Tilburt JC et al. (2008)
26 Fassler M et al. (2009)
27 Linde K, et al. (2014)
28 Howick J et al. (2013)
29 Baghcheghi N, Koohestani H (2011)
30 Fassler M et al. (2010)

far from easy – some studies focused on pure placebos, while others included impure placebos; to make things worse the definition of “placebo” was not standard across the data set, so a treatment that was regarded an “impure placebo” in one survey may not have been considered placebic in another. The timelines for the studies also varied, with some focusing only on the previous 12 months, and others asking whether placebos had ever been used. Nonetheless, the overall conclusions that Fassler et al drew were along the same lines as the individual studies I've included above: The proportion of physicians who reported having used a pure placebo treatment varied between 17% and 80%, and between 41% and 99% if both pure and impure placebos were addressed. On average, about 62% of physicians had reported using a placebo treatment of some kind. The proportion of hospital nurses reporting use of (mainly pure) placebos varied between 51% and 100%, with the average being about 76%.[31]

Fassler et al also looked at the reasons physicans gave for using placebo treatments, and the most common motive reported was the desire of patients to receive a prescription. Other reasons that came up often were to promote the placebo effect, to avoid conflicts with patients, as supplemental treatment or for non-specific symptoms and to avoid telling patients that treatment possibilities were exhausted.[32] While the researchers conclude that the use of impure placebos, like antibiotics for viral infections, and vitamins or analgesics for unknown indications, are much more frequent than the use of pure placebos like sugar pills and saline shots, as there is no agreement as to what constitutes an impure placebo it is difficult to provide any hard data on this point.

I therefore conclude that, based on the available data, doctors and nurses around the

31 Ibid. p6
32 Ibid. p5

world appear to be employing deceptive placebo treatments with some enthusiasm. How might patients respond to this information?

Patients' attitudes to placebo use

There have been a few surveys carried out to assess the attitudes of patients to placebo use in clinical practice. In an American survey, 853 members of the public who had seen a primary health care provider for a chronic health problem in the previous six months were asked about their attitudes towards placebo treatments.[33] Seventy six percent of respondents answered that it was acceptable for doctors to recommend a placebo treatment where they thought it would benefit and not harm the patient, while 50% considered it acceptable even if doctors were uncertain whether it would provide any benefit. Only 22% answered that it was never acceptable for doctors to recommend placebo treatments.[34] In a scenario where a patient presented with moderate stomach pain of unknown cause, 71% of respondents judged that it was acceptable for a doctor to prescribe a herbal treatment in order to promote the placebo effect. This number rose to 79% if the doctor revealed to the patient that she believed the treatment was a placebo. In scenario where a patient is requesting antibiotics for cold symptoms, 66% of respondents answered that it was acceptable for the doctor to prescribe a sugar pill instead. By contrast, 92% of respondents said it was unacceptable for the doctor to prescribe antibiotics as a placebo. In addition, 54% of respondents said it was unacceptable for a doctor to tell their patients that a placebo treatment is active medication.

33 Hull SC et al. (2013)
34 Ibid. p3

While these respondents sound fairly receptive to the idea of placebo treatments so far, the survey returned some interesting results when patients were asked about the revelation of a placebo treatment. Seventy five per cent of respondents said that, in the two scenarios mentioned above, if the placebo treatment worked and patients asked, the nature of the placebo should be revealed. However, 54% thought that this revelation would have a negative effect on the future doctor-patient relationship, even if the placebo worked. This is a very interesting result. I'd guess the 22% of respondents who answered that a placebo treatment was never acceptable make up a good portion of that 54%, but that still leaves a lot of respondents who apparently believe that placebo treatments are acceptable, and also that a successful placebo treatment should be unblinded on request, while also believing that this revelation of an acceptable placebo treatment would damage the doctor-patient relationship!

In a New Zealand study, 78% and 54% of respondents, respectively, answered that a deceptive placebo treatment was acceptable on at least rare occasions in each of two scenarios where the treatment was described as being for the benefit of the patient.[35] A further 13% and 30% responded that a deceptive placebo could be used in these scenarios “as a last option”, with the remaining 10% and 16% responding that a deceptive placebo was “definitely not appropriate” in these scenarios. By comparison, 48% of respondents thought a deceptive placebo was “definitely not appropriate” in a scenario where the intention was to calm a difficult patient and provide some relief for the doctor and her staff, and 72% thought a deceptive placebo was “definitely not appropriate” in a scenario where the intention is to stop a patient from seeking out another doctor. Overall, the authors concluded that “patients were accepting of placebo use in certain clinical

35 Chen GF, Johnson MH (2009)

situations”[36], particularly when the placebo is employed for the benefit of the patient.

A Swiss survey of patient attitudes indicated that patients were generally supportive of the prospect of placebo interventions, but 70% of respondents said that they would want to be explicitly informed of the non-specific nature of an intervention.[37] Interestingly however, 62% of those surveyed judged that a deceptive pure placebo scenario was acceptable where only indirect information was offered (“try this therapy, it has very few side-effects and can help with your nausea”), whereas only 42% judged that the same scenario was acceptable where the placebo treatment was accompanied by a direct lie. Once again, though, the question of unblinding the placebo treatment returned some worrying numbers. 54% of respondents said they would be disappointed to find out they had received a pure placebo, while 44% reported that they would be disappointed to find out that they had received an impure placebo (multivitamin or herbal pill).

Overall then, patients seem fairly receptive to the idea of placebo treatments, even though many think they (or other patients) would not like to learn that they had been given one. The most interesting point for me in all the previous facts and figures, however, is the great disparity between patients' and physicians' attitudes towards antibiotics as placebo treatments. Antibiotics are estimated to be amongst the most common placebo treatments employed, presumably because doctors think that antibiotics are what patients want. And while it may be the case that many patients do feel better walking out of the clinic with a prescription for antibiotics, if Hull et al's survey is anything to go by the average patient would find a bottle of sugar pills a far more ethically acceptable option.

36 Ibid. p35
37 Fässler M et al. (2011)

Conclusion

There is a definite lack of global ethical consensus regarding the use of placebos in clinical practice, and the ethical guidelines (or lack thereof) in individual countries vary widely. Despite this, placebos appear to be used enthusiastically in clinical practice all over the world. Moreover, the attitudes of physicians and patients seem to be fairly consistent across various countries. The ethical question I hope to answer in this thesis, then, is far from hypothetical. It appears that this is a decision being made every day by doctors and nurses around the world.

The two national policies prohibiting placebo use, that of the AMA in the US and the BMA in the UK, are similar enough that I have drawn from them what I think are the three main ethical concerns about deceptive placebo use in clinical practice. I will examine each of these in detail in the following chapters. In the next chapter, I will adopt a Kantian standpoint to analyse the question of whether deceptive placebo treatments are contrary to respect for autonomy. A Kantian ethical framework places an extremely high value on rational agency and is famously inflexible when it comes to acts of deception, so I believe that a Kantian justification of deceptive placebo treatments will be a significant vote in their favour. In chapter five, I will examine the question of whether deceptive placebo treatments could ever meet the standards of informed consent. And in chapter six, I will examine the claim that deceptive placebo treatments undermine patient trust, putting the doctor-patient relationship at risk.

References

American Medical Association (2014) *Code of Medical Ethics,* American Medical Association Press

Baghcheghi N, Koohestani H (2011) “Placebo use in clinical practice by nurses in an Iranian teaching hospital” *Nursing Ethics* 18(3): 364–373

Benedetti F (2009) *Placebo Effects: Understanding the Mechanisms in Health and Disease,* Oxford, OUP

Bostick N, Sade R, Levine M, Stewart DM (2008) “Placebo Use in Clinical Practice: Report of the American Medical Association Council on Ethical and Judicial Affairs,” *Journal of Clinical Ethics* 19(1): 58-61

British Medical Association (2012) *Medical Ethics Today: the BMA's Handbook of Ethics and Law (3rd ed.)* Chichester: Wiley-Blackwell/BMJ Books

Chen GF, Johnson MH (2009) “Patients’ attitudes to the use of placebos: results from a New Zealand survey” *New Zealand Medical Journal* 122(1296): 35-46

Colloca L, Lopiano L, Lanotte M, Benedetti F (2004) “Overt versus covert treatment for pain, anxiety and Parkinson's disease” *Lancet Neurology* 3: 679-684

Diener HC, Dowson AJ, Ferrari M, Nappi G, Tfelt-Hansen P (1999) “Unbalanced

randomization influences placebo response: scientific versus ethical issues around the use of placebo in migraine trials" *Cephalalgia* 19(8): 699-700.

Fässler M, Gnädinger M, Rosemann T, Biller-Andorno, N (2009) "Use of placebo interventions among Swiss primary care providers" *BMC Health Services Research* 9: 144

Fässler M, Meissner K, Schneider A, Linde K (2010) "Frequency and circumstances of placebo use in clinical practice - a systematic review of empirical studies" *BMC Medicine*. 8: 15

Fässler M, Gnädinger M, Rosemann T, Biller-Andorno, N (2011) "Placebo interventions in practice: a questionnaire survey on the attitudes of patients and physicians" *British Journal of General Practice* 61: 101-107

Geers AL, Helfer SG, Weiland PE, Kosbab K (2006) "Expectations and placebo response: a laboratory investigation into the role of somatic focus" *Journal of Behavioral Medicine* 29(2): 171-8.

Holt S, Gilbey A (2009) "Use of placebos by New Zealand doctors" *New Zealand Medical Journal* 122(1298): 84-86

House of Commons Science and Technology Committee (2010) Committee press release (22 February 2010)

House of Commons Science and Technology Committee (2010) *Evidence Check 2:*

Homeopathy London: The Stationary Office

Howick J, Bishop FL, Heneghan C, Wolstenholme J, Stevens S, Hobbs FDR, et al. (2013) “Placebo Use in the United Kingdom: Results from a National Survey of Primary Care Practitioners” *PLoS ONE* 8(3): e58247

Hróbjartsson A, Norup M (2003) “The use of placebo interventions in medical practice--a national questionnaire survey of Danish clinicians” *Evaluation and the Health Professions* 2003(26):153-65.

Hull SC et al. (2013) “Patients’ attitudes about the use of placebo treatments: telephone survey“ *British Medical Journal* 346:f3757

Jütte R, Hoppe JD, Scriba P (2010) “'Stellungnahme des Wissenschaftlichen Beirats der Bundesärztekammer 'Placebo in der Medizin'” *Deutsches Ärzteblatt* 107(Heft 28–29): 1417-1421

Kirsch I, Sapirstein G (1998) “Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication” *Prevention and Treatment* 1(2): 2a

Kirsch I, Weixel LJ (1988) “Double-blind versus deceptive administration of a placebo” *Behavioral Neuroscience* 102(2): 319-23.

Kupferschmidt, K (2011) “More placebo use promoted in Germay” *Canadian Medical Association Journal*, 183(10): E633–E634

Lichtenberg P, Heresco-Levy U, Nitzan U (2004) “The ethics of the placebo in clinical practice” *Journal of Medical Ethics* 30: 551–554

Linde K, et al. (2014) “The use of placebo and non-specific therapies and their relation to basic professional attitudes and the use of complementary therapies among German Physicians – a cross-sectional survey” *PLoS ONE* 9(4): e92938

Mallinckrodt CH, Zhang L, Prucka WR, Millen BA (2010) "Signal detection and placebo response in schizophrenia: parallels with depression" *Psychopharmacology Bulletin* 43(1): 53-72.

Nitzan U, Litchenberg P (2004) “Questionnaire survey on the use of placebo” *British Medical Journal* 2004(329): 944-46.

Pollo A, Amanzio M, Arslanian A, Casadio C, Maggi G, Benedetti F (2001) “Response expectancies in placebo analgesia and their clinical relevance” *Pain* 93(1): 77-84.

Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lanctôt KL (2010) "Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses" *Journal of Clinical Psychiatry* 71: 270–279

Tilburt JC et al. (2008) “Prescribing 'placebo treatments': results of national survey of US internists and rheumatologists” *British Medical Journal* 2008(337): a1938

Vase L, Riley JL, Price DD (2002) “A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia” *Pain* 99(3): 443-52

4. Kant, autonomy and placebos:

Advice from the Categorical Imperative

Truthfulness in statements which cannot be avoided is the formal duty of an individual to everyone, however great may be the disadvantage accruing to himself or another.

– Kant, “On a Supposed Right to Lie from Altruistic Motives”

Introduction

A deceptive placebo treatment inevitably involves deceiving a patient. However, this deception – at least in the ideal case, where of the available treatments a deceptive placebo offers the best, or only, chance of some therapeutic benefit – appears to be carried out in the patient's best interests. Is this deception, carried out in order to promote a therapeutic placebo effect, a violation of a patient's autonomy? This is the key question I hope to answer in this work, and in the present chapter I'd like to examine it in the light of the ethical thought of Immanuel Kant. Now, this might immediately seem to be a bad move – every undergraduate philosophy student is well aware of Kant's strict prohibition on lying; his insistence that truthfulness is a duty we must abide by even if a murderer should be inquiring about the whereabouts of a friend. But I do have my reasons for examining this topic from a Kantian perspective: when an engineer wants to test the integrity of a piece of hardware, she does not merely subject it to the sort of run-of-the-mill wear and tear it could be expected to endure every day, rather she runs it through the most extreme conditions that it could possibly encounter. If the item in question passes this “stress test”, the reasoning goes, it should have no problem enduring the far less exacting pressures of its everyday use. This chapter, then, is a “stress test” for deceptive placebos: a Kantian moral

framework, with its strict moral law and emphasis on the importance of the autonomous choices of individuals, presents the most exacting conditions I can think of under which to test any course of action that involves deception. So while it may be far easier to justify benevolent deception in general, and placebo deception in particular, under different ethical schemes, if I can find a space for deceptive placebos in a much more stringent Kantian ethical framework, I believe that would be a significant vote of confidence in their favour.

In this chapter, then, I will examine the main Kantian arguments against lying and deception, then see how each applies to the placebo case. The first half of the chapter will examine what Kant's Categorical Imperative might have to say about deceptive placebos, and in particular how the Categorical Imperative test as it is interpreted by Christine Korsgaard might apply to the placebo case. I argue that a maxim of deceptive placebo use, properly constructed, can pass this test.

In the second half of the chapter I will examine a competing account of the Categorical Imperative test provided by Barbara Herman, which emphasises the Kantian prohibition on interfering with the will of rational agents – even where this is carried out in order to save a life. While this does not look promising for a maxim of beneficent deception, an analysis of the effect of placebo deception on the will demonstrates why this a maxim of deceptive placebo use might not violate this rule. Moreover, this analysis highlights the features that seperate morally permissible cases of placebo deception from morally impermissible ones, and from cases of paternalistic deception more generally.

Lastly, I will discuss a promising defence of deceptive placebo use offered by Gold and

Lichtenberg, and investigate the ways that Kantian concerns regarding deception and manipulation of the will can be incorporated into this theory. The resulting defence of deceptive placebo use in a clinical context is applicable to a Kantian ethical framework, but I believe it is also much more robust than Gold and Lichtenberg's original defence. It identifies those features that are specific to the placebo case, rather than to deceptive paternalism more generally, and explains how these features mean that the kind of deception employed in the ideal placebo case, and only that kind of deception, might well be meet the exacting terms of Kantian ethical thought.

At the outset I think it would be helpful to distinguish between the concepts of providing false information, lying, and deception. In Thomas Carson's *Lying and Deception*[1], he lays out the distinction between the three (roughly) as follows: Lying and deception are intentional actions. Providing false information is not necessarily lying, because a false statement may be made unintentionally, or without the intention of making another person believe it were true. A false statement made in error or in jest is not a lie. Conversely, one can make a statement one believes to be false, with the intention of making someone believe it were true, but if this statement unexpectedly turns out to be truthful, then one has not lied. A lie, then, is a false statement, that the liar believes to be false, which is provided with the intention of making another person form the belief that it is true. A lie, however, may not actually deceive anyone – if a lie is not believed, no-one will form a false belief because of it. The act of intentionally causing someone to form a false belief, that one believes to be false, is deception. Of course, deception can be carried out without lying: one can make misleading statements which, while true, are intended to make another person form a false belief. For example, say I make the false statement “it's one o'clock”. If my watch is wrong, and I actually think that it is one o'clock, then this unintentionally false

1 Carson (2010) p3-4

statement cannot be called a lie or a deception, for even though it may cause someone to have the false belief that it is one o'clock, it was not my intention to deceive. On the other hand, if I know that it is not one o'clock, and my intention in saying “it's one o'clock” is to cause another person to have the false belief that it is one o'clock, then this statement is a lie. Whether or not it is a deception, however, depends on whether the other person is fooled, and actually forms the belief that it is one o'clock. Lastly, if the time is actually one o'clock, I might nonetheless use the true statement “It's one o'clock” with the intention of causing someone else to have a false belief that follows from this statement. For example, if I am working behind the bar at an establishment that closes at two o'clock, I might nonetheless use the true statement “it's one o'clock, gentlemen” with the intention of giving my patrons to not only the true belief that it is one o'clock, but also the false belief that the establishment closes at one. If I am successful, and the true statement “it's one o'clock” causes this false belief, then I have been deceptive without lying.

Lies and the Categorical Imperative

The first Kantian argument that establishes truth-telling as an absolute duty can be found in the *Groundwork for the Metaphysics of Morals,* and results from the application of the first formulation of Kant’s Categorical Imperative (CI): “act only so that you can at the same time will that the maxim of your action become a universal law”[2]. The Categorical Imperative, a rule which Kant believed all rational beings to be bound by, allows us to test the maxims of our actions to see if they can be “universalised”: consistently willed as universal laws of nature, or of human action. It is an entirely negative test: it can only tell us whether or not a certain maxim violates a duty, not if that maxim has any positive moral

2 Kant (2012) 421

worth. But this is enough for my purposes in this chapter – if I can find a deceptive placebo maxim that is in a patient's best interests, but does not violate our duties to them, then I think I can call it a day.

A maxim of action is a subjective principle upon which we act: it will usually look something like “I will perform act A for purpose P”. An example given by Kant is “I will make a false promise of repayment for the purpose of gaining some money”. How can you determine whether or not such a maxim can be willed as a universal law of nature? In *Creating the Kingdom of Ends* Christine Korsgaard explains how, for Kant, whether or not you *can* will something is not a question of individual ability or desire; the only test appropriate is whether we can will this without contradiction:

> Since the will is practical reason, and since everyone must arrive at the same conclusions in matters of personal duty, it cannot be the case that what you are able to will is a matter of personal taste, or relative to your individual desires. Rather, the question of what you can will is a question of what you can will without *contradiction.*[3]

Kant illustrates the Categorical Imperative test with four examples, showing how contradictions in the universalised maxims generate duties to ourselves and to others. Firstly, a man who feels that his life is not worth living considers whether suicide would be contrary to his duty:

> Now he asks whether the maxim of his action could become a universal law of nature. His maxim, however, is: For love of myself, I make it my principle to shorten my life when by a longer duration it threatens more evil than satisfaction. But it is questionable whether this principle of self-love could become a universal law of nature. One immediately sees a contradiction in a system of nature whose law would be to destroy life by the feeling whose special office is to impel the improvement of life. In this case it would not exist as nature; hence

3 Korsgaard (1996) p77

> that maxim cannot obtain as a law of nature, and thus it wholly contradicts the supreme principle of all duty.[4]

Secondly, a man ponders the permissibility of a false promise of repayment in order to borrow some money:

> the maxim of his action would be as follows: When I believe myself to be in need of money, I will borrow money and promise to repay it, although I know I shall never do so. Now this principle of self-love or of his own benefit may very well be compatible with his whole future welfare, but the question is whether it is right. He changes the pretension of self-love into a universal law and then puts the question: How would it be if my maxim became a universal law? He immediately sees that it could never hold as a universal law of nature and be consistent with itself; rather it must necessarily contradict itself. For the universality of a law which says that anyone who believes himself to be in need could promise what he pleased with the intention of not fulfilling it would make the promise itself and the end to be accomplished by it impossible; no one would believe what was promised to him but would only laugh at any such assertion as vain pretense.[5]

A third man wonders whether he can continue his life of leisure, rather than striving to develop his natural talents:

> Now, however, let him ask whether his maxim of neglecting his gifts, besides agreeing with his propensity to idle amusement, agrees also with what is called duty. He sees that a system of nature could indeed exist in accordance with such a law, even though man (like the inhabitants of the South Sea Islands) should let his talents rust and resolve to devote his life to idleness, indulgence, and propagation—in a word, to pleasure. But he cannot possibly will that this should become a universal law of nature or that it should be implanted in us by a natural instinct. For, as a rational being, he necessarily wills that all his faculties should be developed, inasmuch as they are given to him for all sorts of possible purposes.[6]

Lastly, a wealthy man asks whether he has a duty to assist those less fortunate than

4 Kant (2012) 422
5 Kant (2012) 423
6 Kant (2012) 423

himself:

> he asks, “What concern of mine is it? Let each one be as happy as heaven wills, or as he can make himself; I will not take anything from him or even envy him; but to his welfare or to his assistance in time of need I have no desire to contribute.” If such a way of thinking were a universal law of nature, certainly the human race could exist, and without doubt even better than in a state where everyone talks of sympathy and good will or even exerts himself occasionally to practice them while, on the other hand, he cheats when he can and betrays or otherwise violates the rights of man. Now although it is possible that a universal law of nature according to that maxim could exist, it is nevertheless impossible to will that such a principle should hold everywhere as a law of nature. For a will which resolved this would conflict with itself, since instances can often arise in which he would need the love and sympathy of others, and in which he would have robbed himself, by such a law of nature springing from his own will, of all hope of the aid he desires.[7]

According to Kant, there are two different kinds of contradiction found in the four examples, from which we derive two different kinds of duty. The first two examples illustrate what later came to be called a “contradiction in conception”: merely trying to will the maxim at the same time as you will that it be a universal law creates a contradiction, as these two parts of your will, the maxim and the law, now contradict each other in some way. The latter pair of examples illustrate what came to be known as a “contradiction in the will”: these maxims could conceivably be willed as universal laws without that contradicting the willing of the original maxim, however the willing of the universal law undermines some other feature of the will, or contradicts something that an agent must will. And again, the contradiction in the will is not a question of whether willing the maxim as universal law would clash with a certain agent's desires or inclinations, but whether willing this law would frustrate the things that any rational agent must will, if they are to will anything at all.

7 Kant (2012) 423

As for the two types of duty, Kant concludes that “[w]e easily see that the former maxim [contradiction in conception] conflicts with the stricter or narrower (imprescriptible) duty, the latter [contradiction in the will] with the broader (meritorious) duty”[8]. But while Kant may have been able to “easily see” the distinctions involved, contemporary Kant scholars disagree about exactly how these contradictions arise and the kind of duties generated by each. At this point, then, I'd like to examine a few accounts of the Categorical Imperative and explore the implications of each for deception in general, and placebo deception in particular.

Christine Korsgaard and the contradiction in conception

In her book *Creating the Kingdom of Ends*[9], Christine Korsgaard examines the kinds of contradiction that the Categorical Imperative test could be said to generate, concluding that there are three possible ways that willing the universalisation of an immoral maxim could be said to create a contradiction: the logical contradiction, the teleological contradiction, and the practical contradiction. She contends that, while arguments in favour of all three interpretations can be found in Kant's writings, the practical contradiction interpretation is the least problematic. I will discuss each of Korsgaard's interpretations of the contradiction in conception below.

The logical contradiction: According to this interpretation, the CI test yields a contradiction only if carrying out one's maxim in the world where it has been universalised would be logically or physically impossible. For example, it would be impossible to use a false

8 Kant (2012) 424
9 Korsgaard (1996) p78-102

promise of repayment to obtain money in a world where this maxim was a universal law, as the act of promising becomes impossible. No-one would believe a false promise of repayment in a world where a false promise was a universal way of obtaining money, and this widespread incredulity would mean that there would no longer even be a practice of promising in such a world. The logical contradiction interpretation is good at generating duties for what Korsgaard, following Rawls, calls “conventional” actions or practices, such as promises. These actions depend on certain social rules or conventions for their possibility and efficacy, and while these social conventions can survive the occasional violation of their rules, a universal law violating those rules would make the convention, and the actions that depend on it, impossible, which generates a logical contradiction. Another good example here is any maxim involving theft. Theft entails taking possession of someone else's property, and so depends for its possibility on the social conventions surrounding property and possession, at the same time as it violates the rules of those conventions. So an agent carrying out a maxim involving theft must will the existence of the social conventions of property at the same time as she wills an action that violates them, which creates a logical contradiction when it is universalised.

Applied to conventional actions, then, the CI in this interpretation is a good way of identifying when our actions are “free-riding” on the moral behaviour of others, when we are relying on ourselves being exceptional in some way. On the other hand, however, “natural” actions, like killing, do not depend on any conventions for their possibility, only on the laws of nature, and so they do not generate a logical contradiction when they become universal practice. Korsgaard concludes that “no amount or use of the action of killing is going to make it impossible”[10] (one might argue that totally universal killing would render no-one left alive to kill or carry out the killing, which would make the action impossible, but

10 Korsgaard (1996) p85

it is difficult to think of a purpose for the maxim which results in such carnage upon being universalised). Korsgaard argues that, as some obviously immoral maxims involving killing do not generate a logical contradiction when universalised, we must conclude that this interpretation is mistaken. For example, a maxim of "killing children that tend to cry more than the average, in order to get enough sleep"[11] passes the CI test without generating a logical contradiction. To be more specific, this maxim (and immoral acts of violence in general) does not lead to a contradiction in conception through this interpretation of the CI test. Korsgaard considers that such acts might be dealt with by the contradiction in the will test instead[12], but concludes that it is an inappropriate test for acts of violence. This is because the contradiction in the will test can only derive imperfect duties (or, according to Kant's later work, duties of virtue) which are not enforceable – and this is clearly not the case with murder.

The teleological contradiction: this interpretation of the Categorical Imperative is a little more esoteric. It involves a view of the world whereby the various faculties and attributes of human beings have natural ends or purposes, which is a more or less Aristotelian view of human nature. However, Kant's own explanation is interesting for its almost Darwinian language, coming as it did about 60 years before *On the Origin of Species* was published:

> each of these is a natural end, by which is understood that connection of a cause with an effect in which, although no understanding is ascribed to the cause, it is still thought by analogy with an intelligent cause, and so as if it produced men on purpose.[13]

A teleological contradiction is found when one's universalised maxim results in a law of

11 Korsgaard (1996) p82
12 see my later discussion of Barbara Herman for an analysis of a similar move
13 Kant (1996) 425

nature that requires employing one's faculties or attributes against against their natural end or purpose. In Kant's example, committing suicide out of self-love is contradictory because the natural purpose of the feeling of self-love is the preservation of life. In a world where a maxim of killing oneself out of self-love is universalised, there is a law of nature turning this feeling against its natural purpose, which means this system of nature contradicts itself.

However, even assuming we accept the teleological view of nature that this interpretation presupposes, a teleological contradiction doesn't necessarily mean that the agent's will contradicts itself, and the purpose of the CI test is to find contradictions in the will, not contradictions in a hypothetical system of nature. The only way a contradiction in the will could occur here is if the agent in question happened to will both the natural *telos* of her action or faculty as well as its intended end. And there is no reason why an agent has to will the natural purpose of her action: whereas the lying promiser must logically will the existence of the institution of promising, there is nothing that commits the suicidal agent to willing self-preservation as the end of self-love. Having said this, however, one might nonetheless think that self-preservation is a fairly reasonable end to assume a rational agent has, so another example might make this point more salient. In the *Metaphysics of Morals,* Kant also uses a teleological contradiction to argue that we have a duty to refrain from “defiling oneself by lust”:

> Just as love of life is destined by nature to preserve the person, so sexual love is destined by it to preserve the species... Lust is called unnatural if one is aroused to it not by a real object but by his imagining it, so that he himself creates one, contrapurposively; for in this way imagination brings forth a desire contrary to nature’s end, and indeed to an end even more important than that of love of life itself, since it aims at the preservation of the whole species and not only of the individual.

> That such an unnatural use (and so misuse) of one's sexual attribute is a violation of duty to oneself, and indeed one contrary to morality in its highest degree, occurs to everyone immediately.[14]

Of course, the contradiction Kant finds here does not extend only to acts of "self-defilement", but to any act motivated by feelings of sexual love which does not have procreation as its end. And it is tempting to argue that lust, being as it is a more-or-less constant urge to carry out an act that successfully results in procreation only occasionally, would be almost comically poorly suited to its role if indeed periodic acts of reproduction were the only purpose mother nature intended it to play in human life. But this, however, is beside the point. Even if a much better argument could be mounted for procreation being the one and only natural *telos* of sexual love, it would still be very difficult to argue that, for example, gay or infertile agents must rationally *will* that the end of all their feelings of lust is procreation. But without this act of will there is no contradiction in the agent's will, only in the "system of nature" being proposed by Kant.

Also in the *Metaphysics of Morals*, and perhaps more relevant to the topic of this chapter, Kant uses a teleological contradiction to argue against lying:

> communication of one's thoughts to someone through words that yet (intentionally) contain the contrary of what the speaker thinks on the subject is an end that is directly opposed to the natural purposiveness of the speaker's capacity to communicate his thoughts, and is thus a renunciation by the speaker of his personality, and such a speaker is a mere deceptive appearance of a man, not a man himself.[15]

Once again, this only creates a contradiction in the will if the agent in question happens to will the "natural purposiveness" of their capacity to communicate as well as the end of their

14 Kant (1996) 425
15 Kant (1996) 429

lie, otherwise the contradiction lies in the proposed system of nature rather than the agent's will. I find this example of Kant's interesting, not because it is significantly different to the previous two, but because it seems to me that a better Kantian case against lying could be made here. If you treat lying not as a natural action but as a conventional action, one that depends for its possibility and efficacy on the social conventions surrounding communication, then this would mean that maxims involving lying cause a contradiction when they violate, not the natural *telos* of human communication, but instead the social conventions surrounding communication and truthfulness. After all, a lie only works if it is believed, and I think it is largely due to social convention that people count on being able to believe what others say. For example, say I approach a complete stranger in the street and ask him the time. The reason I can count on him telling me the actual time is not just because he is voicing his thoughts in accordance with the natural purpose of his faculty of speech – if this was the case he might well reply that he hoped I wasn't just striking up conversation so I could try and convince him to sign up to donate money to some charity – but because of the social conventions of politeness and helpfulness that surround communication with strangers in public. (If you'd like to see an example of how the answer to a simple question like “what's the time?” can change depending on the setting and the social conventions, you might like to try a philosophical experiment that I used to enjoy when I was a boy: open your local phone directory and find someone named Mr Wolf. Call him up and ask him that question – then you'll hear a good example of someone voicing his thoughts in accordance with the natural purpose of his faculties of communication).

Approaching communication as a conventional action in this way allows us to argue that a maxim involving lying creates a *logical* contradiction, of the type seen earlier (or a practical contradiction, of the type I will detail next): a lying maxim depends for its possibility and

efficacy on social conventions of communication and truthfulness at the same time as it violates these conventions, and thus a universal law generated from such a maxim would create the same type of contradiction as was seen in the lying promise example. This construal also makes it clear why not all falsehoods are lies: the social conventions surrounding communication and truthfulness are different where, say, the speaker is telling a joke or acting in a play, so falsehoods in these situations do not violate convention.

The approach I propose here might also offer an answer to a “casuistical question” Kant poses in this section of the *Metaphysics*. He asks, “Can an untruth from mere politeness (e.g., the "your obedient servant" at the end of a letter) be considered a lie? No one is deceived by it.”[16] We might answer that “your obedient servant” is not considered a lie, but this is not just because no-one is deceived by it. Rather, the reason that no-one is deceived by this phrase is the same reason that it is not considered a lie: “your obedient servant” has a conventional meaning in this context that is separate from its literal meaning. In this case the social conventions of communication are only violated if the phrase is taken literally – and then it would be the reader who has violated these conventions, rather than the writer. If the conventions of communication and truthfulness that the writer is relying on to have her letter understood are not violated by her inclusion of “your obedient servant”, then there is no contradiction, and no lie, to be found here.

The practical contradiction: This is Korsgaard's preferred version of the CI test. According to this interpretation, a contradiction arises if the maxim in question would be unable to achieve the purpose for which it was originally willed if it was adopted as a universal law. So, in the case of a lying promise to obtain some money, the universalisation of this maxim

16 Kant (1996) 431

creates a contradiction not just because promising is now impossible (as under the logical contradiction interpretation), but because this impossibility means a false promise is no longer an effective means of obtaining money. By willing the universalisation of this maxim you are willing the frustration of your original end, which creates a contradiction in the will. The CI test in this interpretation identifies those maxims which rely for their effectiveness on the agent carrying them out being exceptional in some way, and which would fail if the maxim was used by everybody as what Korsgaard calls the “standard procedure” for obtaining this end.[17] Another way of looking at this is that the hypothetical imperative which you use to create your original maxim, in this case “if you desire some money, you should make a false promise of repayment to get it” is falsified by the adoption of the maxim as a universal law. At the same time as you rely on this hypothetical imperative to create a maxim of action, you will a state of affairs that renders it false.

The practical contradiction version of the CI test is as effective as the logical contradiction interpretation at identifying immoral conventional actions (here probably best described as those which depend for their *efficacy,* rather than their possibility, on the existence of social conventions), and it mostly avoids the problem that the logical contradiction interpretation has in identifying immoral “natural” actions, such as acts of violence, which rely only on the laws of nature. The practical contradiction interpretation is able to identify immoral natural acts because the purpose that an act of, say, killing is intended to achieve will almost invariably be frustrated in a world where a maxim of killing for this purpose is a universal law. For example, an act of murder might be carried out for material gain, or to improve one's station in life, but it would be impossible to secure either of these ends in a world where murder was the standard procedure for bringing them about. Neither one's life nor

17 Korsgaard claims that this interpretation of the CI test ties in with Kant's prohibition on using people as a “mere means”: if an action's effectiveness relies on most people do not doing it, this suggests you are using those people in this way.

one's possessions would be secure if these maxims were adopted as a universal law; the hypothetical imperatives (eg “if you want to possess material things, you should murder someone and take their possessions”) are falsified by the universalisation of the maxims they produce, and thus the agent's will contradicts itself.

However, the practical contradiction interpretation does still have a problem identifying some immoral natural actions – those that have as their purpose not some lasting effect, but only an immediate result. For example, a maxim of killing out of a desire for revenge is clearly immoral, but it does not seem to result in a practical contradiction when it is universalised. In such cases the fact that one's security is threatened by the universal adoption of this maxim has no bearing on its efficacy for achieving its original end, and thus the practical contradiction interpretation is unable to identify such maxims as immoral.

The practical contradiction, then, is Korsgaard's preferred interpretation of the contradiction in conception test, and it does indeed seem to be the least problematic of the three interpretations. But how do we apply this test to a maxim of deceptive placebo use?

Searching for a maxim

Recall if you will my ideal placebo case: a doctor or nurse has good reason to believe that a deceptive placebo treatment is a patient's best, or only, chance at some therapeutic benefit. How could Korsgaard's interpretation of the CI test help this doctor or nurse decide if this course of action violates her duty to her patient? The question is whether the maxim of placebo use could be be universalised without making it impossible to achieve the

original purpose of the maxim. But before we can answer this question, we first need to formulate a maxim which adequately describes the action to be carried out and its purpose, and this task is far from straightforward. A good place to start might be a general maxim of benevolent deception, for example:

M1: In order to benefit others, I will engage in deception whenever that is the best way to benefit the person being deceived

Such a general maxim seems certain to fail the CI test: a universalised maxim of benevolent deception would lead to widespread deceptive behaviour, and correspondingly widespread incredulity; any maxim that relies on other people believing your acts of deception would be certain to fail under such circumstances. And indeed, Kant wrote a whole essay arguing that it was immoral to "lie from altruistic motives"[18]. However, a general maxim of benevolent deception omits those features that make the placebo case such a singular puzzle. It is an inadequate description of the action being carried out and its purpose, and as such the possible world created by the universalisation of a general maxim of benevolent deception is a poor test for the morality of a maxim of placebo use.

If an excessively general maxim is problematic, so is an excessively specific one. Immoral maxims can be "tailored" to pass the CI test by the addition of superfluous details, which will allow the universalised maxim to avoid contradiction. For example, the un-universalisable (and therefore immoral) maxim "I will make a lying promise to gain some money" appears to pass the CI test if more details are included: "In order to gain some money I will make a lying promise to a man named Gussie Fink-Nottle on Wednesday, February 25". The universalisation of the more detailed maxim does not have the

18 Kant (1949) p346-50

widespread effects the more general maxim would; it doesn't seem likely that the adoption of the more specific maxim as a universal law would cause the practice of promising to be done away with, and thus at least one kind of contradiction is avoided.

However we might, following Korsgaard's favoured interpretation of the CI test, imagine the line of people at poor Gussie's front door on the morning of the 25th and ask whether this maxim, once universalised, could still be said to be an effective means of gaining some money. That is to say, the “tailored” maxim might make it through a logical contradiction in conception test, but there is still the question of whether the universalisation of the maxim has caused a practical contradiction. And it is true that universalising this maxim does cause a problem, but this problem is not a contradiction in the agent's will, of the kind that indicates an immoral maxim. Rather, this is a problem that often arises when considering maxims to do with the use of finite resources (in this case, one man's funds and patience) – a universalised maxim inevitably exhausts these. The practical problem that arises in this case is the same that results if we consider the otherwise-innocuous maxim “in order to gain some money I will make an honest promise of repayment to a man named Gussie Fink-Nottle on Wednesday, February 25” raised to the status of a universal law. Universalising this maxim, although it is not immoral, results in the same line of needy (although honest, in this case) people on the same doorstep competing for the same limited funds.

Allen Wood refers to immoral maxims that pass the CI test by the addition of specific details as “false positives”[19], and notes that the more details we include in a maxim the less consequences it has when universalised, and thus “the less likely it is, all other things being equal, that a contradiction will arise in trying to conceive or will the maxim as a

19 Wood (1999) p102-105

universal law of nature”[20]. But how are we to identify the level of generality at which we leave out the superfluous information that would create a “false positive”, but still retain a maxim that adequately describes our action? Onora O'Neill argues that we need to look for what she calls the “underlying principles or intentions by which we guide and control our more specific intentions”[21]. O'Neill describes a situation where she looks to make a visitor feel welcome by offering them some coffee. While this course of action could be described in any level of minute detail, she argues, right down to the number of times she stirs the coffee and the colour of the cup, the one intention underlying all of these actions was that of making a visitor feel welcome. This, she suggests, would be her maxim of action – the making and offering of coffee were subordinate to the maxim “make a visitor feel welcome”, and if, for example, she had been out of coffee, she would have carried out a different course of action towards this same end.

Unfortunately, once you begin looking at underlying principles and intentions it is very difficult to know where to stop. What was the intention underlying O'Neill's maxim of trying to make this person feel welcome? Perhaps she was hoping to foster a personal or professional relationship with them, and if for some reason they had been called away at the last minute, she would have pursued other means towards this end. But what was she trying to achieve by fostering this relationship? And so on down the rabbit hole. By analysing our intentions in this manner, we end up with a nested set of maxims, becoming more general and less descriptive the further we go. Consider the following series of maxims, which could be used to describe the same course of action in increasing levels of detail:

20 Wood (1999) p103
21 O'Neill (1989) p84

- I will act in order to further my own interests.
- I will engage in deception when it is to my advantage to do so.
- I will make a false promise when it is to my advantage to do so.
- I will make a false promise of restitution in order to gain whatever I need.
- I will make a false promise of repayment in order to gain some money.
- I will make a false promise of repayment to a friend in order to gain some money.
- I will make a false promise of repayment to Gussie Fink-Nottle in order to gain some money.
- I will make a false promise of repayment to Gussie Fink-Nottle on Wednesday, February 25, in order to gain some money.

Which is the true, or underlying maxim, the one we should test against the Categorical Imperative? In his maxim, Kant included the type of deception being carried out (false promising) and also a description of the end being sought (money), while he did not include, for example, the amount of money, or the relationship between the deceiver and the deceived, so we might conclude that the former details are relevant while the latter are superfluous. However, it is not immediately apparent why the nature of thing being sought should be a more relevant detail than our relationship with the person it is being sought from, or even that person's identity.

It could be argued that, in the above maxim, the nature of the thing being sought is on a higher level of intention than the identity of the person it is being sought from: it is because I need money that I went to see Gussie, knowing that he is wealthy enough to lend me some. If I was in need of something else – say, a car – and knew that Gussie did not possess one, this would lead me to seek out someone else. In Onora O'Neill's terminology,

the intention of getting some money was underlying the action of talking to Gussie. But this need not always be the case – it might be that the both the nature of the thing being sought and the identity of the person it is sought from are determined by the means I have chosen. For example, say I am in debt to some shady characters, and for whatever reason I decide that the best way to get out of this is through deception. I know my friend Bertrand has a car which I could deceitfully borrow and use to settle my debt, or else my friend Gussie has money which I could obtain with a false promise of repayment. If Bertrand is a naturally suspicious and vengeful sort, whereas Gussie is gullible and forgiving, then it is likely that I will end up making a false promise of repayment to Gussie, rather than making a false promise of return to Bertrand. It was not the fact that I needed money specifically that sent me to Gussie, rather it was his susceptibility to the means I had chosen that determined who it was I approached, which in turn determined what it was that I sought from him. So the level of intention at which these two details reside in the above maxim is definitely not fixed.

However, this is all quite beside the point. Even if it were apparent that, in the above maxim, the nature of thing being sought is at a higher level of intention than the identity of the person it is being sought from, we are still missing a criterion by which to judge that this higher level of intention is relevant to the CI test while the lower one is irrelevant. And while it is certainly possible to propose such criteria (and indeed I am about to), there isn't one that can be derived from the CI itself. It is this problem that leads Wood to conclude that “in addition to [the CI], we need a further specification of the moral laws themselves that [the CI] is commanding us not to violate.”[22] In other words, we need independent moral criteria to allow us to judge which possible features of a maxim are or are not morally relevant to the CI test.

22 Wood (1999) p105

Notwithstanding the lack of direction from the Categorical Imperative on this point, I think it is possible to find a level of detail for our placebo maxim that makes the relevant features of the case salient without going into such specificity that universalising the maxim could only have totally innocuous effects. Firstly, the nature of the relationship between the deceiver and the deceived. In Kant's false promising example no mention was made of this relationship: whether it was a family member, friend, bank manager or loan shark being deceived in this case seems to be irrelevant. And this seems entirely reasonable for a case of deception which takes place within the institution of promising, because the whole point of promising is that it facilitates trust between people no matter what their relationship is otherwise. From this we can perhaps take a lesson: for maxims dealing with conventional actions at least, the conventions surrounding that action can help guide us as to what details are and aren't morally relevant. In the placebo case, however, the deception does not take place within the institution of promising, it takes place in a medical context, and thus the relationship between deceiver and deceived is not necessarily an irrelevant detail. Indeed, I think that the fact that this maxim will invariably involve a health professional deceiving their patient is quite possibly a defining feature of the case, and one which definitely warrants inclusion in the CI test. The medical relationship is a special one, and the conventions and expectations surrounding the medical encounter are quite different to those of everyday life. Moreover, specifying that the relationship is between a health professional and patient does not vastly reduce the consequences of universalising this maxim: this relationship is sufficiently common that universalising a maxim that includes this detail would still have widespread effects, the inclusion of this detail is unlikely to allow an otherwise immoral maxim to slip past the CI test. Including this detail in our maxim, then, gives us:

M2: I will deceive my patient whenever there is good reason to believe that this is the best way to bring about some benefit for them.

Now we must turn our attention to the nature of the end being sought by our maxim. It is essential to Korsgaard's interpretation of the CI test that the end to be achieved is included in the maxim, as the efficacy of the maxim for achieving this end (once the maxim has been universalised) is what we are testing. However, once again there is the question of how much detail to include in the description of this end. In Kant's example, he thought fit to include specific details: money was the end, and a false promise of repayment was the means. I find the choice to include this level of detail a little odd, as it doesn't seem to make a difference to the outcome of the CI test whether it is money or some other benefit, for oneself or another, or just about any other end at all that one wishes to pursue through the means of a false promise. Indeed, in Kant's discussion of the maxim he asks whether a contradiction is created by “the universality of a law which says that anyone who believes himself to be in need could promise what he pleased with the intention of not fulfilling it”[23], suggesting that the original maxim of obtaining some money through a false promise of repayment is equivalent to one involving obtaining anything whatsoever with a false promise of any kind. It seems as if Kant's example is a case where, even for the practical contradiction test, the end to be achieved is actually irrelevant – trying to achieve any end at all through the means of a false promise is going to end up in a practical contradiction upon universalisation, when the institution of promising falls apart.[24]

23 Kant (2012) 423

24 Well, almost. There is one end that I can think of that could be consistently willed as the purpose of an act of false promising: that of the destruction of the institution of promising. Why someone might want to bring this about I don't know, but if this end is what we hope to achieve through our false promises, then it is entirely consistent with the universalisation of a maxim involving false promises, and the ensuing destruction of the institution of promising.

But back to the question of the end of the placebo maxim. While the ultimate purpose to be achieved by this act of deception is that of benefit to the patient, given that we are operating in a strictly medical context it is probably more accurate to say that we are aiming at some therapeutic benefit. However, this purpose underlies, to use O'Neill's term, the more immediate intention of engendering the placebo effect. Is there a justification for including this additional detail, the more specific or superficial purpose of the deception, in the maxim? Is this level of intention equivalent to counting the number of times we stir the coffee, or is it more on the level of making a visitor welcome? One possible way of answering this question is to look at the effect the inclusion or exclusion of this detail will have on the universalisation of our maxim. If we take both of possible resulting maxims to be equivalent, we see that universalising a maxim of placebo use would lead to a law of deceiving one's patient whenever this will achieve some therapeutic benefit. This universalisation is entirely unhelpful however, as I am not currently trying to determine whether it is permissible for a doctor or nurse to deceive a patient any time at all they think it will benefit that patient. I am only trying to determine whether or not it is permissible in the specific case that a placebo presents the best chance of achieving this end. The universalisation of the placebo-specific maxim would give a much better idea of whether or not this specific act is a violation of duty.

“Aha!” you might reply, “this is exactly the kind of argument we would expect from someone who was trying to tailor a maxim to get through the CI test – adding in a specific detail that restricts the scope of the universalisation, with no justification apart from that it's helpful to have the scope of the universalisation restricted!” And you might well have a point. But there are other reasons why the distinction between the end of therapeutic benefit in general, and the end of therapeutic benefit achieved by the placebo effect in

particular, is relevant here. The placebo effect is a peculiar means of achieving therapeutic benefit: one that only occurs in conjunction with deception, and one that cannot actually be willed by the patient themselves. It is entirely different, then, to the standard means of achieving a therapeutic benefit in a medical encounter, which is usually some means a patient can will, but is merely unable to achieve by themselves. For example, a patient might will that they be rendered unconscious and their heart be operated on in the hope of achieving some beneficial outcome. While this is obviously not something that the patient can do all by themselves, they still willed it to happen, in the sense that they knew that their decision, their consent, would cause the whole chain of events that would hopefully lead to this beneficial outcome.

To will something, at least according to Kant, is to regard yourself as a cause of that thing. The will, he says, is “a kind of causality of living things in so far as they are rational”[25]. But one simply cannot will that one's body carry out the classically conditioned or expectation-mediated response that results in the beneficial reaction to a placebo treatment. These bodily reactions are, of course, sub-rational, and I cannot rationally control whether or not they happen, but this is not problematic. I cannot control the digestive processes that occur after I have eaten an apple, but I nonetheless knew that something along those lines was going to happen when I chose to eat that apple, and I can rightly regard myself as the cause of these processes. Similarly, one could say that by swallowing a placebo pill I am the cause of the placebo effect, and this is correct, but I could not *regard myself* as the cause of that effect. This is because knowledge of this particular mechanism precludes it occurring. If I notice, for example, some lessening of pain following the treatment, then I could rightly regard myself as the cause of this, but I cannot regard myself as the cause of the placebo effect that led to it, for the simple reason that I can have no idea that this is

25 Kant (2012) 446

what is going on. Even in retrospect, should I find out that the pill I took was a placebo, I still cannot say that I willed the placebo effect, because at the time that I could have willed it, I could have had no idea that this is what my decision would cause.

This point will become more important when I discuss the Gold and Lichtenberg's defence of deceptive placebo use later in this chapter, but for now it is enough to note that the conflation of a maxim of deceptive placebo use with a more general one seeking therapeutic benefit through deception would mean that the CI test would involve not just a wider scope of application for the same type of deception, but the inclusion of a different type of deception. For example, the universalisation of a maxim of deception for therapeutic benefit would mean that a surgeon, if they judge that an operation would benefit a patient, and that the patient will only consent to this operation if deceived about, say, the likelihood of something going wrong, should deceive the patient in order to bring about this end. However in this case the surgeon is deceiving the patient into willing a means of benefit that they *could* will without deception, but would choose not to. When a doctor or nurse administers a deceptive placebo treatment, they are hoping to deceive the patient into means of benefit that simply cannot be willed by that patient, with or without deception. The distinction between these two types of deception is, I believe, morally relevant, and justifies the inclusion of the fact that it is the placebo effect specifically we are trying to achieve with our maxim. Our maxim then becomes:

M3: I will deceive my patient with a placebo whenever I have good reason to believe that the placebo effect is the best means of therapeutic benefit for them.

What about the nature of the deception being carried out? Is this a morally relevant detail

to include in our maxim? In Kant's example, the fact that the lie took the form of a false promise was definitely relevant – even though other forms of deception do not seem to fare much better in his work. And thus far in this thesis I have argued that there is a morally relevant distinction between deceiving a patient with a true, but deceptive statement (for example, “this is a pill to help with the pain”) and a lie (“this is morphine”). Can I justify including this detail, that the deception will take the form of a true statement, in a Kantian maxim?

An interesting subsidiary question is what exactly Kant would have had to say about true deceptive statements. I do not have nearly the room to do this question justice here, but I will note that, as with many other questions, there are passages in Kant to support just about any position you choose to take. Kant's famous proclamation that “[t]ruthfulness in statements which cannot be avoided is the formal duty of an individual to everyone, however great may be the disadvantage accruing to himself or another”[26], seems to allow for the possibility that true, but deceptive, statements are not a violation of this duty. However, in the *Metaphysics of Morals* he defines lying with a quote from Sallust: *aliud lingua promptum, aliud pectore inclusum gerere*[27] ("to have one thing shut up in the heart and another ready on the tongue"), which suggests that one's intention matters, rather than just the literal truth of one's words.

Lastly, we might look to an example from Kant's own life: Thomas de Quincey's includes in his essay “The Last Days of Immanuel Kant”[28] a tale from Kant's friend Ehregott Andreas Wasianski where Kant himself was forced to decide whether to tell a harmful truth or to resort to a less than wholly accurate version of events. For years Kant's footman Lampe, a

26 Kant (1949) p347
27 Sallust, *The War with Catiline* X, 5
28 Masson D (1968) p323-79

veteran of the Prussian army, had been a faithful and reliable servant to the great philosopher. However, in Kant's advancing old age Lampe became increasingly drunken, disrespectful, and worse – Wasianski believed that Lampe was stealing money from the philosopher. Finally, one morning in January 1802, Lampe committed some trespass so shameful that Kant could not even bring himself to repeat the story to his friend, and Lampe was summarily fired. Kant's dilemma arose when his former servant had the nerve to approach him and demand a reference of good character. Wasianski describes the struggle as follows:

> [Kant] was, however, a good deal embarrassed; his stern reverence for truth being, in this instance, armed against the first impulses of his kindness. Long and anxiously he sat, with the certificate lying before him, debating how he should fill up the blanks. I was present, but in such a matter I did not take the liberty of suggesting any advice. At last, he took his pen, and filled up the blank as follows: '...has served me long and faithfully,'--(for Kant was not aware that he had robbed him,)--'but did not display those particular qualifications which fitted him for waiting on an old and infirm man like myself.' [29]

Although Kant's statement was true, this was a case of having one thing in the heart and another on the tongue if ever I heard one!

Both lies and true deceptive statements are cases of deception – intentionally causing someone else to have a false belief – and it might be argued that this should be all the CI test needs to know. The arguments that are made for a distinction between the two tend to be on broadly utilitarian grounds: an outright lie causes more offence, and to have more impact on trust, that an a true deceptive statement. These, it might be argued, are not at all the sort of considerations that should be included in a formal test like the CI – after all, the CI test is only meant to find contradictions in the will. But this is to fundamentally

29 Masson D (1968) p351

misunderstand the Categorical Imperative, or at least the practical applications of it. While the Categorical Imperative may be an *a priori* principle, the application of it by testing our maxims as universal laws does require us to call on our experience to consider what the consequences of our actions might be, which includes taking into account such messy considerations as other people's feelings. Even the logical contradiction test requires more than purely formal knowledge to find contradictions of the will: there is no logical contradiction to be found in a universal practice of false promising, unless we consider that people tend to remember others' past behaviour and use it to try and predict what they will do in future – particularly, we might add, where this past behaviour has caused them some offence. If, on the other hand, we lived in a society where a widespread cultural or religious belief held that another's past offences should be instantly forgiven, and should under no circumstances pollute future dealings with that person, then a universal practice of making false promises when you are in need creates no contradiction at all. In such a society it would not be the case that, as Kant argues, universal false promising meant that "no one would believe what was promised to him, but would only laugh at any such assertion as vain pretense"[30] – the practice of forgiving past offences would mean there would be no laughter, and no disbelief, and false promises would continue to be an effective means of getting what you need. The difference in these two cases, where a maxim causes a logical contradiction on one hand and doesn't on the other, is not a difference of formal logic, or the nature of the will, but merely a judgement call, a difference in the way we expect people to react to our actions.

The fact, then, that people tend to react differently upon learning that they have been lied to than they do upon learning that they have been deceived by a true deceptive statement may very well mean that this is a relevant distinction to include in a Kantian maxim. And

30 Kant (2012) 423

furthermore, if we consider communication as a conventional action, the social conventions surrounding these two types of deception are vastly different. In courtrooms and parliaments across the world true deceptive statements are almost *de rigeur*, while actual lies are punishable transgressions. This is not to say that social conventions hold doctors and nurses to the same standards as lawyers or, god forbid, politicians, but if one of the things the CI test can tell us is whether a maxim is violating the same social conventions that it relies on to succeed, then at least we should seek to be specific about which social conventions it is that the maxim violates. This final inclusion, then, leaves us with the maxim I wish to test:

M4: I will use only true statements to deceive my patient with a placebo, whenever I have good reason to believe that the placebo effect is the best means of therapeutic benefit for them.

The universalisation of this statement gives us the law “Any healthcare professional will use only true statements to deceive their patients with a placebo, whenever they believe that the placebo effect is the best means of therapeutic benefit for them”. Does the universal adoption of this law make it impossible to use deceptive placebos to effect this therapeutic benefit? I don't think that it does. It is true that this maxim depends for its efficacy on social conventions of the medical relationship, particularly the trust in this relationship, so that the patient believes in the placebo treatment, but I don't think that the use of true deceptive statements in this scenario runs much risk of making patients disbelieve anything their doctors tell them in future. It may be that the more suspicious or easily-offended patients will ask some pointed questions if this situation should arise, and in such a case the truth will out quickly enough, but this is not a problem. The aim was

never to deceive those patients who *actually want to know what's in that pill*, because for such a patient it is simply not the case that a placebo is the best means of achieving some therapeutic benefit. For those patients who are a bit more open-minded perhaps, or simply don't care, a placebo will continue to be a viable treatment in a world where such a maxim is universalised. And thus, while widespread and unrestrained deception of patients would lead to a loss of trust in the medical profession, and therefore a practical contradiction, the restriction of this maxim to only those cases where a placebo is the best, or only treatment, and the deception involved to only that which can be achieved though true statements, this practical contradiction is avoided.

I thus conclude that, when considering a maxim of deceptive placebo use in a clinical context, the inclusion of the morally relevant features of placebo deception mean that this maxim passes the Categorical Imperative test, and thus is permissible under a Kantian moral framework. This is the case at least for Korsgaard's preferred account of the CI test – this maxim does not create a practical contradiction in conception. However, this is just one account of the CI test – a competing account might give a very different answer. In order to cover my bases, then, I will next assess how placebo deception fares under a very different account of Kant's theory – Barbara Herman's interpretation of the Categorical Imperative.

Barbara Herman – Kant and moral deliberation

Barbara Herman, in her book *The Practice of Moral Judgement*[31], takes a different approach to the Categorical Imperative test than the traditional one we have discussed so

31 Herman (1993) p115

far. Herman's understanding of Kantian moral theory has the potential to be an interesting prism through which to view therapeutic placebo use, and I think it warrants closer examination here. She abandons the traditional view of the CI procedure as a test of specific willed maxims from which we derive duties. Instead, she argues, the proper inputs for the CI test are not individual willed maxims of action, but instead "generic maxims". Generic maxims do not describe one specific course of action (for example, "I will do x for y reason") but rather describe a general policy of doing "x-type action for y-type reason".[32] And if these generic maxims are rejected by the CI, test the result is not a strict duty against carrying out certain courses of action, but rather a "deliberative presumption"[33] against all actions of that type. These deliberative presumptions indicate that the type of action in question is impermissible when carried out for the reason in question, but may be permissible if carried out for reasons of a different sort. For example, Herman describes how this version of the CI might apply to Kant's deceitful promising example:

> So, for example, when deceitful promises are rejected in the *Groundwork,* what the CI procedure shows is that reasons of self-interest cannot justify deceitful promises. Agents whose justifying reasons are different (they would make a deceitful promise in order to save a life), could be acting on maxims whose generic pattern is different and possibly permissible.[34]

Initially, then, Herman's approach seems like it might prove more flexible than the traditional Kantian deontology. And this account of the CI test, resulting in a deliberative presumption that might be rebutted by different justifying reasons, certainly sounds more promising for a maxim of placebo deception than one that stipulates a perfect duty not to lie. But there remain a few questions yet to be answered: what would a "generic maxim" of placebo use look like? And what does this "deliberative presumption" against deception mean for the permissibility of deceptive placebo use?

32 Herman (1993) p147
33 Herman (1993) p148
34 Herman (1993) p117

The problem of formulating a “generic maxim” to describe our case of deceptive placebo use potentially raises the same problem of maxim specificity I discussed earlier. Even though Herman's version of the CI does not generate duties, it will still return unacceptably rigorous judgements if these descriptions of action-types and reason-types are too general, and will allow clearly impermissible actions to pass if maxims are too specific. But Herman has an answer for this: the ideal generic maxim for any given type of action, she argues, is a broad description of the action-type in question being carried out for reasons of self-interest.[35] This is because self-interest, in the sense that an agent is only carrying out this type of action because the action or its effect is something the agent desires, is a more or less morally neutral reason for action. A generic maxim of self-interest, then, means that the results of the CI test for a type of action can serve to establish a kind of presumptive baseline: whether or not this action-type is a morally permissible means to a morally neutral end. If the CI test rejects a certain type of action carried out for reasons of self-interest, a *prima facie* judgement is established against this action-type, and any agent seeking to justify an action of this type must establish that the end they have in mind is morally superior to that of self-interest. (Presumably a generic maxim passing the CI test results in a similar presumption in favour of this action-type, which must then be evaluated if one's end is more morally questionable than that of self-interest, but this is not a possibility Herman explores).

So the generic maxim used to describe deceptive placebo use, then, is the same as that for deceitful promises that Herman outlined above: a maxim of deception for reasons of self-interest. This is rejected by the CI test, which means a deliberative presumption is established against this type of action. How do we move from here to considering the

35 Herman (1993) p148

placebo case specifically? Herman sketches out the process as follows:

> Prior to deliberation, the agent must *both* identify her proposed action as of a particular moral kind (this sets the deliberative presumption) *and* determine the nature of her interest in the action (or its end) that is to ground a possible rebuttal of presumption.[36]

This we can do: the moral kind of our proposed action is deception, which there is a deliberative presumption against. Our interest in this action is beneficent: we hope to carry it out in the best interests of another person. What next? Herman makes it clear that rebutting a deliberative presumption against actions of a certain kind involves arguing that the end one has in mind is somehow superior to that of self-interest. The presumption against certain action-types, she argues, "shifts the burden of proof to showing that a particular end, for whose sake we wish to take the action in question, has moral standing for us beyond its being an end in which we have an interest."[37] In the placebo case, this burden of proof is not difficult to shoulder: that our end of beneficence is a morally superior end than self-interest is well supported, both by intuition and by the traditional Kantian duty to help others in need. The question remains, however, whether the superiority of our end is sufficient to rebut the deliberative presumption established against deceptive maxims.

How can we judge whether our morally superior end, beneficence, means that the deliberative presumption against self-interested deception can be rebutted? What are the criteria by which we can judge this? The most immediate solution perhaps, would be to run the CI test again, using a general maxim of beneficent deception, rather than self-interested deception. Would this be sufficient? Unfortunately, it is not this simple: Herman later notes that among what she considers "false positives" (impermissible maxims that the CI fails to identify) are maxims of paternalistic deception.

36 Herman (1993) p150
37 Herman (1993) p149

Herman imagines a case where, in order to prevent a problem gambler from losing the money he has saved for his child's education, she makes a false promise of repayment to obtain the money (Herman believes that there is no morally relevant distinction between maxims of false promising and maxims involving other kinds of deception). She does not need the money, however, and rather than repaying it, she intends to hold on to it until the child can use it.[38] The fact that this maxim of action passes the CI test, she argues, should not lead us to conclude that it is permissible. This maxim passes the CI test, Herman argues, because it does not "free-ride": it does not depend for its possibility on others not acting similarly in similar situations. Although a maxim of paternalistic deceitful promising relies for its possibility on the institution of promising, a universal law of paternalistic deceitful promising would not render promising impossible in the way that is necessary to create a contradiction in conception. However, she argues that this should not lead us to conclude that paternalistic deception is permissible, pointing out that the feature that allows a maxim of paternalistic deceitful promising to "pass" the CI test – its restricted scope – is also found in "morally corrupt or trivial deceitful promises: a deceitful promise in support of racism; a deceitful promise in order to injure; even a deceitful promise in order to surprise you later."[39]

Herman's argument here is the same as my discussion of "tailoring" maxims above – where describing the same action in increasing levels of detail allows it to slip past the CI test – except that in this case the accusation of "tailoring" involves including a description of the *reason* for action, rather than details of the action itself. But this is a little strange: "tailoring", as I argued, involves adding superfluous details to the description of an action

38 Herman (1993) p141
39 Herman (1993) p142

in order to restrict the scope of a maxim. Here, Herman is arguing that describing the intent or end of the action *at all*, unless that end happens to be self-interest, is restricting the scope of the maxim and therefore tailoring it to pass the CI test. But while it is easy to see that the inclusion of the fact that an act of false promising took place on a Tuesday adds nothing morally relevant to the maxim, and serves *only* to restrict its scope, the fact that an act of false promising is carried out for beneficent reasons rather than self-interested ones does more than restrict the scope of this maxim. Indeed, if a maxim takes the form of "I will perform act x for reason y", then changing the reason for which one carries out an action doesn't just change the scope of that maxim, it changes the maxim. Whereas "make a false promise" and "make a false promise on a Tuesday" can describe the same act carried out for the same reason, "make a false promise for reasons of self-interest" and "make a false promise for beneficent reasons" cannot be considered the same maxim.

So, then, this is where we stand: a deliberative presumption against maxims involving deception has been established by the rejection of a generic maxim of self-interested deception from the CI test. We have a maxim of deception towards a different end: a beneficent deception, and we wish to see if this can rebut the deliberative presumption against maxims of this type. Herman argues that this involves showing that our end is morally superior to that of self-interest, which is not difficult. However, the permissibility of our new maxim cannot not be established by showing that it passes the CI test: a maxim of beneficent deception constitutes a "false positive". But, in the absence of the CI test, how are we to assess whether or not this end, which is generally thought to be morally superior to that of self-interest, makes a difference to the permissibility of the maxim? One way around the "false positive" problem might be to subject our new maxim to the second

type of test in the Categorical Imperative: the contradiction in the will (CW) test. The false positive was generated by the contradiction in conception (CC) test, which shows us if there is a contradiction between willing the maxim and willing the universal law that results from it, or a contradiction in the universal law itself. The contradiction in the will test is a further test, which shows us if there is a contradiction between willing the universal law resulting from a maxim and something a rational agent must necessarily will.

Herman demonstrates how a contradiction in the will test might work when discussing the problem the CI test appears to have with generating false positives for “natural” actions, such as killing. I discussed earlier how many intuitively wrong actions of this type seem to pass the CI test without generating a contradiction in conception. Korsgaard discounted the idea that we might solve this problem by instead looking for a contradiction in the will, on the grounds that the imperfect duties generated by this type of contradiction were not sufficient in strength or scope to describe our duty not to murder[40]. Herman, on the other hand, has done away with the idea of the CI test as a source of imperfect and perfect duties, and thinks that the contradiction in the will test might be a way around the problem of natural actions.

Herman takes us through an example of the contradiction of the will test, testing a generic maxim of convenience killing: “to kill whenever that is necessary to get what I want”[41]. Rather alarmingly, this maxim passes the contradiction in conception test – by a traditional Kantian account, then, it appears there is no such thing as a perfect duty to refrain from killing people whenever it suits us. Firstly, it passes the logical contradiction interpretation of the CC test: if we imagine a world in which the maxim “to kill whenever that is necessary

40 Korsgaard (1996) p82
41 Herman (1993) p117

to get what I want" is adopted as a universal law, such a universal law is not self-contradictory. Indeed, Herman imagines that the possible world of this universal law would look much like Hobbes' state of nature. There is no contradiction inherent in the universal law.

So there is no logical contradiction in conception inherent in the maxim "to kill whenever that is necessary to get what I want". What about a practical contradiction? Does the adoption of this maxim as a universal law create a contradiction by frustrating the end of the original maxim? At the outset, it does not seem likely that it would be possible to will, without that will contradicting itself, that something like a Hobbesian state of nature come into existence in order to further our own interests. Hobbes himself famously described such a world as one in which where there would be

> no place for hard work, because there is no assurance that it will yield results; and consequently no cultivation of the earth, no navigation or use of materials that can be imported by sea, no construction of large buildings, no machines for moving things that require much force, no knowledge of the face of the earth, no account of time, no practical skills, no literature or scholarship, no society; and – worst of all – continual fear and danger of violent death[42]

So the universalisation of the maxim "to kill whenever that is necessary to get what I want" would mean willing a state of the world that would frustrate just about every interest that the average person might have – a practical contradiction. A problem arises, however, when we consider someone who is far from average – someone like Genghis Khan. While most rational agents would gladly agree that a Hobbesian state of the world would be woeful, and completely counter to every interest they have, a Genghis Khan-type rational agent might be able to argue that such a state of affairs would actually be conducive to his interests. As such it may be possible for an agent like this to will the adoption of the maxim

42 Hobbes (1996) p89

“to kill whenever that is necessary to get what I want” as a universal law, without creating a practical contradiction. Is there a way to argue against Genghis Khan's assertion that, so long as he has his Mongol horde, whose best interests are served by him staying alive, he can rationally will a world where he (along with everyone else) kills whenever that suits them?

Herman argues against this assertion by finding a contradiction of the will: a contradiction between the universal law of convenience killing and something that any rational agent *must* will. Even a strong man, she argues, could not consistently will a world in which no-one placed a value on their life, because even he could not guarantee that willing this would not result in his death. In the Hobbesian state of nature even the weakest is able to kill the strongest, so even Genghis Khan, when he wills that everybody kills when it is necessary to promote their own interests, could be willing the cause of his own death. To will the cause of one's own death, she argues, is a contradiction: it conflicts with something that all rational agents *must* will, if they are to will at all, which is their own continued existence:

> In the fictional world of the CW test, I will that others not regard my life as a reason to refrain from taking it. Given the Hobbesian condition, I cannot guarantee that I will avoid a contradiction in willing. For if I will anything at all, I must will the necessary conditions of continued agency (or of my continued existence). And, given my inability to guarantee avoidance of the Hobbesian condition or its consequences, I cannot guarantee that I will not also have willed the cause of my loss of life.[43]

And thus any rational agent, even Genghis Khan, could not will a world where the maxim “to kill whenever that is necessary to get what I want” is universalised without a contradiction of the will – and so killing out of self-interest is ruled out, and there is a

43 Herman (1993) p121

deliberative presumption established against actions of this type.

So in this case the contradiction in the will test can successfully compensate for the “false positive” a generic maxim of convenience killing returns from the contradiction in conception test, and the desired deliberative presumption against acts of killing can be established. However, there are a number of questions raised by this result. Foremost among these is the question of the nature of the deliberative presumption generated by this maxim failing contradiction in the will test. Is it equivalent to the presumption that is generated by a maxim failing the contradiction in conception test? Or is a CW-derived presumption different somehow: smaller in scope, perhaps, or easier to rebut than one originating from the CC test? This is not a question Herman offers an answer to.

By a traditional account of the CI test, of course, the duties generated by the CW test are very different to those generated by the CC test. The duty generated by the CW test could establish that a maxim of convenience killing is morally wrong, but only in the same sense that failing to help someone in need is wrong in Kant's example from the *Groundwork* – as a violation of an imperfect duty to another. However, several problems arise if the duty not to kill others is classified in this way, as an imperfect duty. Firstly, imperfect duties are not as strictly binding as perfect duties, they are supposed to leave an agent some leeway as to when and how they are fulfilled. And secondly, imperfect duties are generally thought to identify impermissible ends, not impermissible actions. So Kant's demonstration that we have an imperfect duty of beneficence requires us to adopt a general policy or end of helping others in need, although how we pursue this end, or whether or not we act in a beneficent manner on any individual occasion, is up to us. One would hope that the duty to refrain from convenience killing, however, is a little more strict and specific than this.

Merely adopting a general policy of not killing others for reasons of convenience, and more-or-less adhering to this policy, is not sufficient to the duty we have to others not to kill them for reasons of convenience.

While Herman does not specify whether or not it is the case that the deliberative presumptions established by the CC and CW tests differ in the way that the traditional perfect and imperfect duties do, the results of this process would appear to be problematic either way. The possibility that the deliberative presumptions resulting from the CW test *do* differ from those of the CC test in the traditional way would imply that the presumption against killing others is unacceptably weak, perhaps able to be overridden by the kind of reasons that might justify an isolated instance of nonbeneficence. On the other hand, however, if the presumptions resulting from the CC and CW test *are* on a par, this would mean that the presumption against killing is as strong as we'd like it to be. But this would also imply that the presumptions against not helping others, or against wasting our talents (Kant's other example of an imperfect duty), are much stronger than previously suspected – these presumptions would be on a par with those against acts of deception or suicide. Whereas under the traditional Kantian model, occasionally forgoing beneficence or taking a day off violin practice is acceptable, putting these kinds of acts in the same category as killing suggests that this is not the case. The third option, of course, is that the CW-derived presumption against killing is much stronger than the CW-derived presumption against nonbeneficence, for some reason that has not been discussed. While it is apparent that Herman believes this third option is the case, the justification for this difference is not contained in the CI procedure, nor does she argue for it independently.

While the nature of the deliberative presumptions obtained from the CC and CW tests

remains problematic, Herman argues that there is good reason to believe that killing maxims should be classified alongside beneficence and developing ones talents, as proper subjects of the CW test. The CW, she argues, is a test for the kind of maxim that violates the "conditions of our agency".[44] As we have seen, such considerations include those things a rational agent needs for their continued existence (like not being randomly killed), and those things they need to continue functioning as a rational agent (like help from others in the time of need). On the other hand, the CC test is a test for the kind of maxim that impinges on the "constitutive elements of willed action"[45] – maxims that interfere with the functioning of the will directly. This means, she argues, that the CW test is concerned with the conditions of specifically *human* agency, as the conditions of our agency are tied up with our specific needs and limitations as human beings. On the other hand, the CC test identifies maxims that cannot be willed due to creating contradictions "on grounds that would be the same for all rational creatures".[46] This distinction, she argues, indicates that there are significant parallels between Kant's nonbeneficence case and the case of self-interested violence, good reasons that these cases might be properly classified together. Both cases involve a human agent denying another something that they *must* themselves need as a human being: help from others in times of need, or not to be subject to violence without good reason. These are conditions that a human agent requires to continue functioning as a rational agent, and the CW test identifies the kinds of maxim that impinge on these conditions.

This explanation of the proper subjects of the two tests involved in the Categorical Imperative gives us the answer as to whether the CW test could possibly deal with the "false positive" the CC test returns for a maxim of beneficent deception maxim: no, it

44 Herman (1993) p122
45 Herman (1993) p126
46 Herman (1993) p122f

couldn't. Maxims of deception involve manipulation of the will itself, not interference with the conditions necessary for willing. At least according to Herman, these are considerations that would be the same for any rational agent, and thus are the proper fodder for the contradiction in conception test, not the contradiction in the will test. It should be noted, however, that I disagree Herman on this point. I argued earlier that the logical contradiction that Kant identifies in a universal law of false promising, a type of contradiction that Herman believes would apply to any rational being, is actually based on human psychology, and would not be the case for any rational being that didn't see past trespasses as a reason to disbelieve present promises. But that is rather beside the point: as maxims of beneficent deception involve manipulation of the will *per se*, rather than impinging on the necessary conditions of the will, these maxims are in the domain of the contradiction in conception test, and thus the contradiction in the will test cannot be used to rectify the "false positive" these maxims generate.

According to Herman, then, both the contradiction in conception and the contradiction in the will test are inappropriate to test our maxim of beneficent deception. The contradiction in conception test generates a false positive, while the contradiction in the will test is simply not intended to appraise maxims of this type. This brings us back to our earlier question: how are we to judge whether or not our end of beneficence, which we have good reason to believe is morally superior to that of self-interest, is sufficient to rebut the deliberative presumption established by the rejection of maxims of self-interested deception?

Herman attempts to answer this question by considering an extreme case of beneficent deception: would it be permissible, she wonders, to deceive in order to save a life?

Questions of this type, like the famous “murderer at the door” example, have long been a bugbear of Kantians. Again, Herman finds that the results of the traditional CI test here are insufficient to justify deception in this case. A maxim of deception to save a life might pass the CI test, but “the CI procedure will show that it is also all right to deceive to save slugs”[47]: another false positive. What is required instead is a way to weigh up the results that Herman's version of the CI test gives us for the two relevant generic maxims in this case: deliberative presumptions against acts of deception and in favour of acts of beneficence, respectively. The CI test here tells us that reasons of self-interest do not justify manipulating someone else's will. It also tells us that self-interest does not justify denying aid to someone in need. So we know that these two considerations both outweigh self-interest in the Kantian schema, but how do they compare to one another?

This question presents a problem, because a Kantian framework simply does not support this kind of calculation. As Herman says, the commitments that we get from the CI procedure are not 'scalar': “failing to save the life of a rational agent is not half or twice as bad as deception.”[48] Rather, the Kantian theory of value that Herman outlines accords the integrity of an individual's will an ultimate (or what Herman calls “nonrelative”) value. This means that manipulating another's will cannot be justified by appeal to the value of the end furthered by this manipulation. Even a very laudable end, such as saving a life, she argues, could not justify manipulating someone's will. “The mistake is not in the fact that the manipulator accords the integrity of the will lower value than her ends, but in the fact that the integrity of the will is on the scale at all”.[49]

So we cannot engage in even a small deception in order to save a life? This seems

47 Herman (1993) p152
48 Herman (1993) p155
49 Herman (1993) p156

strongly counterintuitive. But Herman does not think so. “It is not clear” she argues, “that there is good reason to find this conclusion objectionable”[50]. Firstly, she doesn't agree that there is such a thing as a “small deception”: to use such a term is to smuggle consequentialist concerns into the deontological chapel. While there may be deceptions that have larger or smaller consequences, or better or worse consequences, the integrity of an agent's will is not something that can be violated in greater or lesser degrees: the will has integrity or it doesn't.[51] And furthermore, she argues, the constraint on deception should not make us as uncomfortable as we might think: “We are used to thinking that we cannot take someone's house to save a life... [n]ow we have an argument that we may not “take” someone's will.”[52]

Reasons of beneficence, then, cannot justify deception. Even to save a life, deception is impermissible. While Herman might not find this conclusion objectionable, it is somewhat disappointing to find that her approach to the CI is as inflexible as the traditional account, at least on this point. However, Herman does offer an example of a way that deception in this situation might be justified, which might be useful when considering the placebo case. While the ultimate value of rational agency means that it can never be interfered with for beneficent reasons, Herman also argues that there “is no a priori reason to suppose that no act of deception can accord with necessary respect for rational agency.”[53] Manipulating someone's will could be compatible with respect for their agency, she argues, where the will is manipulated to “bring it into conformity with it's own defining practices”[54] – like a kind of volitive chiropracty. If someone is clearly acting on impermissible maxims, and deception could be used to put their will back on the right track, then this deception could

50 Ibid.
51 I will take a closer look at this idea in the next section.
52 Herman (1993) p156
53 Ibid.
54 Ibid.

be consistent with respect for their agency. The fact that this kind of deception does not violate the integrity of the will means that the deliberative presumption established against self-interested deception may not apply.

This concept of respectful deception, manipulating a misguided agent's will in order to straighten out the kinks that lead them to pursue impermissible maxims, gives us a possible answer to the famous “muderer at the door” example. In such a case, deception might permissibly be used – not to save a life, but rather to get the murderer to abandon his impermissible murdering maxim. This means that the obvious lie that might be employed in such a situation: “no, my friend is not hiding in my attic”, is still unacceptable, as the aim of this deception isn't to make the murderer abandon his impermissible maxim, but merely to send him elsewhere in his attempt to carry it out. However, a more creative lie, like “I think he's handing out presents at the children's hospital tonight”, if offered with the intent of manipulating the murderer into abandoning his intent to murder his victim, might be consistent with respect for his agency, and therefore permissible.

While Herman doesn't examine this point in nearly as much detail as I'd like, it is promising that she has identified a kind of beneficent deception that may be permissible under a Kantian framework. These are acts of deception that respect the integrity of an agent's will, and manipulate it not to turn it into a tool of someone else, but to return it to its defining practices. It may sound strange to refer to this kind of deception as beneficent, given that it is not motivated by beneficence in the usual sense of promoting the wellbeing or best interests of others, but it is nonetheless aiming at some improvement in another's life – to be exact, in their will.

These are the two lessons I'd like to take away from my examination of Herman's work, then. Firstly, that violating the integrity of an agent's will can never be justified in a Kantian framework, not even if the benefit that accrues to that agent or another is very great. And secondly, that not all acts of deception are necessarily violations of the integrity of an agent's will. Some acts of deception could possibly be justified under a Kantian framework, then, and in order to determine whether or not a given act of deception is justifiable it is necessary to examine the effect it has on an agent's will. But is it be possible to apply this idea to the placebo case? What is the effect of a placebo deception on an agent's will? In order to examine this, in the next section I will look a little more closely at deception and manipulation of the will, and the way this investigation might add to a particularly interesting argument in favour of deceptive placebo use.

Gold and Lichtenberg's defense of deceptive placebo use

Gold and Lichtenberg[55] argue that in the ideal placebo case, where a doctor or nurse believes that a deceptive placebo is the best or only treatment available, the kind of deception employed is significantly different to the usual cases of deception that we might come across. They illustrate this with an example, which I've adapted a bit here:

Say my flatmate has a job interview, which she needs to get to by 9am. She is very excited by the prospect of this new job, and believes it will make a worthwhile and positive change to her life. However, she thinks that in order to be on time for the interview she needs to leave at 8am, while I believe this would give her insufficient time to make the journey. I have previously tried to convince her that she needs to leave earlier, but she will not listen

55 Gold and Lichtenberg (2014) p219

to my advice on this point. In the morning her phone is out of battery, so she asks me what the time is. It's 7:30am, which is about when I think she should leave the house. If I answer “It's 8am”, I am both lying and being deceptive. But if I answer “It's time for you to get going”, then I'm not lying (I'm giving her information I believe to be true), but I am being deceptive. Although both what I tell her, and the belief that I intend to cause her to have (that she needs to leave the house now), are what I believe to be the truth, I know full well that this information will most likely also cause her to form a false belief (that it is 8am).

Gold and Lichtenberg argue that the type of deception I carry out in this example of is analogous to the kind of deception a doctor or nurse carries out when they administer a placebo treatment in the ideal case. They argue that this type of deception is substantially different from the “standard proscribed sort” of deception, in the following two ways:

- The primary intention is not to make the deceived form a false belief, this is just a side-effect of the deceptive statement.
- The deception is not motivated by self-interest, but by helping the deceived achieve their goal.[56]

The first point is a little weak. While my primary intention in telling my flatmate “it's time for you to go” is not to cause her to have the false belief “it's 8am”, it does seem fairly likely that this false belief is instrumental in her decision to leave the house. She has after all not given much weight in the past to my opinions on when she should leave the house – if she had, then this deception would not be necessary to get her to leave at the time I think she should. Thus, considered on it's own, the belief that I think it's time for her to leave is probably not sufficient to motivate her to the desired action, of leaving the house at the

56 Ibid. p221

correct time. The false belief “it's 8am”, on the other hand, definitely would motivate her to act in this way. But if this false belief is indeed a means of achieving my desired outcome, it cannot be called a side-effect, and this doctrine of double effect-style justification fails. Likewise in the placebo case, if it is the false belief that there is active medication in the treatment that causes the positive expectations that lead to a placebo response in the patient, then this false belief is a necessary means to the end of the deception, and cannot be termed a “side-effect” of placebo deception.

The second distinction that Gold and Lichtenberg point out, however, is more interesting. While the fact that this kind of deception is not motivated by self-interest is important, I disagree with the authors when they argue that this is the quality that makes this kind of deception “significantly and qualitatively different from the classic morally faulty form of deception.”[57] After all, the non-selfish nature of this kind of deception is common to all truly paternalistic acts of deception, and many such acts would be described as “morally faulty” regardless of their lack of selfishness. No, the defining feature of this kind of deception is not just that the deceiver selflessly aims to achieve something the deceiver believes would be beneficial to the deceived, but rather that the deceiver aims to achieve an end *the deceived has chosen for themselves*. This is what separates this kind of beneficent deception from other cases of paternalistic deception. Interestingly, although Gold and Lichtenberg have gone to all the trouble of thinking up an example where deception is carried out specifically with the intent of helping the deceived achieve an end of their own, this detail more or less falls by the wayside in their discussion of paternalism and autonomy, and their argument is the worse for it. Nonetheless, this is the point that I wish to examine in more detail here.

57 Ibid. p222

Why might it be important that the act of deception in question is carried out to help the deceived achieve a goal of their own, specifically? It is not merely because achieving this goal is in their best interests, or would make them feel good, otherwise this kind of deception is no different to any other paternalistic deception, ostensibly justified by the good consequences it brings about for the deceived. Instead, I think the reason this distinction is important is because of the effect this kind of deception has on the will of the deceived. While deception of this sort involves manipulation of the will, one could argue that manipulation carried out with the intention of helping the deceived achieve an end of their own is less invasive than the manipulation involved in other kinds of deception. This kind of deception interferes with the functioning of the will, yes, but it does so in the most innocuous way possible. Recall that Herman argued that there is no such thing as a small or large deception, only deceptions with small or large consequences. While this may well be the case, if we concentrate only on the consequences that deception has for an agent's will, I think we may find a basis to argue for what we might call smaller or larger kinds of deception, greater or lesser violations of the integrity of an agent's will. Broadly speaking, I think that there are four ways that deception, by causing another person to have a false belief, can manipulate their will. They are as follows:

1. A false belief can cause someone to adopt ends they otherwise would not have, or to discard ends that they currently do have. For example, Iago falsely led Othello to believe that Desdemona was being unfaithful to him, and this caused Othello to adopt the new end of avenging the adultery.

2. A false belief can cause someone to employ means that are ineffective at achieving their ends, or at least not as effective as they believe they will be. For example, Kant's case of

borrowing money through a false promise of repayment. The deceiver leads the victim to believe he is taking action towards their shared end of the deceiver's temporary possession of some money, when his action is actually ineffective towards this end, instead being effective towards the hidden end of the deceiver's permanent possession of some money.

3. A false belief can cause someone to employ means that are just as effective at achieving their ends, but somehow different to the means they would have employed otherwise. For example, a Porsche salesman might convince me that purchasing a Porsche is a more effective means than the purchase of a similarly-priced Mercedes towards my end of owning a status symbol, when in fact either purchase would suffice perfectly well.

4. Lastly, as in Gold and Lichtenberg's example above, it is possible that a false belief could cause someone to employ means that are more effective at achieving their ends than those they would otherwise employ.

All four kinds of deception are consistent with selfless as well as self-interested motives. For example, Iago might believe that Othello's relationship is not good for him, and that Othello ridding himself of Desdemona is in his best interests. Conversely, I could deceive someone into employing means that are more effective at achieving their own ends than the means they would employ otherwise, but only do so because I stand to profit from those new means. So it is not the fact that the fourth type of deception can be carried out selflessly that distinguishes it here. Rather, I think it might be argued that this type of deception can be distinguished by its effects on the will. These four examples all involve

manipulating an agent's will, but that does not mean that all manipulations of the will are equivalent. I would argue that, all other things being equal, these four types of deception are presented in descending order of severity: a manipulation that causes an agent to change their ends is more severe than one that leaves those ends intact, but makes it less likely that the agent will achieve them. A manipulation that changes neither an agent's ends nor their chances of achieving those ends is less severe still. And lastly, there is a manipulation that renders an agent's will more effective at achieving its ends – the least severe of all.

What is the basis for this ranking of the severity of manipulations of the will? Again, it is not simply that achieving one's ends is good for an agent, or leads to better consequences. Rather, on a very basic level, choosing and achieving ends is what the will is *for*. An agent's will has integrity, yes, but it also has a purpose, and we can judge the severity of a manipulation by looking at the effect it has on the ability of the will to function in accordance with that purpose. While a manipulation of the will that changes the efficacy of an agent's means does seem invasive, in those cases where this is changed for the better, the manipulated will is still functioning as it should, allowing an agent to achieve their freely willed ends. This is the point that I believe Gold and Lichtenberg picked up on, although they did not argue for it explicitly: an act of deception that leaves an agent's ends intact and furthers their ability to achieve those ends is simply not as severe as the kinds of deception we are more familiar with, where the manipulation of the will renders an agent's means ineffective and leaves their ends in tatters.

There is another justification possible here, though. In cases where deception causes an agent to not merely carry out more effective means to their end, but to actually *will* those

means, we might appeal to the same justification that Herman offered for deception in the murderer-at-the-door case: returning the will to its proper functioning. When Herman argued that deception might be employed to return a will to "its own defining principles", the principle that the will in question had strayed from was the Categorical Imperative. But practical reason is another defining principle of the will. And, at least according to Kant, practical reason requires that if one wills an end, one must also will the necessary means to that end. He argues this in the G*roundwork*:

> Whoever wills the end, wills (so far as reason has decisive influence on his actions) also the means that are indisputably necessary and in his power.[58]

So if deception causes an agent to will a necessary means towards their end rather than an ineffective one, or rather than no means at all, then this deception could be consistent with respect for rational agency, as argued by Herman. By identifying a flaw in the agent's will and manipulating that will to return it to its proper functioning, deception of this type does not violate the integrity of an agent's will but rather restores it. Such an act of deception might successfully rebut the deliberative presumption against deceptive acts, and therefore be permissible under a Kantian framework.

In Gold and Lichtenberg's example, then, if I judge that my flatmate leaving the house at 7:30am is a necessary means to her willed end of attending the job interview, and deception is required to cause her to will this means, then this manipulation could be framed as returning her will to its proper functioning, and might therefore be permissible. However, a few quirks arise in this treatment of the flatmate example. Firstly, Gold and Lichtenberg argue that a defining feature of this kind of deception is that the intention is not to deceive, but to help the deceived achieve their goal. But if we are to apply this Kantian

58 Kant (2012) 417

treatment to this example, this is no longer the case. The deceived achieving her goal is no longer the primary intention, this is just an effect of restoring her will to its proper functioning. Much like in the murderer-at-the-door example, the beneficial consequences for the deceived or others cannot justify the deception: the restoration of the will must be our intended outcome.

Another difficulty arises if we consider exactly what it is that I am trying to deceive my flatmate into doing in this example. After all, it is not enough that my deception causes her to merely pursue the means necessary to achieving her end, she must *will* those means if her will is to be restored to its proper functioning. If I judge that “leave the house at 7:30” is the necessary means that she should will towards her end, I might be able to deceive her into willing “leave the house” at 7:30, which is an effective course of action towards her end, but this is not the same thing as actually willing the effective means, “leave the house at 7:30”. In fact, due to the deception carried out in this example, it would appear that I haven’t actually changed what my flatmate wills at all: her willed means is still “leave the house at 8am”, which is an ineffective means to her end. All the act of deception has done is to cause her to have the false belief that “it is 8am”, a belief which lead her to act in a way that deviates from her ineffective willed means in a manner that is effective towards her end. In order to cause her to actually will the correct means would require a different act of deception, for example telling her that there is a tram drivers' strike starting at 8am. This is an act of deception much more extensive than that proposed by Gold and Lichtenberg, but if it causes my flatmate to actually will the necessary means towards her end, then perhaps it could be more easily justified under a Kantian framework than a true deceptive statement that leads her to merely carry out, rather than will, those means.

Even now, however, I don't believe that we have identified those features that make placebo deception significantly different from other acts of paternalistic deception in a medical relationship. After all, a defining feature of the medical encounter is that the patient has a definite end, that of getting, or at least feeling, better, and that they are seeking a healthcare professional's assistance with achieving that end. With this end in mind, then, the above account of deception in order to cause an agent to will, or at least to carry out, a more effective means to her end, would also justify any case where a doctor or nurse deceives a patient into taking a medication or undergoing a procedure that is believed to be more effective than the one the patient would have chosen in the absence of deception.

By analogy with the “flatmate” example, in any case where a patient has a choice of treatment A or treatment B, if he is leaning towards treatment A when, all other things being equal, the doctor or nurse believes treatment B would be better, then this doctor or nurse would be justified in using deception to change his mind. And if a surgeon judges that an operation is a necessary means towards a patient's end, and that deception about the likely chances of something going wrong is necessary to lead the patient to will those means, then this act of deception would also be covered by this account. Whether an act of deception is selfless, or intended to facilitate an agent's own end, or even to cause an agent to will a necessary means to her own end, whether it is carried out by lying or through true deceptive statements – none of these considerations are specific to placebo deception, rather than paternalistic deception in general. Gold and Lichtenberg, then, and thus far I as well, have failed to identify the relevant features of placebo deception that make it distinct from other cases of paternalistic deception in a healthcare context. It is this that I'd like to examine in the final part of this chapter.

A Kantian defence of placebo deception

I previously argued that, in the “flatmate” example, deception might possibly be justified in a Kantian framework if it was carried out in order to cause my flatmate to will a necessary means to her end. However, this is where the analogy with a case of placebo deception breaks down. In this example, my flatmate has chosen to will an ineffective means towards her end, and my act of deception was an attempt to make her will an effective one instead. In the ideal placebo case, where a deceptive placebo presents the best or only means of achieving a therapeutic benefit, the only means that are available to the patient's will are either sub-optimal or totally ineffective towards the patient's end. There is no more effective means that the patient can be convinced or deceived into willing instead – the most effective means towards the patient's end is an unconscious response that can only be achieved through deception. The patient can be deceived into willing “take this pill”, which will lead to the desired outcome, much in the same way my flatmate can be deceived into willing “leave the house” at the correct time – but this is not the same as willing the necessary means to their end. The actual means towards the patient's end, would be something like “take a placebo in order to provoke the placebo response”, which is not something that a rational agent can will. The effective means in this case is simply beyond the reach of the agent's will.

This is the reason a case of placebo deception is significantly different to other cases of deceptive paternalism in a healthcare context. The placebo case is not one where a patient just needs a deceptive “nudge” to lead them to will effective treatment B rather than the less effective treatment A. In all such cases, as well as the “flatmate” example, it

seems that the right course of action would actually be to explain to the agent why the means you think they should will in this situation are better than the means they currently will. If, after I've explained to my flatmate the reasons why I think she should leave the house at 7:30am, after I've shown her the reports of road works, the weather report, the estimated journey time on Google Maps and the like, if after all this she still chooses to will “leave the house at 8am”, then I believe that I've exhausted my permissible options for changing her will at this point. If reason and argument fail to change my flatmate's will, then deceiving her into acting or willing differently would be a violation of her autonomy. After all, autonomy is the ability to choose our own ends and the means by which we pursue them, and for this choice to be meaningful we must be allowed to sometimes choose means that are less than optimal.

Likewise, in a medical case where a patient has chosen treatment A when a doctor or nurse believes treatment B would be better, if the patient still favours treatment A after being told all the advantages of treatment B and all the downsides of treatment A, then deceiving that patient so they pick treatment B as their willed means would be a violation of their autonomy, and impermissible. But what about where treatment B is a placebo treatment? Placebo deception is not a case where deception is used to manipulate a patient into willing an effective means that they could will, but for one reason or another have chosen not to. The means in question, the placebo effect, is simply not available to the patient in the absence of deception. A doctor or nurse could use persuasive arguments in favour of taking placebo treatment B, much like they could for taking active treatment B: the expected benefits, lack of side-effects, and so on, but these arguments can never cause the patient to will the placebo means in the way they could will an active treatment. Deception cannot cause the patient to will the placebo means either, although deception

can cause them to carry these means out. The nature of the effective means here, the unconscious placebo response, is such that these means are not available to the patient's will, although they are achievable by the patient's body. This is the reason that placebo deception is significantly different to other cases of paternalistic deception in healthcare: manipulation of the patient's will does not merely lead her to choose one of her available means rather than another, but rather it leads her to effect a means that would otherwise be unavailable to her.

Let us concentrate for the moment on cases where a placebo is the only available effective means towards the patient's end. Say I am in pain following an earthquake or some other natural disaster, and there are no active painkillers available. However a doctor knows that an injection of saline is likely to provide significant pain relief, so long as I do not know that it is just saline in the syringe. The placebo injection, and the subsequent response, is a necessary means to my end of pain relief – there is no question that I *should* will it, if I could. It is, as Kant argues “indisputably necessary”. It is also, to some extent “in my power”: it is something that my body can do. It's just not something that I can consciously *will* my body to do. In such a case, where the necessary means to my end is within my power to achieve, but cannot be willed in the usual way, then I think that an act of deception necessary to make these means available to me could be entirely consistent with respect for rational agency. It is entirely possible to recognise the ultimate value that a Kantian assigns to the will, while still acknowledging that the will has a purpose – enabling us to achieve our ends – and that it is limited in its ability to achieve this purpose.

To some extent, then, Gold and Lichtenberg were correct when they argued that deceptive placebo use is justified where it helps the patient achieve their end. But if we look at the

way that placebo deception does so, and in particular the effect it has on the functioning of an agent's will – that process of weighing and choosing ends and means that is so sacred to Kantians – we see why this type of deception is singularly justifiable on Kantian terms. Placebo deception does not hinder or pervert the functioning of an agent's will, but rather enhances it. It is not the case that this kind of deception merely manipulates an agent into choosing one of her available means instead of the one she would have chosen of her own free will. Rather, placebo deception creates a choice of means that would not exist in its absence. When the doctor in the previous example withholds from me the information that the painkilling injection she offers is only a shot of saline, that act of deception gives me a choice of means towards my freely willed end that would not otherwise be available to me. I can now weigh this means against my other options in the usual fashion – this process is not perverted or derailed by the deception, but rather facilitated by it.

Herman argues that the rejection of maxims of deception by the CI test shows us the inappropriateness of maxims that "exploit the vulnerability of human agents to manipulative control" and "fail to respect the integrity and separateness of will that is constitutive of rational agency".[59] However, I would argue that placebo deception respects the integrity and separateness of the human will, and that it is not carried out in order to exert manipulative control over the patient. If it is indeed the case that the integrity of a patient's will is intact following placebo deception, then the grounds on which deceptive maxims fail the CI test do not hold in this case. The "deliberative presumption" against acts of deception can be rebutted in the case of placebo deception.

The absolute value of the will, as Herman argues, allows us to see that someone's will cannot be interfered with as a means to our ends, or even as a means to their own ends.

59 Herman (1993) p154-55

Even saving a life doesn't justify “taking” someone's will in this way, and this is established by the CI test. But placebo deception doesn't involve “taking” someone's will. When a doctor offers me a painkilling injection of saline, without telling me that it is saline, it is not the case that this act of deception subordinates my will to hers, or turns my will into a tool of her own. If my will was subordinate to hers, I would be unable to freely choose for myself the ends I wish to pursue and the means by which to pursue them. But following this act of deception my ends remain my own to choose, and my choice of means is not diminished or perverted by the manipulation. In this way, placebo deception is also consistent with the Kantian imperative that we treat others always as ends in themselves, and never as a means to an end.

In the placebo case, a doctor or nurse recognises that the necessary means towards my willed end is not a willed means, but an unconscious response. This is the reason that manipulation of the will is necessary in such cases – not to lead the patient's will to go one way rather than another, but to cause the patient's body to carry out a response that the patient can't achieve otherwise. This manipulation, then, does not devalue the will, or violate its integrity. Placebo deception is consistent with respect for an agent's will as the faculty by which an agent chooses, and by and large achieves, their ends. However it does require acknowledging that the will is only one of the tools we have for achieving our ends, and that sometimes the means by which we achieve our ends are outside the realm of willed action. And this is not strictly limited to cases where placebos help us to achieve therapeutic benefit – many of the important aims that people have in life cannot be achieved by willed action. Things like falling in love, or being inspired to create, or imagining novel solutions to complex problems, are all important ends that agents have, but the means by which these ends are achieved are not found in the sphere of willed

action. To recognise this is not to deny the value that a Kantian affords the will, merely to acknowledge its limitations as a human faculty.

What about cases where a placebo is not the only necessary means to my end, but merely the best means available? This case is a little less clear-cut. Say I have mild depression, and my doctor has a choice of offering me a placebo treatment, or an antidepressant that has recently been shown to be no more effective than a placebo, and which has some unpleasant side-effects. In this case there is a means toward my end within the sphere of willed action, but the unconscious response that results from placebo deception is a more effective means toward that end. If my doctor offers me a choice of the two treatments, explaining that the effectiveness of the two is roughly on a par, concealing the nature of the placebo and detailing the side-effects of the active treatment, it seems likely that I'm going to opt for the placebo treatment. It would therefore be tempting to argue that the deception in this case, concealing the nature of the placebo treatment, has caused me to choose one way rather than another, and thus has interfered with the free function of my will. But this is not an accurate account of the effect that the act of deception has in this case. It is not that the act of deception influenced my choice in one direction rather than the other, rather it is only because of the deceptive act that I have a choice at all. The decision I make is based on the true information presented to me – the likely effects of the two treatments. The deception was not carried out to influence or interfere with my choice, but only to make it possible for me to choose. Revealing the nature of the placebo treatment would not facilitate the free function of my will, rather it would narrow my options and force me to choose either the active treatment or nothing at all.

Knowledge of the nature of the placebo option precludes it from being a possibility. In the

ideal placebo case, then, providing a patient with this information this removes the most effective or only means they have to achieve their end. I therefore argue that withholding this information, and only this information, is entirely consistent with respect for the patient's rational agency. The alternative would be to reduce the patient's choice of means, and their ability to achieve their ends. To do this in the name of the patient's autonomy, or out of respect for their free agency, simply seems wrong-headed.

This account of placebo deception, where deception is carried only out in order to make the best or only means to an agent's end available to that agent, also supports the distinction I make between deceiving a patient with only true statements, and lying to them. It is true that placebo deception, carried out by making only true statements and withholding only that information that is necessary to preserve the possibility of the placebo response, may well cause the patient to have a false belief that there is active medication in the treatment, and it may well be the case that this false belief is instrumental in effecting patient's the placebo response. However, the fact that the patient has been given only true information about the treatment, its likely effects, method of delivery, and so on, allows them to make a free and informed choice about whether or not to consent to this treatment. While a lie – providing false information that one believes to be false – might also facilitate the placebo response, the false beliefs caused by the falsity of the accompanying information hinder the patient's ability to make a free and informed choice about the treatment. For example, if a doctor offers me a placebo shot of saline, telling me what the likely effects are but concealing its placebo nature from me, I can use that information to decide whether consenting to this shot is in accordance with my ends. It might be the case that I am terrified of needles, and the stress that this shot will cause me outweighs the expected benefits – but I can only make a free and informed decision on

this point if the information I have been given about the treatment, apart from its placebo nature, is true. If, on the other hand, a doctor offers me a placebo shot and tells me it is morphine, my expectations of the effects of the treatment may be entirely divorced from reality. I might expect an unreasonably high level of pain relief from the treatment, and decide that it is worth the high level of stress the shot will cause me, when in fact this is not the case. Conversely, I might be a recovering heroin addict, and turn down a useful and harmless treatment for fear of a relapse. I will examine these concerns in greater detail in the next chapter, but for now it is enough to note that lying placebo deception interferes with a patient's ability to make free and informed decisions in a way that the kind of deception I am arguing for here, withholding only that information necessary to make a placebo treatment an option, does not. Placebo deception carried out through lying to a patient violates the integrity of their will, and it cannot be justified under a Kantian framework.

I therefore offer the following guidelines for the ethical use of deceptive placebos in clinical practice:

- A deceptive placebo may be offered to a patient where there is good reason to believe that the placebo response presents the best or only means to a patient's end.
- The deception involved in a placebo treatment must be limited to withholding the placebo nature of the treatment, where that is necessary to facilitate the placebo response.
- All other information provided about the available treatments to the patient must be true. The practitioner must not lie to the patient.

- The deception must be carried out with the intention of making the placebo response available to the patient as a means to their end, and not in order to otherwise influence their free choice of means towards that end.
- The nature of the placebo treatment must be revealed if the patient specifically asks for this information, or otherwise after the placebo treatment has ceased to be a useful means for the patient.

The arguments I offer in defense of placebo deception do not cover other kinds of paternalistic deception in a healthcare context. Deceiving a patient, even with only true statements, into choosing one active treatment over another, or into consenting to a treatment that they would otherwise refuse, interferes with the functioning of their will in a way that placebo deception does not. This kind of deception is not carried out in order to make a means available to a patient, but to manipulate them into choosing one of their available options rather than another. It therefore is not consistent with respect for their free agency. Likewise, non-beneficent placebo deception is ruled out. Where a doctor or nurse deceives a patient with a placebo treatment merely in order to mollify them, this deception is not carried out with respect for the patient as an end in themselves. Rather the patient here is used as a mere means to the doctor or nurse's end of getting some peace and quiet. This kind of placebo deception, even if it otherwise conforms with the guidelines I have established here, is not justified under a Kantian framework.

I therefore conclude that only in the ideal placebo case, where a doctor or nurse has good reason to believe that a deceptive placebo treatment presents a patient's best or only chance at achieving some therapeutic benefit, the deception involved in a placebo treatment can be carried out in a way that is entirely consistent with respect for the

patient's autonomy. The only kind of deception justified by this argument is withholding only that information necessary to make the placebo response a viable means, and otherwise giving the patient true information about the treatment. Lying to the patient is ruled out, as is continuing to withhold information on the nature of the treatment if the patient has directly asked for it – to do so would be to fail to respect the patient's rational agency. Placebo deception, carried out in order to help the patient achieve their end of therapeutic benefit, does not violate the integrity of the patient's will, interfere with its functioning or subvert it to the will of someone else. Rather it is entirely in accordance with the function of the human will: the free choice of ends and the means by which we pursue them.

Conclusion

I believe that I have achieved the task I set out for myself at the beginning of this chapter. I have found a place for deceptive placebo treatments in the unforgiving environs of a Kantian ethical framework. Placebo deception has passed the “stress test” - the ethical school that is least flexible in its prohibition of deceptive acts has given it a thumbs up. By examining Christine Korsgaard's preferred account of Kant's Categorical Imperative, I have demonstrated how a properly constructed maxim of deceptive placebo use could pass this test. Upon examination of Barbara Herman's interpretation of the same test, however, it appears that the Kantian prohibition on interfering with the will of rational agents – even in order to save a life – presents a larger problem. But an analysis of the effect of placebo deception on the will shows why this particular kind of deception might not violate this rule. Moreover, this analysis highlights the features that separate my ideal

placebo case, where a placebo is administered for its therapeutic effect and the deception involved is limited to withholding information on the inert nature of the treatment in order to facilitate this effect, from less salubrious cases of placebo use or paternalistic deception. Placebo deception, according to this analysis, is only permissible where this can be carried out without lying, and then only in order to make the placebo effect available as a means towards a patient's end. In this way, the act of deception does not seek to pervert or manipulate the functioning of the patient's will, and is therefore consistent with respect for their autonomy.

In the next chapter I will address a related problem: whether placebo deception can be carried out in a way that meets not just the general requirements of patient autonomy, but the specific requirements of informed consent that exist to protect this autonomy in a healthcare context.

References

Carson T (2010) *Lying and Deception: Theory and Practice.* Oxford: Oxford University Press

Gold A, Lichtenberg P (2014) “The moral case for the clinical placebo” *Journal of Medical Ethics* 40: 219-224

Herman B (1993) *The Practice of Moral Judgement.* Cambridge: Harvard University Press

Hobbes T, in Tuck R (ed.) (1996) *Leviathan.* Cambridge: Cambridge University Press

Kant I, in Beck LW (ed.) (1949) *Critique of Practical Reason and Other Writings in Moral Philosophy.* Chicago: University of Chicago Press

Kant I, in Gregor M, (ed.) (1996) *The Metaphysics of Morals.* Cambridge: Cambridge University Press

Kant I, in Gregor M, Timmerman J (eds.) (2012) *Groundwork of the Metaphysics of Morals (revised ed.)* Cambridge: Cambridge University Press

Korsgaard C (1996) *Creating the Kingdom of Ends.* Cambridge: Cambridge University Press

Masson D (ed.) (1968) *The Collected Writings of Thomas de Quincey (vol 4)*. New York: AMS Press

O'Neill, O (1989) *Constructions of reason: explorations of Kant's practical philosophy.* Cambridge: Cambridge University Press

Wood A (1999) *Kant's Ethical Thought*. Cambridge: Cambridge University Press

5. Placebo treatments and informed consent

Perform your medical duties calmly and adroitly, concealing most things from the patient while you are attending to him. Give necessary orders with cheerfulness and sincerity, turning his attention away from what is being done to him... revealing nothing of the patient's future or present condition, for many patients through this course have taken a turn for the worse.

– Hippocrates, *Decorum*

Introduction

While the Hippocratic tradition may have stressed the therapeutic benefits of concealing from a patient as many details about their condition and treatment as possible, contemporary medical ethics generally holds that doing this would violate the principle of respect for autonomy. In a clinical context, the chief practical application of this principle lies in the duty to obtain informed consent from patients before proceeding with a treatment. This is one of the cornerstones of contemporary clinical practice – patients have the right to choose what happens to them in a healthcare context, and the autonomous choices of patients are best protected by the requirements of informed consent. A treatment should not be carried out unless the patient has voluntarily consented to it, and in order to legitimately consent to a treatment, a patient must be adequately informed about that treatment. Thus, informed consent.

Informed consent requirements dictate that it is unacceptable to wilfully conceal from a patient important details of their condition and treatment, but these requirements do not

compel a clinician to reveal absolutely everything that they know about this condition or treatment. To do so would be impractical in the extreme, and such a volume of information would most likely serve to impede patient understanding and autonomous choice, rather than facilitating it. So, given that there is some information that must be disclosed to meet the requirements of informed consent, and some information that can be omitted without violating these requirements, the key question for this chapter is whether the information that a particular treatment is a placebo belongs in the former category or the latter. That is, whether or not it is possible for a patient to give their informed consent to a placebo treatment without knowing the exact (inert) nature of that treatment, or whether withholding this information violates the rules of informed consent. If the obligations of informed consent can be met by informing a patient of the risks, costs, and likely outcomes of the treatment, as well as any alternatives, but not its exact nature, then this presents a possible way to ethically administer placebo treatments without disclosing information that would make the placebo effect impossible.

In order to ascertain whether an undisclosed placebo treatment can meet the requirements of informed consent I will examine different standards of disclosure and investigate how each applies to the placebo case. The “subjective standard” of disclosure, where the level of disclosure necessary for informed consent is judged by reference to the needs of each patient individually, is the most promising standard by which to judge the appropriate level of disclosure in most cases where a clinician may be considering a placebo treatment.

While the level of disclosure required for informed consent will vary from patient to patient under the subjective standard, there is still a minimal level of information, such as the aims and likely risks and outcomes of a procedure, without which an autonomous choice about

that procedure simply does not seem to be possible. Knowledge of the exact nature or mechanism of action of the procedure, however, does not seem to be in this category. And as I shall argue, if informed consent does not necessarily require disclosure of the nature of a treatment, then an across-the-board rule requiring disclosure of the exact (inert) nature of all placebo treatments is inconsistent. Therefore it is possible that, in some cases at least, a deceptive placebo treatment may meet the standards of informed consent. Whether or not this is the case for any given patient will depend on that patient's needs, as well as their history, personality, and so on. Some patients will definitely regard the fact that a proposed treatment is a placebo to be important information to know, but some may not, so in such cases patient autonomy is not best served by a blanket rule requiring disclosure of this information.

What is informed consent?

Consent is the means by which one person authorises another to take an action which would otherwise be in violation of the consenter's rights. For example, an act of consent makes the difference between someone borrowing my bike and their stealing it. In their classic work *A History and Theory of Informed Consent*[1] Faden and Beauchamp identify two distinct meanings of the term “informed consent”. Sense$_1$, which draws on the principle of autonomy, defines informed consent as an act of “autonomous authorisation”:

> [I]nformed consent in sense$_1$ is given if a patient or subject with (1) substantial understanding and (2) in substantial absence of control by others (3) intentionally (4) authorizes a professional [to carry out an intervention].[2]

1 Faden and Beauchamp (1986) p276-287
2 Ibid. p278

Sense$_2$, on the other hand, defines informed consent as a "legally or institutionally effective authorisation":

> Such an authorization is "effective" because it has been obtained through procedures that satisfy the rules and requirements defining a specific institutional procedure in health care or research.[3]

So an act of informed consent in Sense$_1$ would mean that a course of action is morally acceptable, or at least that it does not violate the principle of respect for autonomy, while an act of informed consent in Sense$_2$ would mean that a course of action conforms to the rules and practices of a certain institution. Of course there is a lot of overlap between the two senses – a person can autonomously authorise something in a way that is legally or institutionally effective – but they diverge in some important ways. Someone who is under the legal age of majority may nonetheless have the substantial understanding and freedom from coercion necessary to autonomously authorise a course of action. On the other hand, someone who signs a sheaf of consent forms without reading or understanding them is providing effective authorisation, but could not be said to be providing this authorisation autonomously. In this chapter I am mostly going to be concerned with the question of whether placebo treatments can satisfy the requirements of Sense$_1$ informed consent: whether it is possible for a patient to autonomously authorise a placebo treatment without knowing its inert nature. There can be no single answer to the question of whether placebo treatments can satisfy the requirements of Sense$_2$ informed consent, as these requirements vary depending on the institution or jurisdiction in question.

The concept of consent has a long history in Western economic and social thought, but the history of informed consent in medicine is rather shorter. The doctrine of informed consent

3 Ibid. p280

in biomedical ethics is generally thought to begin with the Nuremberg Code, which was formulated in 1947 in response to the abuse and murder carried out in Nazi Germany in the name of medical research. The primary assertion of the Nuremberg Code is that in all research on humans

> The voluntary consent of the human subject is absolutely essential. This means that the person involved should have legal capacity to give consent; should be so situated as to be able to exercise free power of choice, without the intervention of any element of force, fraud, deceit, duress, over-reaching, or other ulterior form of constraint or coercion; and should have sufficient knowledge and comprehension of the elements of the subject matter involved, as to enable him to make an understanding and enlightened decision.[4]

While the Nuremberg Code was formulated as part of a court judgement, the international medical profession's first major attempt at self-regulation in medical research came with the Declaration of Helsinki in 1964. The Declaration builds on the principles of the Nuremberg Code, and likewise makes informed consent a key requirement of ethical research. However the Nuremberg Code's “absolutely essential” informed consent requirements were relaxed slightly in the Declaration, where allowances are made for research involving incompetent participants to be carried out where proxy consent, such as a legal guardian, is available:

> 3a. Clinical research on a human being cannot be undertaken without his free consent, after he has been fully informed; if he is legally incompetent the consent of the legal guardian should be procured.
>
> 3b. The subject of clinical research should be in such a mental, physical, and legal state as to be able to exercise fully his power of choice.
>
> If at all possible, consistent with patient psychology, the doctor should obtain the patient's freely given consent after the patient has been given a full explanation. In case of legal incapacity

4 *BMJ* (1996) 313:1448

> consent should also be procured from the legal guardian; in case of physical incapacity the permission of the legal guardian replaces that of the patient.[5]

Later revisions of the Declaration made allowances for research involving incompetent participants to be carried out even where proxy consent is not available:

> 29. Research involving subjects who are physically or mentally incapable of giving consent, for example, unconscious patients, may be done only if the physical or mental condition that prevents giving informed consent is a necessary characteristic of the research population. In such circumstances the physician should seek informed consent from the legally authorized representative. If no such representative is available and if the research cannot be delayed, the study may proceed without informed consent provided that the specific reasons for involving subjects with a condition that renders them unable to give informed consent have been stated in the research protocol and the study has been approved by a research ethics committee. Consent to remain in the research should be obtained as soon as possible from the subject or a legally authorized representative.[6]

Early revisions of the Declaration drew a sharp distinction between "research combined with professional care" – where the research is expected to provide therapeutic benefit to patients – and research that is "non-therapeutic", or carried out purely for scientific reasons. While informed consent was necessary for non-therapeutic research, if the research carried the promise of therapeutic benefit for participants then consent requirements could be waived:

> II.5 If the physician considers it essential not to obtain informed consent, the specific reasons for this proposal should be stated in the experimental protocol for transmission to the independent committee.[7]

According to Faden & Beauchamp, this exemption was justified by "the same beneficence-

5 World Medical Association (1964)
6 World Medical Association (2008)
7 World Medical Association (1989)

based principles that support the physician's therapeutic privilege in medical practice"[8] – that is, the idea that sometimes informing a patient can cause them harm, either by by compromising the effectiveness of their treatment or by causing them psychological or emotional harm. In such cases a physician's duty to heal, or not harm, their patient may override their duty to inform them. However, this beneficence-based exemption was removed in the 2000 revision of the Declaration.[9]

Although the doctrine of informed consent was formulated to regulate medical research, it subsequently gained recognition worldwide as a requirement for clinical practice as well. However, extending the informed consent requirements governing research to the sphere of clinical practice was far from simple. While the Helsinki Declaration acknowledges that in some cases research involving incompetent participants must sometimes proceed in the absence of proxy consent, this occurrence is far more routine in clinical practice, where treatment cannot be denied to unconscious or severely mentally impaired people, or the very young, simply because they are unable to give their consent and a legal proxy is unavailable. Furthermore, clinical treatment must often be given under emergency conditions, where the time taken in obtaining consent might compromise the patient's treatment. Public health considerations may also override informed consent obligations, in cases where compulsory quarantine or vaccination is judged necessary to prevent the spread of a disease.

So rather than being "absolutely essential" as stated in the Nuremberg Code, the requirement to obtain informed consent in clinical practice should be seen as a *prima facie* duty, one which can be overridden by more pressing concerns. Furthermore, according to

8 Faden and Beauchamp (1986) p156
9 World Medical Association (2000)

Manson and O'Neil[10] informed consent requirements are not uniform, but should vary according to the nature of the treatment being consented to – the standard of informed consent required for a non-invasive or minimally invasive procedure should be much lower than for a major treatment which could carry serious risks or consequences for a patient. Explicit written consent may be necessary for major surgery, but for a routine blood sample or injection consent inferred from the patient's actions, such as extending an arm to the nurse, may be sufficient. And for very non-invasive procedures, such as the visual appraisal of a facial mole, the need for any informed consent at all is not immediately apparent.

If we accept that the standards of informed consent to a treatment should vary according to the magnitude of the risks of that treatment, the kind of effect it is likely to have on a patient's life, what does this tell us about the duty to obtain informed consent to a placebo? Placebo treatments typically carry less serious risks and side-effects than their active counterparts, so does this mean we can set the standards of informed consent for these treatments correspondingly lower? This would imply that a clinician who believes that a patient could benefit equally from either a potentially addictive painkiller or a placebo treatment has a duty to disclose more information on the active painkiller than the placebo.

It would be a mistake, however, to say that there are no risks to a placebo treatment. Apart from the nocebo effect, which I will examine in a later chapter, the risks of a placebo treatment lie in what may happen if the nature of the treatment is discovered by the patient. This may happen in any number of ways: the clinician may “unblind” the placebo after the course of treatment is over, for example, or the patient may simply investigate out of their own curiosity. On discovering that they have been prescribed a placebo patients

10 Manson and O'Neill (2007) p81-82

may feel embarrassed, angry, or cheated; this may harm their relationship with the clinician in question. More seriously, it may damage their faith in the whole medical profession, and if publicised could bring some disrepute to the institution involved. These risks are fairly minor compared to the actual physical risks of injury, infection, addiction, and the like associated with active medication. But they are non-trivial, and if the likely risks of a treatment do have a place in assessing the informed consent requirements of a treatment, it is these risks that must be considered.

Lying, deception, and disclosure

It is important to note here that when discussing *deceptive* placebo use, the deception involved is non-disclosure, as opposed to outright lying. A clinician who knowingly gives a patient a saline injection and tells them it is morphine is lying – they are giving the patient information they know to be false. A chapter on whether or not this kind of deceptive placebo can meet the requirements of informed consent would be much shorter. It can't. A patient cannot give their informed consent to a treatment if they have been lied to about its nature – consent given to a morphine shot cannot be said to extend to a saline shot given in its stead.

On the other hand, a clinician who gives a patient a saline injection and tells them “this is to help with your pain” is not lying – the intention of the treatment is pain relief, and saline injections have been shown to be effective at this. While they are not giving the patient any false information, however, the clinician is deliberately withholding information on the exact nature of the injection. Furthermore it is reasonable to assume that the patient will form the

false belief that the injection contains an active medication, and that this belief will be the basis of the expectation of benefit that may lead to the placebo effect. Whether or not this kind of deception is consonant with the requirements of informed consent depends on level of disclosure dictated by these requirements. If a patient needs to know everything about a treatment in order to consent to it, then withholding the information required for placebo deception would be inconsistent with informed consent. If, however, it is possible for a patient to give their informed consent to a pain-relieving injection without knowing whether it contains an active painkiller or just saline, then it would be possible to administer a placebo and meet the requirements of informed consent.

How much disclosure?

The Nuremberg Code states that in order to consent to participate in medical research, a participant must "have sufficient knowledge and comprehension of the elements of the subject matter involved as to enable him to make an understanding and enlightened decision."[11] The same standard applies to consent in a clinical context – a patient needs to be provided with, and understand, information about a treatment in order to consent to it in a meaningful fashion. Unfortunately, there is no consensus on how much information a patient ought to be provided with in order to make an "understanding and enlightened decision". Instead, there are a range of answers. At the extreme end of this range is the answer posited by Carl Schneider in *The Practice of Autonomy*[12]. Schneider argues that in order to provide their informed consent to a treatment patients are required to understand all information on the treatment, down to the most technical details, and know all the

11 *BMJ* (1996) 313:1448
12 Schneider (1998)

possible outcomes of it. This standard, which we might call “full disclosure”, is highly impractical – the level of comprehension and foresight required to give consent would be nearly impossible for anyone under ideal conditions to attain, let alone for a layperson in a stressful situation with limited time – and leads to Schneider's conclusion that true informed consent is impossible.[13]

A more moderate answer than Schneider's is that in order for a patient to properly consent to a treatment, their clinician is required to provide them with all the information *material* to their decision. This requirement, which we might call “substantial disclosure” recognises that people do not need to approach omniscience in order to make an intentional decision informed by their values, goals and preferences. This makes informed consent a realistic goal, and brings the informational requirements for informed consent more in line with the standards of free and intentional action that apply in other areas of life, such as financial or legal decisions.

One such “substantial disclosure” standard is Faden and Beauchamp's $Sense_1$ informed consent, which I outlined earlier. They argue that informed consent is given when a patient with:

(1) substantial understanding

(2) in substantial absence of control by others

(3) intentionally

(4) authorises a professional [to carry out an intervention].[14]

13 Ibid. p174-175
14 Faden and Beauchamp (1986) p178

Applying this standard to the placebo case, we are left with the questions of whether a patient's understanding of a treatment could be said to be “substantial” when they don't know the inert nature of the treatment. So, how much disclosure is required for “substantial understanding”? In *Principles of Biomedical Ethics*, Beauchamp and Childress outline three different standards by which adequate disclosure of information may be measured:

The professional practice standard: The practices of the professional community determine the level of disclosure necessary for informed consent. Also known as the “reasonable doctor” standard.

The reasonable person standard: Information is material if a hypothetical reasonable person would deem it significant in deciding whether to undergo a procedure.

The subjective standard: The level of disclosure is judged by reference to the needs of each patient, and thus will vary from case to case.[15]

Of these, the subjective standard, which takes the needs of individual patients into account, is ideal from an autonomy perspective. However, it is less than practical as often neither the patient nor the clinician involved is in a position to know what the patient's particular informational needs are in a given situation. In the US at least, the “professional practice” standard was historically accepted by most courts as the legal standard of informed consent, but the “reasonable person” standard has been steadily gaining popularity.

In Australia, the *Rogers v Whitaker* case produced a ruling that endorsed both the

15 Beauchamp and Childress (2009) p122-123

reasonable person and subjective standards of disclosure, at least as far as the risks of a treatment are concerned. The High Court ruled that, from a legal point of view, a medical practitioner has a duty to warn a patient of material risks inherent in a proposed treatment, and that

> A risk is material if, in the circumstances of the particular case a reasonable person in the patient's position, if warned of the risk, would be likely to attach significance to it or if the medical practitioner is or should reasonably be aware that the particular patient, if warned of the risk, would be likely to attach significance to it.[16]

Applying each standard to the placebo case, we can see that the professional practice standard is not very helpful. Professional guidelines vary: some medical authorities have condemned the clinical use of placebos, others are more supportive, but most have not issued any advice one way or the other. Nonetheless, as I outlined in chapter three, surveys indicate that clinical placebo use is prevalent, and that doctors and nurses around the world regularly use them in clinical practice. This does not tell us, however, that doctors and nurses around the world *should* be using placebos so regularly. It may be that case that placebos should be used far more rarely in clinical practice, or not at all; this may just be an example of poor medical practice that has gained prevalence around the world.

What about a hypothetical “reasonable person”? Would such a person consider the fact that a treatment is inert to be information material to their decision whether or not to consent to it? This is a difficult question. As I also discussed in chapter three, there is some data to support the argument that patients believe undisclosed placebo use to be legitimate, and thus that the inert nature of a such treatments is not material. However, it is not possible to extrapolate from data such as this the conclusion that a hypothetical

16 High Court of Australia (1992)

reasonable person would agree that the nature of a treatment is immaterial. Alternatively it is possible to argue that, in a case where the best available treatment is in fact a placebo, the reasonable course of action is to stay ignorant of this fact, if disclosure of it would effectively nullify the effectiveness of the treatment. But once again, just because it could be reasonable to stay ignorant of a fact, this does not necessarily mean that a hypothetical reasonable person would consider this fact to be immaterial information.

To determine whether the information that a certain treatment is a placebo is material to a hypothetical reasonable person, we have to examine the part that it plays in this person's decision-making process. Consider a reasonable patient's decision whether to consent to treatment A, which has a 35% success rate, and comes with a low cost and no side effects; treatment B, which has a 45% success rate, but is expensive and carries several side-effects, or neither. The information that treatment A is an inert placebo is material to this decision if it has an influence on the patient's decision to consent to A or B or neither. But in assessing this decision-making process we come once again on the problem that we saw last chapter – the information that a treatment is a placebo affects not merely the decision-making process – as material information should – but it affects the actual decision to be made. After the disclosure that treatment A is an inert placebo, the decision whether to consent to treatment A, which has a 35% success rate, or treatment B, which has a 45% success rate, or neither, no longer exists. The information has created a new decision, which is whether to consent to treatment A_1, which is ineffective, treatment B, or neither. Can this piece of information then be said to be material to the hypothetical reasonable patient's original decision, when after the disclosure the original decision is no longer able to be made? This conundrum, created by the unconscious nature of the placebo effect, makes it difficult to assess the placebo case in a hypothetical fashion.

The subjective standard of disclosure, where each patient's history, personality, aims, values, and so on are taken into account in assessing what to disclose, is a more promising standard by which to assess placebo treatments. The oft-stated shortcoming of the subjective standard is that in order to judge the level of disclosure appropriate for each patient, a clinician needs a familiarity with the patient that is impossible, or at least impractical, in many clinical situations. However such a familiarity would be appropriate in most of the clinical situations where a placebo could be used effectively, as knowledge of a patient's history and personality allows a clinician to judge not only the information that it is appropriate to disclose, but also to more accurately judge the likely risks and benefits of a placebo treatment. For example, some people are known to be more susceptible to the placebo effect than others. If a clinician knows that a patient is a “placebo reactor”, then this should be taken into account when considering possible treatments. Furthermore, if a patient is generally mistrustful of the medical profession or has had bad medical experiences in the past, this will not only affect the likely magnitude of their response to a placebo treatment, but may also colour their likely reaction, should the nature of a placebo treatment be intentionally or inadvertently revealed to them. Conversely, if a patient has previously professed a belief in homeopathy, a clinician considering a placebo treatment might be inclined to suggest a homeopathic treatment instead, whilst judging it unnecessary to offer her opinion that such treatments are nothing more than placebos.

So far it seems as though a subjective standard is the most appropriate way to assess the appropriate amount of information that a patient needs to comprehend in order to make an informed decision about a placebo treatment. I believe that this standard of disclosure represents the most promising way to assess the permissibility of placebo treatments. If

both the likely risks and benefits of a placebo treatment, and the level of disclosure appropriate, vary from patient to patient, then it makes sense to determine on a patient-by-patient basis both whether or not a therapeutic placebo is the best treatment option and whether or not a level of disclosure that makes the placebo effect possible is appropriate. And if determining each patient's informational needs involves engaging with that patient and taking into account many of the same considerations that are involved in determining the likely benefits and risks of a placebo treatment, then the subjective standard of disclosure does not seem to place much of an additional burden on clinicians, over and above that which is necessary to determine the best therapeutic treatment option.

Disclosure does not guarantee comprehension

So far I have used the terms “substantial disclosure” and “substantial comprehension” almost interchangeably, but more disclosure on the part of the clinician does not necessarily mean more comprehension on the part of the patient. Indeed, too much disclosure can actually impede comprehension, causing “information overload” or turning the consent process into a meaningless box-ticking exercise. Conversely, if the patient happens to be knowledgeable about their condition and treatment options, for example if they have medical training or they have gained familiarity through a long history with a chronic illness, then disclosure might not be necessary for comprehension at all. Returning to the language of the Nuremberg code, information on a treatment should be disclosed with the goal of fostering “knowledge and comprehension” to facilitate and validate the decision-making process.

A careful explanation of the risks, costs, and likely outcomes of a patient's treatment options, as well as of refusing treatment, is one way of imparting the aforementioned knowledge and comprehension, and this is undoubtedly a valuable part of the decision-making process for many patients. But what are we to make of the opaque descriptions of a placebo treatments that are sometimes offered as a way of disclosing information about a placebo treatment without revealing its inert nature?[17] For example, describing a placebo pill as “working through a mind-body self-healing process that we don't totally understand yet”? Is this statement intended to add to the patient's comprehension of their options, or to obfuscate – to merely make the patient *feel* more informed?

It is hard to imagine that a decision-making process would be facilitated by the knowledge that a certain treatment works through a poorly-understood mind-body self-healing process. It is tempting, therefore, to dismiss this disclosure as useless information at best, a deceptive charade at worst, and to suggest that it has no place in a discussion of treatment options. But it is important to remember that the attitudes and expectations of both clinician and patient are a key component in the effectiveness of placebo treatments (and indeed in the placebo component of active treatments). So if it is the case that this information – that a certain treatment works through a poorly-understood mind-body self-healing process – helps engender confidence in a treatment, then imparting this information could be justified by appeal to its therapeutic benefit, rather than its place in the informed consent process.

17 See for example Lichtenberg et al (2004)

Informed consent standards for active and placebo treatments

Anne Barnhill[18] argues that deceptive placebo treatments are compatible with informed consent, because they meet the same standards of disclosure required to give informed consent to active medication. To assert this she turns to arguments Onora O'Neill made in her 1984 article “Paternalism and Partial Autonomy”. In this article, O'Neill argues that deceptive placebos (as well as “reassuring but inaccurate accounts of expected pain”) can be construed as “non-fundamental but indispensable and so permissible deceptions.”[19] Informed consent requires clinicians to inform patients and obtain their consent to only the *fundamental* aspects of their treatment – obtaining consent to all the non-fundamental aspects of a treatment is not necessary, and indeed is impossible. As O'Neill argues “patients can no more be asked to consent to every aspect of treatment than citizens can be asked to consent to every act of government.”[20]

What basis is there for arguing that the fact that a treatment is a placebo is a non-fundamental aspect of the treatment, and thus that there is no obligation to inform patients of this fact? Barnhill defends this assertion by comparing the standards of consent for placebo treatments with those for active treatments. She argues that while patients are usually offered information on the purpose and risks of potential treatments, along with the risk of nontreatment, they are not regularly informed of the exact mechanism of action of the drugs they are offered. If a patient is offered a choice of two active pain medications, she contends, a clinician needs to explain the risks, benefits, and likely outcomes of taking each, but needn't explain the different mechanisms by which each medication works. This indicates that the mechanism of action of a treatment – whether a certain pharmacological

18 Barnhill (2011)
19 O'Neill (1984) p176
20 Ibid.

or physiological pathway, or the placebo effect – is non-fundamental information.

> Patients are regularly informed of the purpose of potential treatments (e.g. a pain medication's purpose is to relieve lower back pain), the risks of potential treatments and nontreatment (e.g. the pain medication might cause birth defects if you're pregnant) and the likely benefits and outcomes of treatment and nontreatment (e.g. the pain medication will likely reduce pain in the short run, but the pain might go away by itself without treatment). But patients aren't regularly informed of the nature of the treatment beyond the bare minimum. For example, the patient who's prescribed the pain medication might be informed that it reduces pain by reducing inflammation, but she almost certainly won't be informed of the precise physiological mechanism whereby the medication works.[21]

Barnhill thus argues that deceptive placebo use is compatible with informed consent because it meets the level of disclosure routinely offered by clinicians to patients about active treatments. This means that a deceptive placebo meets the professional practice, or "reasonable doctor" standard of informed consent – that the level of disclosure necessary for informed consent is determined by the practices of the professional community. But this leaves open the possibility that the physiological mechanism of action of a treatment actually *is* information that many patients would find material to their decision, and thus is something that clinicians *should* routinely disclose (according to the subjective and reasonable person standards). The failure of clinicians to routinely offer information on the mechanism of action of various treatments could therefore be seen as an example of pervasive bad practice, rather than an indication that this is is an ideal standard of disclosure. Whether a similar argument can be made on the "reasonable person" standard – whether a hypothetical "reasonable person" would find the exact nature or mechanism of action of a treatment to be information material to their decision – is a matter for debate. However I have previously argued that the subjective standard is the best standard for our

21 Barnhill (2011) p237

purposes, and as this standard varies from person to person it cannot be used as a basis for a sweeping judgement that mechanism of action of a treatment is to be considered non-fundamental information for the purposes of informed consent. It may well be the case that the precise mechanism of action is fundamental information for some patients, and not for others.

Nonetheless, with reference to the subjective standard of disclosure it is possible to make a softer version of Barnhill's argument: that in at least some cases, disclosure of the exact nature or mechanism of action of an active treatment is not required for informed consent. In some cases, for example, it might be judged that this information is overly technical, and that it would serve to confuse the patient rather than to aid their decision-making process. In other cases it might simply be that the patient does not care *how* exactly their treatment works, so long as it works. But if this is true for informed consent as it applies to active treatments, then an across-the-board rule stating that the exact nature of all placebo treatments must be disclosed would mean that the standards of disclosure for placebo treatments are higher than those for a corresponding active treatment. In other words, it may be possible for a patient to give their informed consent to an ibuprofen pill without knowing that it's an anti-inflammatory, but it would be impossible for that patient to give informed consent to a sugar pill without knowing that it's a placebo. What could be the basis for such a difference in the informed consent requirements between placebo treatments and active treatments? Barnhill argues that there is none:

> Unless we believe in magic or mind-body dualism, there is a physiological mechanism underlying the effective use of placebos... Why would it be a fundamental aspect of a placebo pill that it works via one physiological mechanism (the placebo effect physiological mechanism) rather than another physiological mechanism, when it isn't a fundamental aspect of ibuprofen or

Celebrex that it works via one physiological mechanism rather than another?[22]

Placebos are, of course, *necessarily* deceptive. Informing a patient of the exact nature of an ibuprofen pill will probably not alter its potency very much, whereas revealing the same information about a placebo will negate its effectiveness. This prompts the conclusion that declining to disclose the nature of a placebo pill is deceptive in a way that declining to disclose the nature of an ibuprofen pill is not. However, it cannot be the case that the same level of disclosure – disclosing the aims and risks of a treatment but not its exact nature – is deceptive on the one hand and not deceptive on the other. If a patient can validly consent to an ibuprofen pill without knowing its exact nature or mechanism of action, then it must be possible to validly consent to a placebo pill with the same level of comprehension.

Placebo prohibition

As I have discussed previously, the American Medical Association and the British Medical Association have both issued ethical rules prohibiting the use of deceptive placebos in clinical practice. These rules effectively assert that the only permissible way to administer a placebo treatment in clinical practice is if the inert nature of this treatment, or the fact that it lacks a specific pharmacological effect, is disclosed to patients. The AMA rule, for example, asserts that "the use of a placebo without the patient's knowledge may undermine trust, compromise the patient-physician relationship, and result in medical harm to the patient."[23] By contrast, asking a patient to consent to possibly receiving an inert placebo treatment in the future "respects the patient's autonomy and fosters a trusting

22 Ibid.p238
23 AMA (2014) Opinion 8.083 - Placebo Use in Clinical Practice

relationship… and improves health outcomes"[24]

So it is argued that failing to disclose to a patient the (lack of) specific pharmacological effect of a placebo treatment violates their autonomy, undermines their trust, and can damage the patient-physician relationship. However, as we have seen the "specific pharmacological effect" of active treatments is information that is regularly not disclosed to patients, and in at least some cases withholding this information this appears to be consistent with patient autonomy and trust. A requirement that the specific pharmacological effect (or rather, the lack thereof) of a placebo treatment always be disclosed, then, suggests that placebo treatments specifically require a higher standard of disclosure than corresponding active treatments. But it does not seem that, from an autonomy perspective, the level of comprehension required to autonomously authorise an active treatment is less than that required to autonomously authorise a placebo treatment. And although there are risks associated with placebo treatments, they are generally lower than those associated with active treatments, so a standard of disclosure based on the level of risk would not seem to support this distinction either.

There is, however, a different *kind* of risk associated with placebo treatments – the risk that a patient, on learning they have been given a placebo, will feel hurt, or cheated, and lose faith in their physician, the institution, or the medical profession more generally. This risk does apply specifically to the placebo case – it is hard to imagine someone reacting in a similar way on finding out that the analgesic pill they were given contains ibuprofen. Could this, then, be a basis on which to rule that deceptive placebo treatments are always unacceptable? Perhaps, but are the stated aims of respecting patient autonomy, fostering the patient-physician relationship, and improving health outcomes best served by such a

24 Ibid

rule? As I argued earlier, the risk of a patient reacting negatively upon discovering that they have been given a placebo varies from person to person, so allowing clinicians to assess this risk on a person-to-person basis should minimise these negative outcomes, while still allowing for deceptive placebo treatments in some cases where these promise the best therapeutic benefit – an option that is lost when a blanket rule against placebos is enforced.

The AMA rule requires disclosure every time a treatment that "the physician believes has no pharmacological effect on the condition being treated"[25] is provided to a patient. I have argued that this creates an inconsistent standard of disclosure for placebo and active treatments, but this problem is compounded when we consider that many potential treatments are not cut-and-dried examples of sugar pills or saline injections, but exist in a kind of grey area between placebo and active medication. A few examples might help to illustrate this point:

Probable placebos: Acupuncture has been shown to be more effective than standard treatments for pain relief in some conditions.[26] It is widely thought to be a placebo treatment, but it may yet be shown to have some specific mechanism for pain relief. If a clinician believes that a patient would benefit from acupuncture, how should she describe the treatment in order to avoid deception? She could either describe it as a placebo, as is the general consensus, or as a treatment that's probably a placebo, but might have an as-yet unknown mechanism of action. Or she could put a positive spin on the information, and merely say that the mechanism of action is not yet known, but it has been shown to be effective. It is not clear that any of these descriptions is necessarily more or less deceptive

25 Ibid
26 See for example Cherkin et al (2009)

than the rest, or that patient trust and autonomy is better served by one rather than the other. It seems likely, however, that describing the treatment as a placebo will not promote the best therapeutic outcome.

Possible placebos: Some recent studies have indicated that antidepressants are no more effective than placebos in the treatment of mild and moderate depression.[27] This raises the possibility that, despite the purported mechanism of action of these drugs, they actually have no specific effect for these conditions, and any therapeutic benefit they provide is entirely due to the placebo effect. If a clinician is prescribing one of these drugs to a patient with mild depression, should she disclose the possibility that the treatment is a placebo? In other words, should she adhere to an "active medication" standard of disclosure, where the mechanism of action may not be considered material information, or a "placebo" standard, where it is definitely is? The AMA statement's wording suggests that disclosure would only be necessary if the clinician *believes* that the treatment has no specific effect, but it offers no advice as to what to do if she is unsure, or merely acknowledges the possibility that this may be the case.

Placebo-like effects: Even active medications have a significant placebo component to their therapeutic effect – morphine administered by a nurse at a patient's bedside is more effective than the same dose administered covertly by a machine.[28] If informed consent requirements mean that clinicians have to disclose the mechanism of action of placebo treatments, do they also have to disclose the placebo-like nonspecific therapeutic effect that accompanies the administration of active medications? Charlotte Blease addresses a problem similar to this in her paper "The principle of parity: the 'placebo effect' and

27 Kirsch et al (2008)
28 Colloca and Benedetti (2005)

physician communication."[29] She argues that the placebo effect should be considered part of a larger therapeutic phenomenon called the "positive care effect", which also encompasses the beneficial effects of a clinican's bedside manner, communication style, non-verbal cues, and the like. She argues that any obligation to inform patients of the placebo effect must also extend to informing them of the other therapeutic aspects of the "positive care effect". If keeping the therapeutic benefits of the placebo effect hidden from a patient is a violation of their autonomy, the same can be said of failing to disclose the part that positive communication, and other aspects of the therapeutic encounter, can play in a patient's recovery. If, however, we think that requiring a doctor or nurse to disclose that they are only acting in such a warm, assured and friendly manner because that's the best way to promote your recovery is absurd, then the same is true of a requirement to disclose the nature of a placebo treatment.

In all of the above examples I would contend that the AMA's stated aims of respecting patient autonomy, fostering the patient-physician relationship, and improving health outcomes are best served by allowing clinicians to determine the level of disclosure appropriate on a case-by-case basis, taking into account each patient's history, beliefs, preferences, and so on. It is not clear how a categorical rule prohibiting placebo use serves to foster patient autonomy and trust in many cases, and it rules out a course of action which may sometimes lead to the best therapeutic outcome. I therefore argue that these guidelines should be revised.

29 Blease (2012)

Alternative Approaches to Informed Consent

While practice of informed consent is usually thought to be one of the ways that the principle of respect for patient autonomy manifests in medicine, there are some who argue that the relationship between autonomy and informed consent is not so simple. For example, Onora O'Neill challenges the notion that what she calls "the ritual of informed consent" serves to secure patient autonomy:

> Informed consent procedures protect choices that are timid, conventional, and lacking in individual autonomy (variously conceived) just as much as they protect choices that are self assertive, self knowing, critically reflective, and bursting with individual autonomy (variously conceived).[30]

Conversely, O'Neill argues that the real purpose of informed consent is to ensure that healthcare professionals meet certain basic obligations to their patients – in particular, the obligations to refrain from deception or coercion. Informed consent procedures, she argues, provide "reasonable assurance that a patient (research subject, tissue donor) has not been deceived or coerced."[31] These basic obligations not to deceive or coerce others, then, give a clincian reason to ensure that a patient is informed about, and consents to, any treatment they administer: "Our aim in seeking others' consent should be not to deceive or coerce those on the other end of a transaction or relationship: these are underlying reasons for taking informed consent seriously."[32]

O'Neill believes that many of the problems with informed consent arise because of its nature as a propositional attitude:

30 O'Neill (2003) p5
31 Ibid.
32 Ibid. p6

> consent is a propositional attitude, given in the first instance not to another's action, but to a proposition describing the action to be performed... Propositions may be more or less specific, and some limit has to be drawn to the amount of detail included. The inclusion of excessive or technical detail, for example, will eventually overtax even the most energetic, and undermine the possibility of informed consent. On the other hand, consent that is too vague and general may also fail to legitimate action.[33]

So rather than consenting to a treatment, a patient really consents to a proposition describing that treatment, and problems arise when the proposition that the patient has in mind and the one that the clinician has in mind differ in ways that the patient finds important. O'Neill cites an example a situation like that of the Alder Hey organs scandal, where the parents of a dead child consent to tissue being removed from their child's body, but object when they find out that whole organs have been taken. The clinician in this case knows that organs are composed of tissues, and believes that consent to the removal of tissues entails consent to the removal of organs. However, for the parents the distinction between tissues and organs is an important one, and consent to the removal of tissues does not imply consent to the removal of organs.[34]

So how can O'Neill's ideas apply to placebos? The challenge for a clinician is to present their patient with a proposition describing a treatment in sufficient detail that it is neither deceptive nor coercive, whilst still omitting the fact that the treatment is inert. If it is indeed possible to do this, and the patient then consents to this proposition, consent to the placebo treatment can legitimately be implied.

Is it possible for a clinician to non-deceptively present a proposition describing a placebo treatment, while leaving out the inert nature of that treatment? According to O'Neill, the

33 O'Neill (2003) p5
34 Ibid.p6

amount and level of information given should be dictated by the patient, rather than by the physician. So it is possible that some patients will judge the fact that their treatment is a placebo to be information material to them, while others will not. “Genuine consent”, she argues, “is apparent where patients can control the amount of information they receive, and what they allow to be done.”[35] This conclusion dovetails in an interesting way with the subjective standard of disclosure I have been promoting in this chapter: while the clinician must decide how much information to disclose in each case, each patient's desire for information must be a determining factor. The clinician still has a kind of gatekeeper role – as they are in possession of the knowledge about a treatment, and charged with disclosing an appropriate level of it. But allowing each patient's desire for information to determine the appropriate level of disclosure both respects the patient's autonomy, and reduces the burden the subjective standard can place on clinicians by making them the sole arbiter of the patient's informational needs.

This conclusion has much in common with a 1983 ruling from the Supreme Court of South Australia in the case of *F v R,*[36] which is cited in the ruling of *Rogers v Whitaker* I mentioned earlier. In *F v R*, it was ruled that that “the amount of information or advice which a careful and responsible doctor would disclose depended upon a complex of factors: the nature of the matter to be disclosed; the nature of the treatment; the desire of the patient for information; the temperament and health of the patient; and the general surrounding circumstances.”[37] This conclusion also has much in common with an approach that Ulrik Kihlbom has proposed, called negatively informed consent.[38]

35 Ibid.p6
36 King CJ (1983)
37 Ibid.p192-193
38 Kihlbom (2008)

Like O'Neill, Kihlbom argues that the doctrine of informed consent is not necessarily the best way to ensure that patient autonomy is respected. Rather, he argues, informed consent requirements force patients into a certain role – that of "informed consenter" – denying them the choice to be less informed, should they so wish. It is possible, he argues, to be informed of the aims of a treatment but autonomously choose not to know about the means, methods, or even the risks of it. In such a case, the patient has given what he calls "negatively informed consent":

> We might say that a patient has given a negatively informed consent when the patient:
> a. is competent, and
> b. has the capability of understanding the information
> c. has received information of:
> – 1 Purpose of the treatment
> – 2 That it is possible to receive more information if wanted
> – 2[sic]That the treatment is voluntary
> – 3 That the consent can be withdrawn at any time
> d. has well founded beliefs that the physician will choose the treatment that best promote his/her values
> e. has well founded beliefs that the physician will choose the treatment, the risks of which are in accordance with his/her attitudes towards different kinds of risks.
> f. on the basis of this gives his/her voluntary and explicit consent to undergo the treatment and express his/her voluntary and explicit wish not to have more information.[39]

The patient who gives negatively informed consent to not receive more information regarding the treatment, then, explicitly chooses to trust the physician to promote the best possible treatment. And this trust should be well-founded – for Kihlbom, "this rules out negative [informed consent] in situations where the physician and patient know little about each other."[40]

Even in a situation where a physician and patient know each other well, however, Kihlbom's assertion that information on the likely risks of a treatment is not needed for a patient to provide negatively-informed consent seems a little extreme. By comparison, the

39 Ibid.p147
40 Ibid. p146

patient must be told of the purpose of a treatment, and presumably this is so the patient can assess that that purpose is in agreement with their own ends. As Kihlbom says, “to exercise your autonomy, you do not need to know *how* your ends are realised, given that you have good grounds to believe that they will be realised.”[41] But if the patient is required to assess whether the purpose, which is to say the likely positive outcome, of a treatment is in accordance with their ends in order to give negatively informed consent, what is the justification for omitting the possible negative outcomes – which are sometimes fairly likely to occur? It is understandable that some patients might autonomously choose not to hear horror stories of what can happen if things go spectacularly wrong, but if certain negative outcomes or side-effects are not particularly rare, then these are things that patients undergoing such treatments should perhaps be expecting, or planning for, rather than merely worrying about. If a patient is to truly assess whether a treatment is in accordance with his ends, then he must also weigh up the effects the likely risks and side-effects of this treatment will have on those ends, and this will be impossible if those risks are not disclosed to him. For example, say that a patient is informed that the purpose of an antidepressant is to help with their mood swings, and he judges that this is concordant with his end of patching up his relationship. However a major side-effect of these antidepressants is to suppress libido, and this could be antithetical to that same end – the patient cannot judge whether or not the treatment is in accordance with his ends without knowing this.

Kihlbom might argue here that in order to give negatively informed consent in this case the patient in question must be able to trust that the physician would be aware of their end of patching up their relationship, realise that the risk of a suppressed libido would be unacceptable for them, and thus not prescribe the antidepressants. However it would

41 Ibid. p147

require a staggering level of insight on the part of a physician to be aware of all a patient's ends and the effects that the possible risks of a treatment could have on these – far more than is required by the subjective standard of disclosure. A close friend might be able to achieve this kind of insight into someone's ends, but it is unrealistic to expect it of a physician. It does not seem that a patient's autonomy is likely to be best served unless they are made aware of all the likely outcomes – both positive and negative – of a treatment, and can assess for themselves whether these outcomes are concordant with their ends.

While I think that the minimum standard of information required to protect a patient's autonomy is rather higher than Kihlbom believes, his wider assertions that more disclosure does not guarantee more autonomy, and that the decision to not be informed about some subjects can be an autonomous one, have important implications for the placebo case. The idea that it's possible to autonomously authorise a placebo treatment, whilst not knowing that it's a placebo, presupposes the possibility of autonomously choosing to remain uninformed about something – in this case, autonomously authorising the placebo entails authorising non-disclosure of it's inert nature. Kihlbom argues that choosing not to know something is not a choice to waive one's autonomy, rather it is an exercise of that autonomy. A non-medical example of this might be if I record a football match, and choose not to find out the final score before I watch it. I avoid the newspaper and news websites, shun all discussion of the match around the office, and leave the radio turned off. These are all autonomous choices to not be informed about something. Someone who informed me of the final score against my will would not be aiding my autonomy by informing me, rather they would be violating it. In this example the benefit I gain, the excitement of watching the match without knowing the outcome, is only gained if I stay uninformed of the

final score. In the case of a patient being given a placebo, the therapeutic benefit of the placebo effect is only gained if they remain uninformed of the exact nature of the treatment.

Of course, a patient who consents to a pill, but chooses not to know exactly what's in it, probably doesn't know whether or not disclosure of the nature of the pill would impact on the effect of the pill in the way that I know that disclosure of the final score will ruin the excitement of the game. A careful clinician might hint as much, however: D M Shaw, discussing negatively informed consent as a possible framework for the clinical use of placebos[42] , suggests a doctor could present a placebo pill to a patient as follows: “This pill has no side-effects, but studies have shown you that the more I tell you about how it works the less effective it will be.”[43] Nonetheless, the decision to not be informed about the nature of a treatment is not antithetical to autonomy in the way that a decision to not be informed about the likely risks and side-effects of a treatment is. If a patient decides to remain uninformed in this way, provided they comprehend the basic information required to autonomously authorise the treatment, then respect for their autonomy requires a clinician to refrain from further disclosure.

Conclusion

In this chapter I have argued that the likely risks and benefits of a placebo treatment vary from patient to patient, and the level of disclosure required for a patient to consent to a placebo treatment should likewise be assessed on a patient-by-patient basis. Attempts to

42 Shaw (2009)
43 Ibid. p98

establish an across-the-board standard of disclosure for the clinical use of placebos by reference to a professional practice standard or a hypothetical “reasonable patient” fail, and a categorical prohibition of the use of placebos in the name of patient autonomy, trust, and medical benefit is not be the best way to promote these aims. While it is true that placebo treatments are necessarily deceptive, and carry risks that are specific to this, a rule against all such treatments fails to take into account the cases that are neither definitely placebo nor active treatment, but may be classed as either, as well as the non-disclosed non-specific therapeutic component of all active treatments. Assessing the level of disclosure appropriate on a case-by-case basis does require a clinician to have a high level of familiarity with patients, but this level of familiarity is also helpful in determining the likely risks and benefits of a placebo treatment. Furthermore, cases where a clinician will not have the amount of time or interaction with a patient necessary to become familiar with them are simply unlikely to be cases where a placebo could be used effectively.

According to the proposed standard of informed consent, a clinician is tasked with determining for every patient the minimum level of understanding required for a patient to provide their consent to a placebo treatment. Above that minimum level, however, respecting a patient's autonomy means respecting their wish to know as much, or as little, as they want about a treatment. We therefore arrive at three ways in which a placebo treatment could be judged inappropriate. Firstly, a clinician could judge that, given a patient's history, personality, beliefs, and so on, there is a high risk of their reacting badly on discovering that they have been given a placebo treatment, and that this risk outweighs the likely therapeutic benefit of a placebo treatment. Secondly, a clinician may decide that a particular patient's informational needs are such that the minimum level of disclosure appropriate makes a placebo treatment unfeasible in this case. Thirdly, a clinician may

have presented what they believe to be an appropriate level of information on a treatment, but a patient's desire to know more means that further discussion is necessary. If this means that the nature of a placebo treatment is revealed, and that the possible therapeutic benefit of this treatment is thereby lost, then this is the price that must be paid for respecting that patient's autonomy. And again, the nature of the cases where a placebo treatment is likely to be under consideration means that this price is unlikely to be very high. Even where a placebo is judged to be the best available therapy, the loss of a therapeutic placebo effect is not going to be a life-or-death situation. In most cases where a placebo treatment may be under consideration, however, both the patient's autonomy and their therapeutic interests are best served by an open and mutual discussion.

References

American Medical Association (2014) *Code of Medical Ethics,* American Medical Association Press

Barnhill A (2011), “What it takes to defend deceptive placebo use” *Kennedy Institute of Ethics Journal* 21(3): 219-250

Beauchamp TL, Childress J (2009) *Principles of biomedical ethics (6th ed)*, New York: Oxford University Press.

BMJ (1996) “The Nuremberg Code (1947)” *British Medical Journal* 313: 1448

Cherkin DC et al (2009) “A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain” *Archives of Internal Medicine* 169: 858-66

Colloca L, Benedetti F (2005) “Placebos and painkillers: is mind as real as matter?” *National Review of Neuroscience* 6(7): 545-52

Faden RR, Beauchamp TL, King NMP (1986) *A history and theory of informed consent.* New York: Oxford University Press.

High Court of Australia (1992) "Rogers v Whitaker" *The Australian law journal* 67(1): 47-55

Jones WH (ed) (1923) *Hippocrates with an English translation*, Vol. II. London: William Heinemann

Kihlbom U (2008) “Autonomy and Negatively Informed Consent” *Journal of Medical Ethics* 34: 146-149

King CJ (1983) “F v R” 33 *SASR* 189 SASC

Kirsch et al (2008), “Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration” *PLoS Medicine* 5(2): e45.

Lichtenberg P, Heresco-Levy U, Nitzan U (2004) “The ethics of the placebo in clinical practice” *Journal of Medical Ethics* 30: 551–554

Manson NC, O'Neill O (2007) *Rethinking informed consent in bioethics.* Cambridge: Cambridge University Press

O'Neill O (1984) “Paternalism and partial autonomy” *Journal of Medical Ethics* 10: 173-178

O'Neill O (2003) “Some limits of informed consent” *Journal of Medical Ethics* 29: 4-7

Schneider C (1998), *The practice of autonomy: patients, doctors, and medical decisions.* New York: Oxford University Press

Shaw DM (2009) “Prescribing placebos ethically: the appeal of negatively informed

consent” *Journal of Medical Ethics* 35: 97-99

World Medical Association (1964) “The Declaration of Helsinki” *British Medical Journal,* 313: 1448-1449

World Medical Association (1989) “The Declaration of Helsinki” https://history.nih.gov/research/downloads/helsinki.pdf

World Medical Association (2000) "Helsinki Declaration VI" *Bulletin of Medical Ethics* 162: 8–11

World Medical Association (2008) “The Declaration of Helsinki” http://www.wma.net/e/policy/b3.htm

6. Trust, betrayal, and placebo use in clinical practice

A man he is of honesty and trust

To his conveyance I assign my wife

– Shakespeare, *Othello*

Introduction

A common objection to the use of placebos in clinical practice is that the deception involved in administering a placebo treatment is a betrayal of the trust between practitioner and patient, a trust which is an essential part of the medical relationship. A good example of this objection can be found in Sissela Bok's *Lying:*

> the giving of placebos is a waste of a very precious good: the trust on which so much in the medical relationship depends. The trust of those patients who find out they have been duped is lost, sometimes irretrievably. They may lose confidence in physicians and even in bona fide medication which they may need in future. They may obtain for themselves more harmful drugs or attach themselves to fad cures[1]

As mentioned previously, a similar argument against the use of deceptive placebos can be found in both the American Medical Association and the British Medical Association policies on the use of placebos in clinical practice.[2] [3] The same argument can be found in papers by Howard Brody[4] Beatrice Golomb[5], Richard Kanaan[6], and others.

1 Bok (1978) p63
2 AMA (2014) Opinion 8.083 - Placebo Use in Clinical Practice
3 BMA (2012) p565
4 Brody (1982) p114
5 Golomb (2009) p35
6 Kanaan (2009) p29

The idea that honesty and trustworthiness are closely linked, and that deception constitutes a failure of trust, is strongly intuitive. However, upon examination of several philosophical accounts of trust I find that it is not immediately clear why this is the case – the relationship between honesty, trust and trustworthiness is not something that has been examined in much depth. I therefore propose my own theory of the relationship between trustworthiness and honesty, and find that, while it is difficult to classify lies as a betrayal (outside of situations where honesty is explicitly promised), lies definitely constitute a failure of trust, or a failure to be trustworthy. This can be explained by reference to a virtue theory of trustworthiness, or by reference to Karen Jones' account of rich trustworthiness.

While lies constitute a failure to be trustworthy, the deception involved in a placebo treatment does not need to involve lying. By examining the connection between honesty and trustworthiness, I argue that it is possible to draw out some guidelines for administering deceptive placebos in a way that is consistent with a trusting medical relationship. This involves being, and being seen to be, as open and honest as possible, and withholding only that information necessary to facilitate the placebo effect. In addition, the information that was withheld, the nature of the placebo treatment, should be disclosed to the patient once the treatment has ceased, along with an explanation of why this treatment was thought to be the best option. In this way practitioners will show themselves to be trustworthy with regard to the patient's therapeutic needs, but also open and truthful where this does not conflict with those needs.

Placebos and the place of trust in medicine

When faced with the argument that a deceptive placebo constitutes a breach of trust in a medical relationship, we might well ask whether this trust is really necessary at all. Isn't the idea that patients should trust healthcare practitioners just a throwback to the bad old days, when patients were powerless and subservient to their godlike, patriarchal doctors? What use has an empowered, autonomous patient for trust, when the medical relationship is supposed to be seen as a partnership? Well, even a perfectly equal partnership has to involve some trust – contracts, vigilance, the threat of sanctions, and so on, can only go so far. And furthermore the nature of the medical relationship, whereby a patient's vulnerability and lack of knowledge puts them at an inevitable disadvantage, means that entering into such a relationship requires a lot of trust on the patient's part. As Merrilyn Walton argues, “Trust alone enables us to hand over responsibility for our own wellbeing and the wellbeing of others into the care of doctors. Without it we would not contemplate such an action”[7].

As well as being a necessary foundation of the medical relationship, trust between practitioners and patients can promote good outcomes through a number of instrumental benefits, neatly summed up here by Jessica Miller:

> well-placed patient trust enables good outcomes: patients who trust physicians are more likely to seek care at the first sign of trouble, to communicate important information about their symptoms and feelings, to reveal any fears or concerns that could sabotage treatment, to comply with medical recommendations, and in general, experience less anxiety around issues of health.[8]

7 Walton (1999) p17
8 Miller (2007) p52

In addition to these instrumental benefits, an atmosphere of trust between practitioner and patient may also directly affect clinical outcomes through mechanisms akin to the placebo effect[9]. Conversely, a loss of trust in the medical profession is not only detrimental to the relationship between the practitioner and patient involved, it can also have widespread negative consequences for society. An atmosphere of distrust in medicine is fertile ground for conspiracy theories and quackery to flourish: witness the current hysteria over vaccines, where doctors and medical organisations are accused of everything from covering up evidence of adverse reactions to using vaccines to introduce mind control agents into the population. Such beliefs are simply incongruous with trust in the medical profession, and their acceptance beyond the fringes of society, along with the resulting drop in vaccination rates and increase in the mortality and morbidity of vaccine-preventable diseases, can be directly attributed to a wider sense of distrust in medicine. Bok's objection is therefore a serious one: if a deceptive placebo constitutes a violation of the trust in a medical relationship, this loss of trust could have a detrimental effect on the patient in that particular relationship and also on society more widely. As Bok says, “even the health professionals who do not mislead their patients are injured by those who do; the entire institution of medicine is threatened by practices lacking in candour”[10].

If it is the case, then, that a deceptive placebo violates the trust between practitioner and patient, this is a powerful argument against placebo use in clinical practice. But does a deceptive placebo violate the trust between practitioner and patient? One way of answering this question would be to treat it as an empirical matter. The trust a patient has in their physician can be measured using the Trust in Physician Scale developed by Lynda Anderson and Robert Dedrick in 1990[11]. Measuring a patient's trust in this way involves

9 See for example Barnhill (2012)
10 Bok (1978) p68
11 Anderson and Dedrick (1990)

reading the patient a series of of 11 statements (for example, “I doubt that my doctor really cares about me as a person” and “I trust my doctor so much that I always try to follow his/her advice”) and asking the patient to rate how strongly they agree with each statement on a scale of one to five. From this we can obtain a numerical value representing the strength of that patient's trust in their physician. We could therefore design a trial where the trust of a group of patients is measured before their physician administers them a deceptive placebo, and once again after the nature of the placebo treatment has been revealed. These figures could be compared with those of a control group (who would probably be treated with active medication, in a reversal of the usual roles of placebo and active treatments in controlled trials), and if the trust value of the placebo group is found to have dropped by a statistically significant amount following the placebo then Sisella Bok is right, a deceptive placebo is detrimental to patient trust. Of course, carrying out such a study is beyond my means as a poor philosophy graduate student. But even were I to carry it out, and find that the result did support Bok's argument, I don't believe it would shed any light on the far more interesting question of *how* or *why* a deceptive placebo might violate patient trust. The study I have described would only give us an answer as to whether or not deceptive placebos as used in the study cause a loss of patient trust. An investigation of how a deceptive placebo might violate patient trust, on the other hand, might possibly show us what the trust-busting aspects of a placebo treatment are, and whether it might be possible to use a deceptive placebo in a way that doesn't violate patient trust.

How, then, does Sisella Bok believe a deceptive placebo violates patient trust? Unfortunately she doesn't tell us. Bok seems to assume the truth of this proposition that placebos destroy patient trust, rather than arguing for it. Granted, if you believe that a

placebo treatment is just an ineffective substitute for real medicine, then it may indeed seem obvious that duping a patient into taking a placebo is a betrayal of their trust. But as I have argued previously, we have good reason to believe that placebos can offer a real therapeutic benefit, and so in some situations a deceptive placebo treatment may offer the best chance, or indeed the only chance, of achieving some therapeutic benefit for the patient involved. If this is the case, then what is it about a deceptive placebo that constitutes a betrayal of that patient's trust?

Administering a deceptive placebo to a patient will inevitably involve some deception on the practitioner's part, and if we accept that a placebo can sometimes be an effective treatment, then it is the deception involved that looks like the most likely culprit for violating a patient's trust. By most accounts lying to someone is (at least *prima facie*) wrong – could this be where the betrayal lies? It is possible, but there are other ways healthcare professionals can wrong their patients – by overcharging them, perhaps, or making fun of them behind their backs, and these aren't widely cited as violations of trust. Is there something about deception, then, that is particularly toxic to a trust relationship? In order to examine this question, I will first need to look at the nature of trust and trustworthiness in general, and trust in the medical relationship more specifically.

Trust

What is trust, and how can it be betrayed? What differentiates trust from other social attitudes, and betrayal from other disappointments? Annette Baier[12] has written an influential account of trust, which I'd like to examine here. According to Baier, trust is

12 Baier (1986)

necessary because, as human beings, we inevitably need to rely on others to look after the things that we value in our lives. Trust, she argues, is the mechanism that allows us to do this in the absence of an absolute guarantee of another person's reliability:

> We trust those we encounter in lonely library stacks to be searching for books, not victims. We sometimes let ourselves fall asleep on trains or planes, trusting neighbouring strangers not to take advantage of our defencelessness. We put our bodily safety into the hands of pilots, drivers, doctors, with scarcely any sense of recklessness.[13]

For Baier, however, there must be more to trust than “mere” reliance. For trust is betrayed, whereas reliance is merely disappointed. And even where we do not trust someone, we may still rely on them to act in accordance with our wishes, perhaps in the hopes of personal gain or out of the fear of punishment. Baier defines trust as a particular kind of reliance: reliance on a person's goodwill towards one to motivate that person to carry out that which they are trusted to do. If I trust a friend to return something that they've borrowed in good condition, then I am counting on their goodwill towards me to motivate them to look after and return the item in question. On the other hand, if I am counting on not their good will towards me, but their fear of my anger to motivate them, then it would be incorrect to say that I trust them. Unless goodwill is involved, then the relationship in question is one of reliance, rather than trust:

> We can still rely where we no longer trust. What is the difference between trusting others and merely relying on them? It seems to be reliance on their good will toward one, as distinct from their dependable habits, or only on their dependably exhibited fear, anger, or other motives compatible with ill will toward one, or on motives not directed on one at all. We may rely on our fellows' fear of the newly appointed security guards in shops to deter them from injecting poison into the food on the shelves, once we have ceased to trust them.[14]

13 Ibid. p234
14 Ibid. p234

There is more to trust, however, than being able to count on another's goodwill towards one. One must also believe that the trusted person is competent to carry out that which they are trusted to do. I may trust the person who walks my dog, and be confident of her goodwill towards me, even to the extent that I give her the key to my house so she can let herself in when I am not home. But would I trust her to pilot the jumbo jet on which I am flying overseas? My reluctance to do so is not due to a suspicion of ill-will towards me on her part, merely a doubt that she is actually capable of looking after this particular good (in this case my bodily safety, when strapped into an economy class seat with a paperback and a gin and tonic), if entrusted with it. This gives us a three-place account of trust:

> Thus, there will be an answer not just to the question, Whom do you trust? but to the question, What do you trust to them? What good is it that they are in a position to take from you, or to injure? Accepting such an analysis, taking trust to be a three-place predicate (A trusts B with valued thing C) will involve some distortion and regimentation of some cases, where we may have to strain to discern any definite candidate for C, but I think it will prove more of a help than a hindrance.[15]

Trusting inevitably creates a vulnerability on the part of the truster. If I trust someone with a certain thing of value to me, then I place myself in a position where that person is able to harm that thing, either through a lack of goodwill towards me, or through well-willed but negligent or incompetent actions on their part. However, trusting also makes us vulnerable to the actions of people who, through miscommunication or misunderstanding, fail to grasp the limits of that which they were trusted with. Karen Jones, building on Baier's work, refers to the limits of that with which someone is entrusted as the “domain” over which that trust extends.[16] Overstepping this domain and taking on more than one was originally trusted with can also constitute untrustworthy behaviour, even though there may be no ill-

15 Ibid. p236
16 Jones (1996) p6

will behind the overstep. Baier gives the example of a babysitter who decides the nursery will look better if painted purple, and takes it upon himself to do so: “[He] will have acted, as a babysitter, in an untrustworthy way, however great his good will.”[17]

So when I trust someone to look after a certain thing, I am counting on their competence to do so and their goodwill towards me, as well as their judgement as to the limits of that with which they are entrusted. Within these limits, however, trusting someone to look after a certain thing entails allowing them a certain discretion as to how this is best carried out. The level of “discretionary power”, as Baier calls it, afforded to a trusted person varies widely according to the situation. A child entrusted with a simple task will generally have very little power to decide the way that task is done, whereas a surgeon carrying out a complex operation is entrusted with extensive powers to judge how best to carry out that operation and deal with any complications that arise. The limits of a trusted person's discretionary powers, like the limits of the domain over which that trust extends, are often not made explicit, but are a matter of judgement for the truster and the trusted, and are likewise susceptible to miscommunication and misunderstanding. But whereas, according to Baier, it is clear that a trusted person fails to be trustworthy when they overstep the domain of that trust and take it upon themselves to care for more than they were trusted with, it is less certain that exceeding one's discretionary powers constitutes a breach of trust:

> The more extensive the discretionary powers of the trusted, the less clear-cut will be the answer to the question of when trust is disappointed. The truster, who always needs good judgement to know whom to trust and how much discretion to give, will also have some scope for discretion in judging what should count as failing to meet trust, whether through incompetence, negligence, or ill-will.[18]

17 Baier (1986) p236
18 Ibid. p238

Say for example a friend borrows my car. Because he is my friend, I trust him to return it to me in good condition. If he sells the car and disappears with the money, he has betrayed my trust by acting with ill will towards me. If the car is stolen because he left it unlocked with the keys still in the ignition, then through this negligent or incompetent act my friend has also failed my trust. If my friend returns the car to me, but I find that he's modified the engine so that my car is now a gas-guzzling monster that rattles windows all through the neighbourhood, then my friend has acted in an untrustworthy manner by taking on more than I trusted him with. But what if my friend returns my car in good condition, but I learn that he's ensured the car was not stolen by sitting on the bonnet all night brandishing a shotgun, threatening anyone who came near? He has looked after the good I entrusted him to look after, and no more or less than that, however the means by which he did this go beyond what I consider reasonable in that situation. Is this a breach of trust, then? The true test of this, I suppose, is whether or not I would let my friend borrow my car again. And I think that I probably would – after clarifying what I consider reasonable behaviour in looking after the car, I'd probably still trust him with it (this is, of course, discounting the fact that for the purposes of this example I've made my friend sound like a gun-waving lunatic that nobody would want anything to do with).

Baier agrees that oversteps or abuses of discretionary power are not incompatible with further trust, speculating that malicious agents may try to disguise their deliberate violations of trust as well-meaning oversteps of their discretionary powers, and thereby retain our trust:

> Similarly with our surgeons and plumbers – *just* what they should do to put right what is wrong is something we must leave to them. Should they act incompetently, negligently, or deliberately against our interests, they may conceal these features of their activities from us by pretense that

> whatever happened occurred as a result of an honest and well-meaning exercise of the discretion given to them. This way they may retain our trust and so have opportunity to harm us yet further. In trusting them, we trust them to use their discretionary powers competently and nonmaliciously, and the latter includes not misleading us about how they have used them.[19]

So in Baier's analysis, then, person A trusts person B with (at least some of) the care of valued good C where:

- A believes that B is competent to take care of C
- A believes B's goodwill towards A will motivate B to take care of C
- A believes that B is aware they have been entrusted with the care of C[20]
- A affords B a certain discretionary power in taking care of C

How might this picture of trust apply to a medical relationship, and to the placebo case in particular? Given that we are examining the charge that deception is a betrayal of trust in a medical relationship, perhaps we can say that 'C', the good that patients trust their practitioners with, is the truth, or their autonomy – in short, that patients trust their practitioners to be honest. Although such a picture would make it clear how an act of deception would constitute a betrayal – it would be a direct failure to look after that which one has been trusted with – I do not believe it is an accurate picture of the medical relationship. People do not go to the doctor seeking autonomy, and truth is not the *telos* of the medical profession. Furthermore, while patients might expect honesty from their practitioners, or even count on it, it's not accurate to say they *trust* their practitioners to be honest, at least according to Baier's account of trust. If I say I trust someone to be honest, this implies that I count on them to tell me the truth out of some goodwill towards me, not

19 Ibid. p239-240
20 Jones (*op. cit.*) argues that A also has to believe that B will be directly motivated to act by this awareness (p6).

that I expect them to be honest as a matter of course. But this is not the case in a medical relationship – insofar as honesty is expected from practitioners, it is expected across the board, not as a matter of individual goodwill. Of course, it may very well still be wrong to be dishonest to patients who are expecting or relying on your honesty, but it would be inaccurate to describe this as a betrayal of trust. (I will examine this point in more detail in the next section of this chapter). I propose that the likely candidates for 'C', the good which patients count on their practitioners to take care of, are bodily or mental health or wellbeing, recovery from illness, freedom from pain, or the like. The actual object of patient trust will vary from case to case, however I think “wellbeing” is sufficiently broad to be used as a catch-all in this examination.

So say patient A trusts practitioner B with his wellbeing, believing in B's competence and goodwill and affording her certain discretionary powers to take care of that wellbeing, what are the ways in which practitioner B can betray this trust? According to Baier's account, B could betray A by failing to act in the interests of A's wellbeing due to ill-will, perhaps administering an experimental treatment intended more to further B's research than restore A to health. B could also betray A's trust by being incompetent or negligent in her duties, perhaps by failing to keep up with new treatments which would be more effective or less harmful to A. Another way B could betray A's trust is by taking a much broader view of “wellbeing” than A does, perhaps offering to set A up on a blind date with a friend or critiquing A's dress sense.

What about a case where a practitioner treats a patient with a deceptive placebo? This does not immediately seem to be a betrayal of the kinds mentioned above. As I have argued before, in some cases a placebo may be the best available treatment – a deceptive

placebo may be the only treatment with a chance of having a therapeutic effect, or it may be just as effective as the best available active treatment, whilst avoiding the specific side-effects of an active treatment. In such a case it would seem that administering a deceptive placebo could be consistent with looking after a patient's wellbeing, and not necessarily indicative of negligence, incompetence, or ill-will on the part of the practitioner, nor does it seem to overstep the boundaries of the domain the practitioner was entrusted with.

Without being negligent, incompetent, or ill-willed, however, the deception involved in a placebo treatment may well exceed the discretionary power a practitioner is afforded by their patient. One of the constraints that many patients will expect their practitioners' discretionary power to be limited by is respect for autonomy, and deception is generally thought to be incompatible with this (although I have argued in the previous chapters that placebo deception could be an exception to this rule). A patient who discovers that they have been treated with a deceptive placebo, then, may appreciate that their practitioner has attended to their wellbeing as best they could, but feel that they overstepped their discretionary powers in doing so. Is this overstep a betrayal of trust? Annette Baier doesn't seem to think that it is, or at least she seems to allow that it is a less serious matter than a betrayal due to negligence, incompetence, or ill-will. Indeed, Sisella Bok's original objection – that a deceptive placebo is a betrayal of trust that could cause patients to lose confidence not just in individual physicians but in the entire institution of medicine – implies a breach of trust much more serious than that suggested by a “well-meaning overstep of discretionary power”.

What about other cases where, in the interests of looking after a patient's wellbeing, their autonomy is violated? If, by Baier's account, all paternalistic acts are classified as merely

being “oversteps of discretionary power”, we may be tempted to draw the conclusion that this is simply not a very useful model of trust, at least as far as medical relationships are concerned. However, if we examine more severe cases of paternalism, those which we would usually want to cite as betrayals of patient trust, I think we can usually identify one or more features that allow us to do so. Take for example the case of Dax Cowart, who was admitted to hospital with severe burns over two thirds of his body following a gas explosion in 1973. Despite his repeated requests that he be let die, Cowart was made to undergo arduous but life-saving procedures over a course of ten months. During this time he was judged competent by a psychiatrist, but nonetheless he was treated against his will. He was also denied access to legal representation, by which he might have challenged his treatment in court. Eventually he recovered enough to be released from hospital, although he was left blind and without the use of his hands. Despite this he has since gone on to earn a degree in law, and he is now an outspoken advocate for patient's rights. He remains adamant that his physicians should have provided him with pain relief and let him die, rather than putting him through such a painful and protracted course of treatment.[21] Can Cowart's physicians be said to have betrayed his trust by saving his life in this paternalistic fashion? I believe they can. While they may have acted in what they believed were the interests of their patient's long-term wellbeing, in this case long-term wellbeing is not actually the good that that their patient was counting on them to look after. Cowart has subsequently argued that what he was counting on his physicians to provide, indeed what he begged them to provide, was pain relief while he died. His medical team, however, not only took charge of more than they were trusted with, and therefore failed to respect the limits of Cowart's trust, but in doing so they also failed to look after the good with which they were actually charged. This breach of trust therefore goes well beyond an “overstep in discretionary power”, and so the actions of this medical team can clearly be

21 Burt and Cowart (1998)

construed as a betrayal under Baier's account of trust.

While Baier's account of trust does not give *carte blanche* to all acts of paternalism, it has so far failed to tell us with any certainty whether or not a deceptive placebo constitutes a failure of trust. We can say that it is not a betrayal, in that it is not a failure to look after that which one was trusted with, or an attempt to look after more than one was trusted with, however it may be an overstep of discretionary powers, something which may or may not involve a loss of trust. A loss of trust is not the same thing as a betrayal, although the two are closely associated. Not every betrayal results in a loss of trust, and not every loss of trust involves a betrayal. For example, someone might choose to forgive a minor betrayal entirely, and find that their trust in the betrayer has not been shaken. On the other hand, someone who wrongly believes that they have been betrayed may totally lose trust in another person even though no betrayal has taken place.

A lie told in the interest of a patient's wellbeing may be, by this account, an overstep of discretionary power, but not a betrayal. However, this assertion – that there's no particular connection between lying and being untrustworthy – is a remarkable departure from the way honesty and trustworthiness are commonly considered. Indeed, in many of the discussions I read in the preparation of this chapter, the link between dishonesty and betrayal is taken to be so straightforward it doesn't even need spelling out. By contrast, in Baier's account of trust honesty is notable by its absence – Baier's paper contains no mention of the term “honesty” whatsoever. I think at this point a more thorough examination of the relationship between honesty and trustworthiness may shed some light on whether or not this particular “overstep of discretionary power” is likely to involve a loss of trust.

Truth, trust and trustworthiness

What is the connection between honesty and trustworthiness? The two terms are often mentioned in the same breath, much as they are in the lines from Othello with which I opened this chapter. However, honesty and trustworthiness are far from synonymous, and one is not necessarily implied by the other – a police officer trusted with an undercover assignment to infiltrate a criminal gang, for example, will have to be routinely dishonest in order to be trustworthy. And if friend A tells me a piece of gossip about friend B in confidence, and friend B asks me if friend A has told me this thing, I am faced with a choice between being honest or being trustworthy. Nonetheless, common wisdom tells us that these two qualities are closely linked.

In everyday speech, if I say that I trust someone to speak the truth and that I believe they are honest, then I have more or less repeated myself. This is not good enough for us here, however – according to Baier's account of trust these two statements mean quite different things. If I believe that what someone says is true because I believe them to be habitually honest, then it would be correct to say I am relying on them to speak the truth, but unless I believe they have some goodwill towards me to motivate this truth-telling, then it would be mistaken to say I trust them to do so. To borrow one of Baier's examples, we can compare the case of someone who is habitually honest to Immanuel Kant, whose morning walks were so regular that his neighbours could set their clocks by him. Although these neighbours may have come to rely on Kant to walk past their front window at the same time every morning to keep their clocks accurate, and they may well have felt disappointed had he slept in one morning, it would be wrong to say he had betrayed them by doing so. Similarly, if someone's reputation for honesty leads me to rely on the truth of what they

say, and I then find out they have misled me, I may well feel let down, but this is not a betrayal – unless I believe that some ill-will on this person's part has led them to lie to me. It is only if I believe that another person's goodwill towards me is going to motivate them to be truthful with me that I can say that I trust them to do so. A close friend may not be habitually honest, but I might nonetheless trust them to be truthful with me. Should they fail to do so, my betrayed reaction might be “How could you lie *to me?*”. In a similar situation involving someone not particularly close to me, but with a reputation for honesty, my disappointed reaction would be more along the lines of “How could *you* lie?”

Nonetheless, in the medical relationship as well as elsewhere, whether or not a person is seen to be honest appears to be a determining factor in the perception of their trustworthiness. In 1997 Bruce Campbell and David Thom carried out a series of focus groups on patients, where participants were asked to give an account of past instances that had positively or negatively affected their trust in a physician. From these accounts Campbell and Thom were able to identify nine broad categories of experience likely to affect patient trust, of which one was “Demonstrating honesty and respect for the patient”. Examples where physicians were perceived to have been dishonest were few, they noted, but these were “particularly detrimental to trust when they did occur.”[22] (Interestingly, there were no accounts where a display of honesty on a physician's part was said to *increase* trust – it appears that honesty is taken as a given, or perhaps only noticed in its absence.)

In *Truth, Trust and Medicine* Jennifer Jackson sketches out a link between honesty and trust by arguing that honesty is a prerequisite for the institution of promising, which in turn is a foundation of social co-operation and trust:

22 Campbell and Thom (1997) p173.

> The disposition to be truthful is a prerequisite for the device of promising to be workable. We could not sensibly rely on promises if we could not rely on people (in general) to be truthful (most especially, true to their word)... in short, truthfulness matters because because it is necessary to support trust and co-operation, without which social life would be radically impoverished – in Hobbes' memorable words: 'nasty, poor, brutish and short'".[23]

While this serves as a good explanation of how honesty is necessary for people to be able to rely on one another and co-operate, it stops short of specifically making the link with trust. Again, I can rely on someone who's generally honest to do what they say they're going to do in our co-operative endeavour without going so far as to trust them, or feeling betrayed if they fail to do so.

It is tempting to say that being trustworthy must at least involve being honest about one's trustworthiness as regards the domain over which one is trustworthy, but this is too simple – the thing a trustworthy undercover police officer is *most* likely to have to lie about is whether or not she is a trustworthy undercover police officer. However, recall the three-place account of trustworthiness: a person is trustworthy as regards a certain good or domain with respect to a certain person or people. Does being trustworthy necessarily entail being honest about one's trustworthiness, as regards the domain over which one is trustworthy, to the person or people to which one is trustworthy? I believe it does. A trustworthy undercover police officer can be dishonest about her capacity as a police officer to the criminal types that are her targets, but should she start being dishonest to her handlers or superiors on this topic, then her trustworthiness in this domain becomes suspect.

For a person to be trustworthy, she must have a way of indicating to the people who can

23 Jackson (2001) p30

trust her in a certain respect that this is the case. This communication is rarely in the form of an explicit “trust me!” statement – indeed, such statements usually indicate a failure to convey trustworthiness through more conventional channels – but nonetheless, in order to be trustworthy it is necessary to convey that trustworthiness to those who can trust us. This is what Karen Jones refers to as “rich trustworthiness”:

> Rich trustworthiness is two-place in structure: B is richly trustworthy with respect to A just in case (i) B is willing and able reliably to signal to A those domains in which B is competent and will take the fact that A is counting on her, were A to do so, to be a compelling reason for acting as counted on and (ii) there are at least some domains in which B will be responsive to the fact of A’s dependency in the manner specified in (i).[24]

If someone fails to signal their trustworthiness in this way, then this will make it more difficult for people to trust them, which in turn would make it more difficult for them to be trustworthy. Rich trustworthiness requires us to invite trust, and to do so honestly. It also requires us to be honest about our shortcomings, to help prevent trust from being placed in us unwisely:

> As finite agents, we have limited information and limited time to search for more. We want the competent who can be counted on in the ways we need to identify themselves, and we want those who are not up for a particular form of dependency, whether because they lack the competence or the inclination, to identify themselves before we count on them in ways that are apt to be disappointed. We want those we can trust regarding a particular domain to signal their trustworthiness to us, so we can work out where — and where not — to turn.[25]

So by this account, at least, there is an intimate connection between honesty and trustworthiness – in order to be trustworthy, one must at least be honest about the quality and extent of one's trustworthiness to those to whom one is trustworthy.

---

24 Jones (2012) p74
25 Ibid. p74

Where does this leave us on the question of honesty and trust in a medical relationship? For Jennifer Jackson, the connection between trustworthiness and honesty in a healthcare context is simple – patients need to be able to trust their practitioners to tell the truth. If a doctor tells a patient that it is necessary to take their full course of antibiotics, the patient needs to be able to trust that this is actually the case, otherwise there is nothing to compel them to keep taking the pills when they are feeling much better after only half the bottle. This, however, is too simplistic. As I argued before, construing the medical relationship as one where the patient trusts their practitioner to tell the truth makes it easy to explain how a deception is a betrayal of that trust, but it is far from obvious that this is an accurate picture of the relationship. Were I to ask a patient why they trust a certain practitioner with their wellbeing, they might well answer that this person is highly trained in looking after people's wellbeing, and furthermore that to do so is the very *raison d'etre* of their profession. Truth-telling, on the other hand, is not. Insofar as dishonesty interferes with the promotion of wellbeing, it can easily be construed as a betrayal of the trust patients have in their practitioners. In cases where a deception does not interfere with the promotion of wellbeing, however, it is far from obvious that this counts as a betrayal of patient trust. Simply adding truth-telling as another domain that patients trust their practitioners with at this point looks like an *ad hoc* solution to this problem, at least without further argument to back it up.

Jackson draws a distinction between honesty and truthfulness, where the former is a general duty to refrain from lying, and the latter is a role-dependent duty to inform. Healthcare professionals, she argues, have an absolute duty to be honest – as does everyone else. They also have a duty to be truthful, derived from their duty of care and the

duty to respect patients' autonomy, which varies from case to case. Confusion between the duty to be honest and the duty to be truthful has, she argues, led to a perception amongst practitioners that their duty to inform and their duty not to lie *both* vary from case to case, and this is something that needs rectifying:

> If I am right, the teaching that is needed should repudiate lying altogether. This is a teaching for all time. It is so because it is the only teaching that is adequate to sustain trust in the word of nurses and doctors. That trust is and always has been a necessary supporting pillar of the covenant between the patient and the health professional.[26]

Elsewhere she notes a conundrum: deception is said to destroy trust between doctors and patients. However, the principle of respect for autonomy, and its attendant duty to tell the truth, have only relatively recently been recognised as important for medicine, and doctors nonetheless managed to be trusted by their patients for centuries before this became commonplace. The answer, according to Jackson, again lies in the distinction between honesty and truthfulness, between lying and non-disclosure:

> [Beauchamp and Childress argue that] lying and failing to disclose information undermines trust. Those who opposed lying to patients in earlier times also made this claim – in regard to lying. But they did not see that failure to disclose information was necessarily subversive of trust. Why should it have been – unless disclosure had been promised?[27]

Here again an account of exactly why or how lying undermines practitioner-patient trust, (and why historical cases of non-disclosure did not) is missing. And this means that, assuming a practitioner hasn't tacitly promised not to lie, the important features that might differentiate the case of lying from a case of non-disclosure are left for us to guess. Moreover, it suggests that the answer to the conundrum of how the medical profession remained trustworthy in the pre-informed consent era is that, while they may have been

26 Jackson (2001) p157
27 Ibid. p25

less forthcoming with information than they are today, doctors refrained from lying. Perhaps my placebo-centric reading has left me with a somewhat biased view of medical history, but I don't think there's much evidence to say that this was the case.

Unless practitioners tacitly promise to be honest, then, we are still lacking an account of how telling a lie is a betrayal of trust. However, the following is what I would like to propose as an account of how dishonesty can undermine trust, in general and in a medical relationship more specifically. I agree with Karen Jones that judging the trustworthiness of others is a difficult task, and that we therefore look to the trustworthy to indicate themselves to us. While this process of indication is seldom as simple as stating “you can count on me me to do x”, it is nonetheless something that can be done honestly or dishonestly – a good confidence trickster will be as adept at mimicking the unspoken cues of the trustworthy as they are at glib talking. Often, then, the problem of judging a person's trustworthiness entails judging whether the signals they are sending about their trustworthiness are honest. Thus the problem of judging trustworthiness can collapse to one of judging honesty. Judging honesty, while not without its difficulties, is a simpler affair than judging trustworthiness, and is something most people are fairly comfortable with, having been doing it since they were children. And one of the key factors when considering the problem of another's honesty has to be whether or not that person has been honest in the past.

If I lie to another person, then, this may affect her overall judgement of my honesty, which in turn may affect her judgement of my trustworthiness. This in itself is enough to erode trust in a relationship, but following Karen Jones' account of rich trustworthiness we can propose that this lie, insofar as it is a failure to signal my trustworthiness to others, is in

itself a failure to be (richly) trustworthy. If a necessary part of trustworthiness is conveying our trustworthiness to others, then even if the lie we tell is not in itself a betrayal, if the act of lying nonetheless indicates to others that we are not trustworthy, then this is a failure to *be* trustworthy. In this way lying may be a failure of trustworthiness, and this may damage a trust relationship even if the lie was told in good faith or has little or nothing to do with the good or domain of the trust in question. To return to an earlier example: the fact that I do not trust my dog walker to pilot me in a jumbo jet has no bearing on whether or not I trust her to walk my dog. If, however, my dog walker boasts to me of her jumbo jet-flying prowess, and I know these claims to be completely fictitious, then I may find myself also looking askance at her more outwardly reasonable claims of dog-walking prowess. In this way the link between trustworthiness and honesty serves to connect the three-place account of trustworthiness, where a person's trustworthiness to a certain person in one domain has no bearing on their trustworthiness to a second person in another domain, with a more common sense account whereby mistrust is far more contagious, and can spill over from one domain or person to another.

How does this account apply to the example of the trustworthy undercover police officer, who needs to lie to infiltrate a criminal gang? Her lies, if exposed, will definitely affect how the people to whom she has lied judge her honesty. All the signals she has been sending about her trustworthiness as a criminal will be (rightly) judged to have been fraudulent by those who trusted her in this respect, and these people will therefore no longer consider her a trustworthy criminal. While it is tempting to say that her betrayals in the criminal domain will have revealed her to be a trustworthy police officer, it is entirely possible that those whom she betrayed will subsequently consider her to be untrustworthy in a wider sense. Should the betrayed criminals find themselves in need of a police officer in future,

they may avoid calling on this particular officer, lest her assurances of trustworthiness in this domain prove as false as her assurances of trustworthiness in the criminal domain did. However, assuming she was honest to the people to whom she was actually trustworthy, her colleagues and perhaps her family, the lies she told in the course of her work should not negatively impact their judgements of her trustworthiness. The signals she was sending were on the one hand fraudulent signals of trustworthiness as a criminal minion, but to those to whom she was actually trustworthy, these same signals marked her as a trustworthy undercover police officer. For those in the know, then, the lies she told to others may have counted as an honest signal of her trustworthiness in this domain.

While this account of honesty in trustworthiness is unable to endorse Jackson's assertion that all lies are a betrayal of trust, it does support her distinction between honesty and truth-telling in medicine – a case where someone is known to have lied will generally have a greater influence on future judgements of their honesty than a case where someone is known to have withheld information. This means that withholding information, especially doing so judiciously, is less likely to constitute a failure of trustworthiness than an outright lie. This distinction is important for the placebo case, because a deceptive placebo need not involve an outright lie. In cases where a placebo is administered dishonestly, for example where a patient is told that a sugar pill is an antibiotic, we can concede that this is a failure of trustworthiness on the part of the practitioner. However if a practitioner does not lie, but merely withholds the crucial information that the pill is a placebo, we may escape this objection. Jackson argues that if there is a more general admission that placebo treatments might be used, then the withholding of information involved in administering one could be consistent with a trusting medical relationship. She envisions a situation where:

> [Practitioners] could be open with patients about their *general* deceptive and secretive practices and policies. It could be made known, for example, that placebos are sometimes prescribed – and that patients can sometimes be helped by these. Patients knowing of this possibility could always choose to ask directly on this or that particular encounter if they wanted to know if they were being offered a placebo there and then. Then, since, as I have argued, doctors must not lie, the patient would have to be told: they then could not be helped in this way. But that would be their choice. Some patients might prefer not to know... Acknowledging the practice publicly would not give patients reason to mistrust their doctors – provided, of course, that doctors were also openly committed to not lying.[28]

This is a similar scenario to that proposed by the American Medical Association, whereby patients could consent to possibly being given a deceptive placebo at some time in the future. However, I argued in chapter three that such a disclosure could adversely affect patient expectations, and thereby weaken not only the placebo treatments on offer, but also the placebo-like effects that make up part of the therapeutic effect of any active treatments available. A blanket disclosure that deceptive placebos are sometimes used is therefore not in patients' best interests. Moreover, I argued in the previous chapter that this approach is also not the best way to respect patient autonomy. Rather than disclosing this information to all patients, regardless of whether or not they might actually want to know this, or whether this information might help them in their decision-making, I would argue that patient autonomy is best respected by tailoring the level of disclosure to the needs of each individual patient. If a patient is judged likely to react badly to the suggestion of a placebo treatment, then deceptive placebo treatments should be off the table – but so also should the disclosure that they are a possibility. If, on the other hand, some patients “prefer not to know”, then their therapeutic needs, as well as their informational needs, may well be best met by withholding the general information that placebo treatments may be administered, as well as the specific information that an individual treatment is a placebo.

28 Ibid. p104

But in cases where this withholding of information is judged appropriate, how can it be done in a way that is consistent with trust in a medical relationship? Firstly, it is essential that a practitioner avoids lying to their patient. The discovery that a practitioner has lied could severely affect the judgement of her trustworthiness, which would be toxic to the trust relationship she have with her patient. Withholding information is more innocuous in this regard, but it should nonetheless be kept to a minimum. A patient should be told as much about the treatment as they wish to know, only withholding that information necessary to facilitate the placebo effect, and even this information should only be withheld where it is apparent that the patient would not wish to know it. Withholding this information from a patient who does wish to know it is inappropriate on autonomy grounds, but this level of deception might also affect judgements of a practitioner's trustworthiness. And while disclosing the nature of a placebo treatment removes the possibility of therapeutic placebo effect, the potential harm that could come from a loss of patient trust far outweighs the potential benefit this presents, so the patient's therapeutic interests are best served by this disclosure.

Lastly, I would suggest that where it is judged appropriate, the nature of the placebo should be carefully revealed following the cessation of treatment. This revelation should be accompanied by a discussion of the reasons why the practitioner judged that a placebo treatment was the best option in this situation, the nature of the placebo effect and the patient's own experience and outcome of this particular treatment. In this way practitioners will show themselves to be not only trustworthy with regard to the patient's therapeutic needs, but open and truthful where this does not conflict with those needs. Both of these things should serve to signal a practitioner's trustworthiness in this area, in the way

required to be richly trustworthy.

The assertion here that placebo treatments should be retrospectively revealed to patients might prompt an objection: isn't the revelation that a placebo treatment *has* been used essentially the same as disclosing that placebo treatments *might* be used in future? And I have argued that this latter disclosure is not in patients' best interests, as it could affect patient expectations and so lessen the effectiveness of placebo and active treatments both. However, these two disclosures are not equal in this respect. Disclosing, or asking a patient to consent to, the possibility of a placebo treatment in the future will serve to colour that patient's expectations before any treatment has even taken place. While the revelation of a placebo following the cessation of treatment will be one of the factors that shapes expectations of future interventions, in this case the patient will also have his own experiences of the treatment to draw on. Where the revelation of the placebo is accompanied by a discussion of those experiences, the nature of the placebo effect, the justification for the treatment and the like, patient expectations of future treatments, whether inert or active, can hopefully be maintained.

Trustworthiness as a Virtue

I have so far examined what are called “three-place” accounts of trust, whereby trusting someone involves relying on their goodwill and competence to look after a certain thing of value. However, according to a virtue account judging someone's trustworthiness involves more than a judgement of someone's goodwill and competence in a certain domain, instead it is a judgement of that person's character. By such an account, medical

practitioners foster patient trust not just by acting with goodwill towards them, but by having a trustworthy character. Many works of bioethics have cited the importance for medical practitioners of cultivating a trustworthy character – Beauchamp and Childress[29], for example, include trustworthiness as one of the "focal virtues" a physician should cultivate, while Pelligrino and Thomasma[30] cite "fidelity to trust" as one of the key virtues necessary for medical practice.

In *How Can I Be Trusted?*, Nancy Nyquist Potter combines a three-place account of trust with a virtue account of trustworthiness. Trusting someone, she argues, involves an expectation or belief that the trusted person has good intention with regard to the care of something we value, and the ability to look after it in the way that is expected of them. However, looking after a certain thing that another person values in the way that they expect is not sufficient for someone to be judged as having a trustworthy character:

> What many of us really want to know – at least some of the time – is not just, Can I trust this person to put up flyers about the meeting next week, but: Could this person still be counted on to be trustworthy if she were backed into a corner? When faced with a conflict of loyalties, would that person still be trustworthy? When the wolf comes knocking at her door, what will she do? When the wolf comes knocking at *my* door, what will she do?[31]

Potter therefore proposes that a trustworthy person is "one who can be counted on, as a matter of the sort of person he or she is, to take care of those things that others entrust to one and (following the Doctrine of the Mean) whose ways of caring are neither excessive nor deficient."[32]

29 Beauchamp and Childress (2008)
30 Pellegrino and Thomasma (1993)
31 Potter (2002) p13
32 Ibid. p16

So we are left with two definitions of trustworthiness: trustworthiness in the specific sense and trustworthiness in the general sense. Being specifically trustworthy is being trustworthy with respect to a certain person and a certain good, and this may have little or no bearing on your trustworthiness relative to other people or goods. Being trustworthy in the more general sense, however, involves having a trustworthy character, what Potter refers to as being "fully trustworthy":

> [F]ull trustworthiness requires much more than that one be trustworthy in the specific sense. The second sense, then, is that of being fully trustworthy. Being trustworthy in the more general sense involves being the sort of person who can be counted on, given who one is in relation to diverse others, to have the right feelings toward the right sort of things, to deliberate and make choices, and to act from a trustworthy disposition.[33]

Unlike Karen Jones' account of rich trustworthiness, Potter's definition of "full trustworthiness" does not include the requirement to invite trust or signal one's trustworthiness. This is a problem for my theory of the connection between trustworthiness and honesty, which is entirely based on this requirement to signal one's trustworthiness. By omitting this requirement, however, Potter's definition raises the problem of unwanted or uninvited trust – if someone unreasonably trusts me to do something that I am not willing or able to do, and I fail to do it, does this make me untrustworthy? For example, a friend might trust me to help him rob a bank: he might believe that I am capable of helping him rob a bank, that my goodwill towards him will motivate me to do so, and that my generally trustworthy disposition means that I can be counted on. However, I may not have given him any reason to believe that I would be willing to risk getting shot in order to earn him some money, and I may have even tried to warn him off trusting me with this task. But according to this account of trustworthiness my failure to help him would still mean that I am untrustworthy in this respect.

33 Ibid.p25

Intuitively it does not seem that being untrustworthy in this particular respect, as regards uninvited trust to carry out dangerous criminal activity, should affect judgements of whether or not I am “the sort of person that can be counted on”, as Potter says, or whether I have a generally trustworthy disposition. But unfortunately Potter's theory doesn't give us a way to differentiate between this failure of trust and the kind that would constitute a stain on my character. According to a rich trustworthiness account, by contrast, my failure to help my friend would only be untrustworthy if I had previously signalled to him that I *was* trustworthy in this respect.

How does Potter's account of trustworthiness apply to a medical relationship? It would suggest that patients, as well as counting on their medical practitioners to look after a certain valued good with goodwill towards them, also count on them to be of a trustworthy character. Pelligrino and Thomasma argue along similar lines, stating that “the intimacy, specificity, and personal nature of relationships with physicians compel us to be concerned with personal qualities – with personality, but most of all with character.”[34] By this account, a deceptive placebo could be argued to be a failure of trust if the deception involved was judged to be a stain on the practitioner's character. Much like the way judgement of a person's honesty affects their rich trustworthiness, if the deception involved in a placebo treatment is incompatible with good character, then a patients' trust in their practitioner may well be disappointed – even if that deception was carried out with good will as part of carrying out that which the practitioner was trusted to do.

Potter includes an interesting case of deception and betrayal from her experience as a crisis counsellor, where she felt she may have betrayed a client despite acting in the best

34 Pellegrino and Thomasma (1993) p68

interests of her wellbeing. One night a young woman rang up Potter's crisis line, and over the course of the conversation it became apparent that she had taken an overdose of pills. Judging her client's life to be in danger, Potter arranged for the phone call to be traced so that an ambulance could be sent. However, during their conversation the client demanded to know if the phone call was being traced, stating that she would hang up if it was. Potter assured her it wasn't, and

> the phone call was traced, the ambulance was dispatched, and she was transported to a hospital where her life was saved. However, when the ambulance arrived and she learned that I had acted against her express wishes and in addition had lied to her, she indicated that she felt betrayed, and I wondered whether, in an important sense, I had in fact failed to be worthy of her trust.[35]

Potter is in no doubt that her actions were justified – that telling a lie was the only option she had to save the young woman's life, and that saving a life more than outweighs the *prima facie* negative value lying has. Moreover, she argues that her actions would stand up to public scrutiny (as Sisella Bok argues that moral actions must), and were well within the discretionary powers she was afforded as a crisis counsellor – although she notes that in this situation her discretionary powers are afforded to her by the institution she is part of, rather than by her client. Nonetheless, the question of whether she had been trustworthy to her client remains:

> no matter how justified one is in having lied to a patient or client, by any norm for rationality, other moral concerns about the harms involved in broken trust may remain. For example, even if the patient or client were to come to believe that the lie was justified, she may not be able to trust the practitioner again... My point, then, is that questions of trustworthiness do not reduce to questions of justification for what one has done.[36]

35 Potter (2002) p43
36 Ibid. p50

Even if the client were to come to accept the justification for the lie, Potter argues, she may still rightly feel betrayed. But if the client accepts that Potter acted in her best interests and was justified in doing so, how exactly did Potter betray her? How might a lie, told to save a life, betray the trust of the person whose life it saves? I've speculated that, by a virtue account, dishonesty might affect a person's trustworthiness by being a stain on their character, but a lie told to save a life would surely not stain one's character – to the contrary, remaining truthful at the cost of another's life would surely indicate a vice of excess. This is not what Potter argues in any case. The lie is a betrayal, she argues, because “[t]he client, at the very least, expects and trusts that the counselor will be honest with her”[37]. But I have already argued that such an account of trustworthiness and honesty is problematic. How, then, did Potter's client trust her, and how was this trust betrayed by Potter's lie?

Given that Potter's client was a suicidal teenager calling a crisis line, it is tempting to argue that on some level she must have actually wanted Potter's help, that she was therefore trusting Potter with her long-term wellbeing, that and by acting in the interests of her client's wellbeing Potter therefore did not betray her trust. However for the purposes of this discussion let's assume that Potter's client's purpose in calling the crisis line was just as she stated, to have someone to talk to in her final hours, and that she was therefore only trusting Potter with her short-term emotional needs, rather than her long-term wellbeing. By looking after her long-term wellbeing Potter was taking on more than her client trusted her with, and therefore we can find that she acted in an untrustworthy manner, in the three-place sense. But this would still have been the case had Potter not had to lie – had her client never asked whether she was tracing the call, running the trace and sending the ambulance would still have been a betrayal of trust in the three-place sense. And it is not

37 Ibid. p52-53

tracing the call and sending the ambulance that Potter argues was a betrayal, it is the lie that she had to tell to do it. I think a rich trustworthiness account can show us why this is the case. Potter's client trusted her in a way that was uninvited: presumably neither Potter nor her organisation had given assurances in the past that this was the kind of place where you could find people who can be trusted to listen passively while you die at your own hand, so her client had no grounds for believing her to be trustworthy in this way. Nonetheless, this is how Potter's client trusted her. And once it became apparent to Potter that her client was trusting her in a way in which she wasn't trustworthy, then the (richly) trustworthy thing for Potter to do would have been to disabuse her of this notion. However, Potter did not do this. In order to maintain her client's false trust in her, and thereby save her life, she told a lie. This is how Potter's lie was a betrayal of trust – she lied in order to give the impression that she was trustworthy in a way that she wasn't.

The above analysis suggests that there were two trustworthy courses of action available to Potter: she could either act in accordance with her client's trust, or she could tell her client that she was not trustworthy in the way her client thought she was. But given that the likely result of either option was her client's death, I would argue that these would be an even greater failure of trust on Potter's part. Again, a rich trustworthiness account can show us how. While Potter's client in this case rang the crisis line with unfounded hopes of finding someone she could trust to provide her with short-term emotional support at the expense of her long-term wellbeing, many other clients who ring the crisis line have well-founded expectations that Potter will be trustworthy with respect to their long-term wellbeing. Potter's position as a healthcare professional requires her to be trustworthy with respect to her clients' health and wellbeing: her trustworthiness in this area would have been well supported by assurances made by Potter, her institution, her certifying body, and the like.

Even if this particular client was not trusting Potter with her wellbeing, to let the client die would be to effectively communicate that Potter was no longer trustworthy in this particular domain, which would constitute a massive failure of trustworthiness for a healthcare professional.

Granted, in this case Potter has shown herself to be untrustworthy in one specific domain, and has given one false assurance of her trustworthiness in that domain. But apart from that one occasion, no assurances she has given in the past, and nothing about her position as a healthcare professional, suggest that she *should* be trustworthy in this way. This specific failure of trust, then, has far less serious implications for Potter's trustworthiness than showing herself to be untrustworthy with respect to her clients' wellbeing, which would suggest that her past assurances of her trustworthiness in this domain (along with those of her institution, her school, and the medical establishment more widely) were false.

There are parallels between Potter's case study and the case of Dax Cowart, which I discussed previously. In both cases a life was on the line, and in both cases a betrayal was necessary to save it. However, voluntary euthanasia and end-of-life decisions are special cases, where the failure of a practitioner to be act in the interest of a patient's long-term wellbeing may reinforce, rather than erode, their trustworthiness as a healthcare professional. In addition, this is the case only under very strict circumstances. Cowart was judged to be mentally competent, and he made repeated and consistent requests to be let die over a period of time, as his only other alternative was suffering extreme pain. Potter's client was a suicidal teenager calling a crisis line, and Potter simply did not have the time or the means to judge whether her attempt to die was a voluntary, competent or well-

considered act. Given this situation, Potter's position as a crisis counsellor, and the fact that her client had reached out to her, Potter would have failed to be trustworthy in a far more serious manner than she actually did had she simply listened passively and let her client die.

This, then, seems like an odd case to cite in an examination of trustworthiness. Potter was faced with a choice between betraying her client by telling her a lie to save her life, or failing to be trustworthy in a far worse fashion by letting her client die. The course of action she took was morally justified, and I argue that it is the least untrustworthy option, but it appears there was no trustworthy course of action to take. We might therefore be tempted to say that in this case one should simply disregard the question of trustworthiness, and do what is morally justifiable. This, however, is precisely what Potter argues against:

> Within moral systems where right action is taken to be primary, the counselor might be assured that her moral integrity has not been compromised by her having lied to her client. But nagging doubts and worries about broken bonds of trust call for a richer analysis of lying and institutional discretionary power than the process of justification seems to suggest or most theories of right action seem to provide.[38]

Unless measures are taken to repair broken trust, she argues, then the client who has been betrayed in this fashion may not be able to trust that counsellor again, even if they come to accept the justification for the betrayal and believe the counsellor was right to do so. In order to be trustworthy, a counsellor would have to "feel regret for having betrayed her client's trust and would have to do what she can to restore that trust".[39]

So how could Potter go about repairing the broken trust between her and her client in this

38 Ibid. p51
39 Ibid. p55

situation? She has lied to her client in order to save her life, and although her client realises she was justified in doing this, she still feels that her counsellor has proven herself to be untrustworthy. In the specific sense, of course, Potter remains untrustworthy – her client in this case trusted Potter in a way that she did not invite, and her client expected her to be trustworthy in a way that she could not be. And furthermore, should Potter be in a situation in future where she must lie to this person in order to save her life, then this is presumably exactly what she would do. The particular trust that was broken, then, could not be rebuilt honestly. However, discussing the situation with her client, and giving an honest indication of those areas in which she *can* be expected to be trustworthy, might go some way towards rebuilding an impression of her trustworthiness more generally. If Potter shows that she regrets having to lie, and that she is generally honest about her capacities and shortcomings, this will make her assurances of trustworthiness ring more truly.

Unlike Jennifer Jackson, Potter does does not argue that all lies in a healthcare context are unjustified, but she has shown the way that even a justified lie can undermine trust. This recognition is important when considering the case of a deceptive placebo: I can argue that deception involving withholding information is justifiable where an outright lie is not, but we have to consider the possibility that even a well justified deception may result in a loss of trust. And while a life-or-death situation like Potter's case might mean that concerns of personal trustworthiness are outweighed by the greater trustworthiness that is tied with one's role obligations, a deceptive placebo is never going to be a case of life or death. However we have also seen that practitioners can take steps to repair trust that is broken by justified deception.

I have argued that a practitioner could administer a deceptive placebo in a trustworthy

manner if they avoid lying, are as forthcoming with information as the patient wishes, and where appropriate they reveal the nature of the placebo following the cessation of treatment. With a careful explanation of the nature of placebos and the placebo effect, along with an account of why one was necessary in this particular case, and a discussion of what (if any) effect the treatment had had on the patient, hopefully an understanding can be reached. However, we must acknowledge the possibility that even if a patient understands that a placebo treatment was the best course of action for their practitioner to take, they will nonetheless feel deceived, and may lose trust in their practitioner. In this case it is incumbent on the practitioner to take steps to repair this trust, to show that they are trustworthy, with regards to the patient's wellbeing and in the more general sense, and that they are honest and as forthcoming with information as possible.

Trusting patients: the other side of practitioner-patient trust

So far I have been discussing trust in a medical relationship entirely from one point of view: it is necessary for patients to trust their practitioners, and this trust is something that can be be gained or lost by the practitioners involved. However, trust in a medical relationship needs to flow in both directions. A practitioner must trust that her patient will adhere (more or less, perhaps) to the regimen she prescribes, just as the patient needs to trust her advice. If a practitioner suspects, for example, that a patient is faking symptoms in order to gain access to pills that can be sold at a high profit, this mistrust can be as corrosive to the doctor-patient relationship as patient mistrust can. Jennifer Miller describes the way being the object of mistrust can affect a patient:

> being the object of distrust, whether deserved or not, has a corrosive effect on one’s moral self-

> esteem, such that it can become a self-fulfilling attitude, in which an agent decides that there is no use in trying to communicate openly and truthfully with his physician when everything he says is liable to misinterpretation. And this effect may be especially pronounced in cases where the distruster has the moral authority and power of a physician.[40]

How is this perspective relevant to the placebo case? One of the most well-known (and still common) uses for deceptive placebos in clinical practice is to mollify difficult patients or hypochondriacs. And there is a common perception that, as a placebo is not a "real" treatment, anything treated with one is not a "real" condition. Either of these perceptions could lead a patient who learns that he has been given a placebo to conclude that his practitioner thought he was presenting with an imaginary condition, that the symptoms and discomforts he had described were judged to be all in his head, or even outright lies. In short, learning that that he has been given a placebo treatment may well lead a patient to conclude that his practitioner did not trust him. In such a case, the corrosive effect that the experience of distrust has is aggravated by the knowledge that this is unfairly so. The belief that one has been unfairly distrusted, especially in a situation where one is already at a disadvantage, is an intensely unpleasant experience, and can incite feelings of anger, shame, and resentment. Such an experience could plausibly make someone more reluctant to seek out medical help in future, especially from the practitioner or institution which, in their estimation, may well have labelled them a hypochondriac or "problem patient".

As it stands, some of the patients who are given placebo treatments are justified in thinking their practitioner does not trust them. In our ideal placebo case, however, where a placebo is administered because the placebo effect presents the best chance of achieving a therapeutic benefit, a patient would be mistaken in this belief. Nonetheless, in this case

40 Miller (2007) p59

the perception that a failure of trust has taken place can be just as damaging to a medical relationship as a real failure of trust. So what steps can be taken to assure patients who have been given a placebo that their illness is being taken seriously, and to avoid the risk of patients mistakenly concluding that their practitioner mistrusts them? Many of the suggestions that have been already made to reinforce patient trust can be helpful here. If the patient is given as much information as appropriate about the nature of their condition, the alternative treatments (or lack thereof) and the likely outcomes of the treatment, this will reduce the perception that they are not being taken seriously, or are being treated like they don't have a “real” illness. Where appropriate the placebo should be carefully revealed to the patient following the treatment, and they should be provided with an explanation as to why this was the best treatment at the time. This should help with the perception that they have not been given “real” treatment. A discussion of the patient's experiences of the treatment and any outcomes could also help confirm that these are not being treated as merely figments of the imagination.

Conclusion: deception, betrayal, and the loss of trust

In an article discussing his struggle with alcoholism, the great film critic Roger Ebert describes the physician who helped him overcome his addiction:

> What a good doctor, and a good man, Jakob Schlichter was. He was in one of those classic office buildings in the Loop, filled with dentists and jewelers. He was a gifted general practitioner. An appointment lasted an hour. The first half hour was devoted to conversation. He had a thick Physician's Drug Reference on his desk, and liked to pat it. "There are 12 drugs in there," he said, "that we know work for sure. The best one is aspirin."[41]

41 Roger Ebert, “My Name is Roger, and I'm an Alcoholic”, Chicago Sun-Times August 25, 2009

Although he does not specifically mention it, it is obvious that he trusts his physician deeply. And this trust is due not just to Dr Schlichter's skill as a physician but also to his being a “good man”, something that is immediately illustrated by a remarkable example of honesty – the doctor's candour about the limits of medical science. In this chapter I have largely been searching for an account of how the reverse position might come about – how a lie, even one told in a patient's best interests, could cause a patient to lose trust in a medical practitioner. Although it seems to be generally assumed that the connection between lies and betrayal is straightforward, I have not found this to be the case. Nonetheless, I have identified two general scenarios where a lie can clearly be described as a betrayal of trust. The first, which I do not examine in any detail, is where telling a lie means breaking a promise to be truthful. The second involves telling a lie to someone who trusts you, in order to mislead them about your trustworthiness relative to the domain in which they trust you. Neither of these scenarios describes the case of a deceptive placebo administered with the intention of provoking a therapeutic placebo effect, so I conclude that the deception involved in a placebo treatment is not necessarily a betrayal of patient trust – at least according to the accounts of trust I have examined here.

While a deceptive placebo treatment might not be a betrayal, it may still lead to a loss of trust. I have described a way in which, if a deceptive placebo treatment was seen to be dishonest, this could result in a loss of trust in a practitioner. As judgements of honesty are closely intertwined with judgements of trustworthiness, dishonest acts may mark a person as untrustworthy, and this could make others more reluctant to trust that person in future. And following Karen Jones' account of rich trustworthiness, a failure to communicate one's trustworthiness also constitutes a failure to be trustworthy. In this way an otherwise trustworthy practitioner, if seen to be dishonest, could lose her patients' trust and in this

way become untrustworthy.

In order to prevent such a loss of patient trust, then, efforts must be made to ensure that a deceptive placebo is as honest as possible, and is seen to be so. This means that lying in order to administer a placebo treatment, for example telling a patient that a sugar pill is an antibiotic, is inappropriate. Withholding information on a placebo treatment, on the other hand, may be less problematic, but only if steps are taken to preserve patient trust. As much information as appropriate should be disclosed about the patient's condition, the treatment, any alternatives, and the possible outcomes – including, if they wish (but only if they wish) the nature of the placebo. Revealing the nature of a placebo treatment in this way will mean that it is lost as a possible treatment option, but the prospect of therapeutic benefit from a placebo treatment is far outweighed by the ill effects that a loss of patient trust could have. If, on the other hand, the nature of the placebo can be withheld, the dishonesty associated with this practice is removed, and the chance that a patient who learns that they have been given a placebo will feel duped, unfairly treated, or distrusted is minimised. Finally, patients should be carefully debriefed as soon as possible following the cessation of the treatment. This will not only minimise the deception involved in a placebo treatment, it will also give patients a chance to bring up any misgivings or ill feelings they have about the treatment, so that these can be discussed in an honest and respectful manner. In this way, deceptive placebos may be employed in a manner that is consistent with retaining the trust on which the medical relationship depends.

References

American Medical Association (2014) *Code of Medical Ethics,* American Medical Association Press

Anderson LA, Dedrick RF (1990) “Development of the Trust in Physician Scale: a Measure to Assess Interpersonal Trust in Patient-physician Relationships” *Psychological Reports* 67(3): 1091-1100

Baier, A (1986) “Trust and Antitrust” *Ethics* 96(2): 231-260

Baier, A (1990) “Why Honesty Is a Hard Virtue” in Flanagan OJ, Rorty A (eds) *Identity, Character, and Morality.* Cambridge: MIT Press 259-282

Baier, A (1992) “Trusting People” *Philosophical Perspectives* 6: 137-153

Barnhill, A (2012) “Clinical Use of Placebos: Still the Physician”s Prerogative?” *Hastings Center Report* 42(3): 29-37

Beauchamp TL, Childress J (2009) *Principles of Biomedical Ethics (6th ed)*, New York: Oxford University Press

Bok S (1978) *Lying: Moral Choice in Public and Private Life* New York: Pantheon Books

British Medical Association (2012) *Medical Ethics Today: the BMA's Handbook of Ethics*

and Law (3rd ed.) Chichester: Wiley-Blackwell, BMJ Books

Brody H (1982) “The Lie that Heals: The Ethics of Giving Placebos” *Annals of Internal Medicine* 97: 112-118

Burt R, Cowart D (1998) “Confronting Death: Who Chooses, Who Controls?” *The Hastings Center Report* 28(1): 14-24

Campbell B, Thom DH (1997) “Patient-physician Trust: An Exploratory Study” *Journal of Family Practice* 44(2): 169-176

Golomb BA (2009) “Doctoring the Evidence: The case against lying to patients about placebos” *American Journal of Bioethics* 9: 34-6.

Goold SD (2001) “Trust and the Ethics of Health Care Institutions” *Hastings Center Report* 31(6): 26-33

Hall MA, Dugan E, Zheng B, Mishra AK (2001) “Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter?” *Milbank Quarterly* 79(4): 613-639

Hardin R (2006) *Trust.* Cambridge: Polity

Holton R (1994) “Deciding to Trust, Coming to Believe” *Australasian Journal of Philosophy* 72(1): 63-76

Jacobs AK (2005) “Rebuilding an Enduring Trust in Medicine: A Global Mandate: Presidential Address American Heart Association Scientific Sessions 2004” *Circulation* 111(25): 3494-3498

Jones K (1996) “Trust as an Affective Attitude” *Ethics* 107(1): 4-25

Jones K (2012) “Trustworthiness” *Ethics* 123 (1): 61-85

Kanaan R (2009) “When Doctors Deceive” *American Journal of Bioethics*, 9: 29-30

Miller J (2007) “The Other Side of Trust in Health Care: Prescribing Drugs with the Potential for Abuse” *Bioethics* 21(1): 51-60

Nickel PJ (2007) “Trust and Obligation-ascription” *Ethical Theory and Moral Practice* 10(3): 309-319

O'Neill O (2002) *A Question of Trust: The BBC Reith Lectures 2002.* Cambridge: Cambridge University Press

Pearson SD, Raeke LH (2000) “Patients' Trust in Physicians: Many Theories, Few Measures, and Little Data” *Journal of General Internal Medicine* 15(7): 509-513

Pellegrino ED, Thomasma DC (1993) *The Virtues in Medical Practice.* New York: Oxford University Press

Potter NN (2002) *How Can I Be Trusted?: A Virtue Theory of Trustworthiness.* Lanham: Rowman & Littlefield

Scanlon T (1990) “Promises and Practices” *Philosophy & Public Affairs* 19(3): 199-226

Skirbekk H, Middelthon A, Hjortdahl P, Finset A (2011) “Mandates of Trust in the Doctor-patient Relationship” *Qualitative Health Research* 21(9): 1182-1190

Thom DH, Hall MA, Pawlson LG (2004) “Measuring Patients' Trust in Physicians When Assessing Quality of Care” *Health Affairs* 23(4): 124-132

Thomas L (1990) “Trust, Affirmation, And Moral Character: A Critique Of Kantian Morality” *Nous*, 24:294-295

Walton M (1998) *The Trouble with Medicine: Preserving the Trust Between Patients and Doctors.* St Leonards: Allen & Unwin

7. Non-deceptive placebos and the nocebo effect.

Our doubts are traitors,

and make us lose the good we oft might win

– Shakespeare*, Measure for Measure*

Introduction

In this chapter I will examine in more detail the two main assumptions I have made for the majority of this thesis: that placebos do not carry the risk of side-effects, and that placebo treatments require deception to be effective. These assumptions were made for the sake of simplicity, in order to focus on the ethical question of whether a deceptive placebo treatment would be ethically acceptable if that was the best or only treatment available. However, in real-world cases of placebo administration both assumptions might turn out to be false: a placebo treatment has the potential to lead to adverse effects through the nocebo effect, where the factors that are usually responsible for a beneficial placebo response instead produce an adverse reaction, and some recent studies have indicated that a placebo treatment may have a beneficial result even when patients are informed about the treatment's inert nature.

In this chapter I examine the evidence that might lead us to reject these two assumptions, and discuss the implications that rejecting these assumptions might have for the conclusions I have drawn so far. The possibility of a nocebo effect, I argue, is not a problem that is specific to placebo treatments, as the nocebo effect has the potential to take place in active treatments as well. Therefore, the risk of nocebo side-effects does not

constitute a reason to avoid administering a placebo treatment, if a placebo treatment is believed to be the best or only treatment available.

The prospect of effective non-deceptive placebo treatments, on the other hand, has the potential to overturn many of the arguments I have made in favour of placebo deception. So far I have argued that it is permissible to engage in only the minimum level of deception necessary to facilitate a therapeutic placebo effect, so if this therapeutic effect can be achieved without deception, then no deception is justified. However, I find that at present there is insufficient evidence to support the conclusion that non-deceptive placebos are therapeutically effective.

Placebo side-effects: The nocebo effect

Walter Kennedy first coined the term “nocebo” – Latin for “I shall harm”, roughly the opposite of “placebo” – in 1961, a few years after the publication of Henry Beecher's “Powerful Placebo” paper.[1] Kennedy defined the nocebo effect as any unpleasant response to a real or dummy treatment that is not attributable to the pharmacological action of the medication. The nocebo effect occurs when those factors that are usually responsible for a beneficial placebo response – expectancy, conditioning, and the rest – instead produce an adverse reaction to a treatment, often through negative emotional reactions like fear and anxiety.[2] As I discussed in chapter three, negative beliefs and expectations can reduce the effectiveness of even active treatments, by weakening that component of the therapeutic effect of the treatment that is attributable to the placebo

1 Kennedy WP (1961)
2 See for example Barksy et al (2002)

effect. The nocebo effect, however, involves more than just the loss of a therapeutic placebo effect – in some cases, expectations of a bad outcome can actually lead to a bad outcome. Negative expectancy, conditioning, or other factors can create adverse reactions and side-effects that would not otherwise exist, and can result in a treatment that would usually be inert or beneficial actually being harmful to the patient. Commonly reported nocebo effects include headache, nausea, fatigue, dizziness, and insomnia.[3] There is extensive evidence to support the existence of this phenomenon, from the incidence of adverse reactions in the placebo arms of clinical trials, to studies of the impact of expectancy on the outcomes of active treatments, to studies of the negative effects of belief and expectancy outside of a medical context.

A classic demonstration of the nocebo effect was carried out by Schweiger and Parducci in 1981. These researchers told participants that an electrical current, which might produce a mild headache, would be passed through their heads while they carried out a task. No such electrical current was actually applied to the participants in the study, but over two-thirds of the participants reported headaches nonetheless. The headaches occurred at the same rate regardless of whether the headache was emphasised as the main likely effect of the current, or whether the emphasis was on the task, with the headache mentioned as only a possible side effect.[4]

Evidence of the nocebo effect causing adverse reactions to inert medical treatments is provided by Armanzio et al, who carried out a systematic review of 73 trials of different types of anti-migraine medication. These researchers found that participants in the control arms of these trials frequently reported adverse effects, including vomiting, diarrhea,

3 Barksy et al (2002) p623
4 Schweiger and Parducci (1981)

abdominal pain, dizziness and memory difficulties, from the placebo pills they were taking.[5] Those side-effects associated with the active drug being tested were significantly more likely to show up in the corresponding placebo arm, suggesting that the expectations of patients and possibly of researchers could be behind these reactions.[6] A systematic review of 143 placebo controlled trials of antidepressant medications by Rief et al returned a similar result, with the adverse effects reported in the various placebo arms conforming closely to those expected from the drug being trialled.[7] A review of clinical trials of various medications carried out by Weihrauch and Gauler also found that the profile of adverse reactions reported in the placebo groups was largely similar to the side effect profile of the active treatment being studied.[8]

The results of these reviews, where the nocebo side-effects observed in the placebo arms of clinical trials match the active side-effects expected in the treatment arms, suggest that a major cause of these nocebo effects is the participants' expectations of adverse reactions. These expectations presumably result from being warned about the risk of these side-effects during the informed consent process, or being asked about them by researchers during examinations. However, there is no reason to suppose that these expectation-mediated nocebo side-effects only occur in the placebo arm of these trials – the nocebo effect can arise in active treatments just as it can in placebo treatments. For example, in a trial of aspirin for the treatment of angina, Myers et al found that including a statement in the consent form outlining possible gastrointestinal side-effects from the treatment resulted in a sixfold increase in the number of participants reporting these symptoms – in both the placebo group and the active treatment group alike.[9]

5 Amanzio M et al (2009)
6 Ibid.p266
7 Rief et al. 2009
8 Weihrauch TR, Gauler TR (1999) p385
9 Myers et al (1987)

Negative expectations, then, can make active treatments less effective and can make side-effects more likely. However, the nocebo effect also has the potential to make otherwise therapeutic treatments actually harmful. For example, Dworkin et al showed that the anaesthetic effect of nitrous oxide could be totally reversed when patients were led to expect that the gas would increase awareness of and sensitivity toward bodily sensations. For the group who were told to expect this, the treatment actually served to increase the perception of pain, rather than to dull it.[10] Moreover, the impact of negative expectations like these on pain perceptions can be long-lasting: Rodriguez-Raecke et al showed that a single instance of negative information could increase participants' perceptions of pain over an eight-day period. Compared with a control group, the participants who were told, on just one occasion, to expect increased pain levels reported higher perceptions of pain in response to stimuli delivered on each of the following eight days.[11]

Some have argued that the negative effects of belief and expectancy can go well beyond hyperalgesia and other unpleasant side-effects, and might even be capable of leading to death. The anthropologist Walter Cannon's accounts of “voodoo death” amongst tribal people, where the fear and expectation of death following a curse or the violation of taboo leads to the deterioration and death of an individual,[12] could perhaps be called an extreme example of nocebo response. More recently, Eaker et al found that the belief that one is susceptible to heart attacks is itself a serious risk factor for death by coronary condition – over the 20-year course of the study, women who believed that they were more likely than others to suffer a heart attack were actually 3.7 times more likely to die of coronary conditions than women who didn't have this belief. This was independent of any

10 Dworkin SF et al (1983)
11 Rodriguez-Raecke et al (2010)
12 Cannon W (1942)

recognised risk factors, such as smoking or high blood pressure, that might have given individuals good reason for this belief.[13]

Implications of the possibility of harm for deceptive placebo treatments

There is good reason, then, to believe that an inert treatment, accompanied by negative beliefs or expectations, could possibly lead to a negative outcome for a patient. Is this a reason to avoid administering a deceptive placebo treatment? I do not think so. Firstly, it is important to remember is that it is not the placebo treatment itself that causes the nocebo effect. As Walter Kennedy noted when he coined the term, the effect is due to a "quality inherent in the patient, not the remedy".[14] It is the patient's reaction to being given the treatment, to the way the treatment is presented, and to the context of the treatment, that causes the nocebo effect. This was neatly demonstrated in a classic study by Luparello et al, where nearly half of asthmatics who were exposed to a mist of inert saline solution, and told they were inhaling irritants or allergens, experienced significantly increased airway resistance and reduced lung function. Twelve of these subjects developed full-blown asthma attacks, which were subsequently relieved by the same saline mist, presented this time as a therapy.[15] Whether these participants' reaction to the treatment was positive or negative, whether the saline mist acted as an irritant or a therapy, a placebo or a nocebo, depended entirely on the way the researchers presented it, and the expectations and attitudes this presentation engendered.

The possibility of a nocebo effect, then, is not a reason to avoid using placebo treatments,

13 Eaker E et al (1992)
14 Kennedy WP (1961) p203
15 Luparello et al (1968)

it is a reason to avoid engendering low expectations or negative beliefs in patients. While discussing the likely risks of a treatment, along with a patient's concerns and worries about it, is an important part of the therapeutic encounter, and the informed consent process (I will return to this point shortly), care should be taken to avoid the promotion of an excessively negative outlook. This could reduce the effectiveness of placebo treatments and active treatments alike, and can even lead to harmful outcomes. Avoiding ill-founded negative beliefs and expectations, and encouraging positive ones, should be a part of good clinical practice, no matter what kind of treatment a clinician is administering. Good clinical practice, then, should serve to maximise the placebo benefits and minimise the placebo risks of any treatment administered, at least as much as is possible while still meeting the requirements of informed consent.

The fact that nocebo effects occur in active treatments as well as in placebo treatments is another factor to consider. In a situation where a doctor or nurse is weighing up the relative merits of a deceptive placebo treatment and an active treatment, it is not the case that the possibility of a nocebo effect is a risk that only applies to the placebo treatment. Negative expectations or beliefs have the potential to affect either treatment, possibly lessening, negating or even reversing any potential benefit that either might provide. This is simply another factor that must be considered when deciding whether or not a deceptive placebo treatment presents a patient's best chance at achieving some therapeutic benefit.

It might be argued here that there is less chance of a nocebo effect resulting from an active treatment than from an inert placebo treatment, simply because it is easier for a doctor or nurse to promote positive expectations towards, say, antibiotics than towards a sugar pill. However, I would argue that this will not be the case in any situation where a

placebo treatment should be under serious consideration. If a clinician is considering administering a placebo treatment instead of an active one, it is either because the active treatment is likely to be no more effective than the placebo, or because the specific side-effects of the effective treatment are likely to outweigh any specific therapeutic effect it might have. In these circumstances it is not apparent that it will be easier to promote positive expectations towards the active treatment, at least not while meeting the requirements of informed consent.

More generally, however, a clinician who is not confident of their ability to promote positive expectations towards a placebo treatment is just not someone who should be considering such a treatment in the first place. Engendering positive attitudes and expectations towards a treatment is as essential to the promotion of a positive placebo response as it is to avoiding a negative nocebo effect, so any clinician who thinks that their patient's negative expectations might lead to a nocebo effect should also not consider that these expectations do not leave much chance for a positive placebo effect. If a clinician is unsure of their ability to promote positive attitudes and expectations towards a placebo treatment, then, it seems unlikely that they would have good reason to judge that a placebo treatment is a good option for their patient.

I would therefore suggest that the chance of nocebo effects occurring when a clinical placebo treatment is administered well are fairly low. This is contrary to data from clinical trials, which suggests that around 20% of healthy volunteers taking placebos will report adverse effects.[16] However, in the context of a clinical trial, the researchers administering the treatments are not likely to be trying to actively encourage the placebo effect the way a doctor or nurse in clinical practice should. And those practices that serve to promote a

16 Rosenzweig et al (1993)

positive placebo effect – reassuring and encouraging positive expectations and attitudes towards a treatment – are the same that are required to avoid a nocebo effect. We should therefore expect the incidence of nocebo effects in the placebo arms of clinical trials to be substantially higher than that found in clinical practice.

Doctors and nurses who are skilled at promoting positive attitudes and expectations towards a treatment will be well equipped to help their patients benefit from positive placebo effects and avoid negative nocebo effects. Clinicians with the ability to do this will be well equipped to employ a placebo treatment, should they judge that it is the best option, but they will also be able to ensure their patients get the most out of the active treatments they administer. The skills necessary to minimise the risks of nocebo effects, then, should just be considered part of good clinical practice.

The nocebo effect and informed consent

The possibility of a nocebo effect resulting from a treatment has interesting implications for informed consent, as it relates to both deceptive placebo treatments and active treatments. This is because the likely side-effects of a treatment is information that most patients would probably consider material, and is therefore something that they should be informed about. However, there might be a very fine line between informing a patient of potential side-effects and encouraging these same side-effects through the nocebo effect. The challenge for the clinician, then, is to carry out the disclosure of possible side-effects in such a way that information on the likely risks of a treatment is conveyed reliably, but the possibility of producing harm through negative beliefs and expectations is minimised.

However, the problem of how to disclose to a patient the possible nocebo side-effects of an inert placebo treatment creates a kind of catch-22 situation, where the likely side-effects of the treatment should be disclosed, but the most likely source of these side effects is the disclosure process itself.

It should be noted, however, that this dilemma is not unique to placebo treatments. Take for example the drug finasteride, used to treat an enlarged prostate. In a study, nearly 45% of participants who were told that the treatment "may cause erectile dysfunction, decreased libido, or problems of ejaculation, but these are uncommon" subsequently reported experiencing sexual dysfunction. By comparison, only 15% of participants not warned about these adverse effects reported problems.[17] This result would suggest that the proportion of participants in the first group experiencing adverse effects due to the nocebo effect is around 30%, and raises the unpleasant prospect that being warned about possible adverse effects of this drug is twice as likely to cause these adverse effects as the drug itself is.

So how should a clinician who is considering prescribing finasteride to a patient approach the disclosure of possible side-effects in this situation? On the face of it, the best way to promote a beneficial outcome would be to simply not mention the possibility of side-effects. However, a 15% chance of adverse reactions is not negligible, and is a risk that most patients would probably want to be informed about. Omitting this information would therefore be likely to violate the standards of informed consent. But on the other hand, to inform a patient of the risk of an adverse effect is to possibly triple the chance of an adverse effect occurring. Problems like this lead Wells and Kaptchuk to argue that full disclosure of the possible adverse outcomes of a treatment, by making those outcomes

17 Mondaini N et al (2007)

more likely, could violate the principle of nonmaleficence.[18] Instead, they propose what they call “contextualised informed consent”, where a clinician avoids mentioning vague or non-specific side-effects, and tailors disclosure of specific drug side-effects to the needs of individual patients, taking care to frame this information carefully to minimise unnecessary negative expectancies. This, they argue, balances the need for disclosure with the obligation to minimise the harm that can result from this disclosure.

However, it is not immediately obvious how Wells and Kaptchuk's solution helps the clinician in the finasteride situation. The 15% chance of sexual dysfunction is a specific side-effect that is likely to be material information for most patients, so in most cases this is going to need to be disclosed. The key feature of this approach, then, is that the disclosure must be framed positively. Telling patients that there's an 85% chance that their sex life will be unaffected, and perhaps emphasising the positive effects that the prostate treatment can have in this area, could be an effective method of avoiding negative expectations and attitudes around the treatment, and thus lowering the incidence of nocebo side-effects.

Colloca and Miller provide an alternative approach to this problem. They argue that, in cases where side-effects are not severe, rather than clinicians deciding how much disclosure is appropriate on their patient's behalf, clinicians should warn patients about the possibility that disclosure will negatively affect them, and let them decide for themselves whether or not they want to hear about possible side effects.[19] They call this “authorised concealment”, and propose that a clinician might pitch it to a patient along the lines of the following:

> A relatively small proportion of patients who take Drug X experience various side effects that

18 Wells and Kaptchuk (2012) p25
19 Colloca and Miller (2013)

> they find bothersome but are not life-threatening or severely impairing. Based on research, we know that patients who are told about these sorts of side effect are more likely to experience them than those who are not told. Do you want me to inform you about these side effects or not?[20]

I'm not sure why Colloca and Miller see the need to use such vague language here. Oblique references like this can be useful when trying to inform a patient about a placebo treatment without tipping them off about the inert nature of that treatment, but in the case of the nocebo effect I don't see any reason not to call a spade a spade. I would therefore suggest that a possible solution here is to be open about the nature of the risk, and inform the patient about nocebo effects and the role of their attitudes and expectations in the outcome of the treatment. Coupled with a positive framing approach, this could minimise the risk of nocebo side-effects. The disclosure in the finasteride case, then, might look something like like the following:

“This treatment is most likely not going to affect your sex life, and in fact if it works you're probably going to see some improvement in that area. The likelihood of getting a good result from this treatment is greatly increased if you adopt a positive attitude and try to expect the best outcome possible. While there is a small chance of an adverse reaction, which I will tell you about if you wish, any adverse reactions are actually more likely to come from your own expectations than the treatment itself, and this risk is something you can influence with your own attitude.”

If educated about the role of their beliefs and attitudes in the treatment, patients can hopefully take steps towards encouraging the beneficial side of these while at the same time minimising the chances of negative outcomes.

20 Ibid.p283

A disclosure along the lines of the above, I argue, would also be appropriate for placebo treatments. Where a doctor or nurse believes that a deceptive placebo presents a patient's best or only chance at some therapeutic benefit, that chance can be maximised by emphasising the importance of positive attitudes and expectations, while acknowledging the risk of drawbacks that might result from negative ones. As the nocebo effect is a phenomenon affecting active treatments and placebos alike, patients can be informed about this possibility without being tipped off about the inert nature of the proposed treatment. In this way, then, the patient can be informed about the treatment and its possible risks as well as maximising their chances of a beneficial placebo response.

Non-deceptive placebo treatments

The second major assumption I make in this thesis is that a placebo treatment has to be administered deceptively in order to be effective – that a patient's knowledge of a placebo treatment's inert nature precludes the possibility of a therapeutic placebo effect. This is the received wisdom about placebo treatments, and intentional placebos in clinical practice are, and always have been, administered deceptively. Because of this, however, there have been very few studies into the effect of deception on the efficacy of placebo treatments, and there is no definitive historical study that a clinician could point to that establishes that a placebo treatment must be administered deceptively to be effective. In fact, the few studies that have been done on deception and placebos suggest that the received wisdom might be wrong about this, and that placebo treatments administered in an open or non-deceptive fashion may nonetheless lead to a therapeutic placebo effect.

The first such study was carried out by Park and Covi in 1965.[21] Fourteen neurotic patients showing signs of anxiety were given a bottle of placebo pills, to be taken three times a day, and informed about the treatment as follows:

> Mr. Doe, at the intake conference we discussed your problems and your condition, and it was decided to consider further the possibility and the need of treatment for you before we make a final recommendation next week. Meanwhile, we have a week between now and your next appointment, and we would like to do something to give you some relief from your symptoms. Many different kinds of tranquilizers and similar pills have been used for conditions such as yours, and many of them have helped. Many people with your kind of condition have also been helped by what are sometimes called “sugar pills,” and we feel that a so-called sugar pill may help you, too. Do you know what a sugar pill is? A sugar pill is a pill with no medicine in it at all. I think this pill will help you as it has helped so many others. Are you willing to try this pill?[22]

After a week of the placebo treatment, 13 of the 14 patients showed improvement over a range of self-reported and psychiatrist-measured symptoms. The average improvement in symptom scores reported by the group was more than 40%, which the researchers note is larger than the improvement they had observed in trials of active drugs using the same measures.[23] The researchers conclude that patients can be willing to take an open placebo treatment, and that knowledge of the nature of a placebo treatment does not preclude the possibility of improvement from taking the placebo.

While this is a remarkable result, and this study has been widely cited, I am reluctant to endorse Park and Covi's conclusion here. The small sample size and short duration of the study make it doubtful that these results can be generalised, and the lack of a no-

21 Park and Covi (1965)
22 Ibid. p337
23 Ibid. p338

treatment control group makes it impossible to tell how much of the improvement reported is due to factors like natural progression, the cessation of medication, or regression to the mean. However, the major flaw with this conclusion is revealed in the post-treatment interviews the researchers conducted with patients: the researchers had simply failed to successfully communicate the nature of the placebo treatments to the patients. At the follow-up appointment eight patients stated that they believed the pills were placebos, with three patients stating that they were certain of this. Six of the returning patients thought the pills contained drugs, with two patients certain that they did. So Park and Covi's conclusion – that knowledge of the nature of a placebo treatment does not preclude the possibility of a therapeutic placebo effect – can only actually be said to apply to three of the participants in the study. The remaining 11 participants either expressed some doubt as to the nature of the pill or expressed a belief that it contained active medication. In these cases the improvement measured cannot be said to take place despite being informed about the inert nature of the pill.

A few other small studies have also produced promising results for non-deceptive placebos when employed for ADHD,[24] pain,[25] and symptoms of irritable bowel syndrome.[26] However, the first major study to compare a non-deceptive placebo treatment to a no-treatment control group was carried out by a team headed by Ted Kaptchuk in 2010.[27] These researchers recruited 80 patients with irritable bowel syndrome and divided them randomly into two groups. One group received no treatment, while the other took a placebo pill twice a day. The placebo pills came in a bottle labelled “placebo pills”, and were explained to patients as follows:

24 Sandler and Bodfish (2007)
25 Chung et al (2007)
26 Vase et al (2003)
27 Kaptchuk et al (2010)

> The provider clearly explained that the placebo pill was an inactive (i.e., "inert") substance like a sugar pill that contained no medication and then explained in an approximately fifteen minute *a priori* script the following "four discussion points:" 1) the placebo effect is powerful, 2) the body can automatically respond to taking placebo pills like Pavlov's dogs who salivated when they heard a bell, 3) a positive attitude helps but is not necessary, and 4) taking the pills faithfully is critical. Patients were told that half would be assigned to an open-label placebo group and the other half to a no-treatment control group. Our rationale had a positive framing with the aim of optimizing placebo response.[28]

After 21 days, the outcomes were measured using questionnaires, and the placebo group reported significantly higher symptom improvement scores than the no-treatment group. The researchers conclude that "patients given open-label placebo in the context of a supportive patient-practitioner relationship and a persuasive rationale had clinically meaningful symptom improvement that was significantly better than a no-treatment control group with matched patient-provider interaction."[29]

While this is a very promising result, much more research is needed before the prevailing wisdom on the necessity of deception in placebo treatments can be overturned. One study on 80 participants with one condition is simply not enough evidence to countermand a clinical practice that possibly dates back thousands of years. The researchers themselves acknowledge the limitations of their trial, describing it as a "proof-of-principle pilot study" that needs to be replicated with a larger sample size and longer follow-up,[30] and that further research is needed into open-label placebo treatments for other conditions.

One thing that the previous studies of open-label placebos have in common is that they focus on the expectancy side of the placebo effect: the researchers mostly seek to

28 Ibid. p2
29 Ibid. p4-5
30 Ibid. p6

promote a non-deceptive placebo response by framing the treatment in positive language and engendering positive expectations towards the placebo treatment. One recent study, however, has indicated that a conditioning-mediated non-deceptive placebo response is also a possibility.[31] In this study, participants were conditioned with pain relief (in reality a reduced pain stimulus) from a placebo cream for either one or four days before the nature of the placebo was revealed to them. Following the reveal, the participants reported that they subsequently expected no pain relief from the placebo at all. However, in the test that followed the analgesic effect of the placebo cream persisted for the four-day conditioning group, but not for the one-day conditioning group. This result suggests that the analgesic effect of the open placebo was not mediated by conscious expectancy, but instead by unconscious conditioning resulting from the history of exposure to the placebo treatment.

While this is an interesting result, it is difficult to see how a non-deceptive placebo treatment that is reliant on a long period of deceptive placebo conditioning is going to be helpful in a clinical context. The authors suggest that these results could be applicable to a placebo treatment openly substituted for an active painkiller after a similar length of time,[32] however this is a substantially different situation to the one in the trial. The participants in the trial were conditioned with a treatment that was revealed to be a placebo, then administered the same treatment again. In the case of an active painkiller, the treatment revealed to be a placebo is different to the treatment that the patients have been conditioned with. This is not to say that a conditioning-mediated non-deceptive placebo response in such a situation is impossible, however I don't think that the possibility of this can be extrapolated from the results of this study.

31 Schafer et al (2014)
32 Ibid. p419

In summary, then, there is perhaps reason to be cautiously optimistic about the therapeutic prospects of non-deceptive placebo treatments. In particular, if the general public is becoming more aware of the therapeutic potential of the placebo effect, then it will become more and more feasible to reveal the inert nature of a placebo treatment to a patient without deflating the expectations of therapeutic benefit on which the placebo effect is thought to partially depend. And this spread of awareness is something that seems to be taking place, perhaps facilitated by extensive media coverage of the work of placebo researchers like Ted Kaptchuk, Irving Kirsch, and others whose names appear frequently in this thesis. Indeed, wider knowledge about placebos and more positive attitudes towards placebo treatments is one possible reason why the placebo effect observed in clinical trials seems to be getting stronger.[33,34] However, much more research is needed – it is far too early yet to declare that placebo treatments can be effective without deception.

The implications of non-deceptive placebo treatments

If non-deceptive placebo treatments were definitively shown to have a therapeutic effect akin to that of deceptive placebo treatments, the conclusions I have reached thus far in this thesis would have to be overturned. To pick one example, my conclusion that placebo deception could be consistent with a Kantian respect for rational agency was predicated on the idea that knowledge of the inert nature of a placebo treatment obviates the placebo effect as a means towards that agent's end. If this is not the case, if knowing about the nature of a placebo treatment can be consistent with employing this treatment as a means to one's end, then withholding this information cannot be justified. My discussion in the

33 Leucht et al (2013)
34 Tuttle et al (2015)

informed consent chapter likewise centred around the tension between a patient's informational needs and their therapeutic needs. The prospect of non-deceptive placebo treatments could dissolve this tension entirely.

What if non-deceptive placebo treatments are shown to be effective, but not as effective as deceptive ones? If it was the case that a deceptive placebo was more likely to have a therapeutic effect, or likely to have a greater therapeutic effect, than a non-deceptive placebo, then this would appear to fit my ideal deceptive placebo case, in that the deceptive placebo presents the best chance of a therapeutic benefit. And indeed, if the expected therapeutic benefit of a deceptive placebo is substantially larger than that of a non-deceptive placebo, I would agree that deception could be justified in such a case. However, I do not think that such a substantial difference is likely to be demonstrated. The expected magnitude of the placebo effect is generally fairly minor, so for non-deceptive placebos to have a statistically significant therapeutic effect at all, their expected placebo effect will probably have to be in the same ballpark as deceptive ones. Any additional expected benefit from a deceptive placebo will probably therefore not be sufficient to justify the deception.

However all of this is merely speculation. It is far too early yet to tell if non-deceptive placebo treatments will indeed be an effective replacement for their deceptive counterparts. Until a lot more research is done, then, I would suggest that deceptive placebos can be used in an ethically acceptable fashion as outlined in the rest of this thesis.

Conclusion

In this chapter I have considered the effects that the rejection of the two main assumptions I make in this thesis would have on my conclusions thus far. The possibility of placebo side-effects occurring through the nocebo effect presents an interesting problem, but this is not a problem that is specific to placebo treatments. The nocebo effect is something that should be taken into consideration in all therapeutic encounters, and the possibility of a nocebo effect does not constitute a reason to avoid administering a deceptive placebo treatment where this is believed to present a patient's best or only chance at some therapeutic benefit. Nor does the possibility of a nocebo effect create a problem for informed consent as it applies to placebo treatments, or at the very least it creates no problem for placebo treatments that it does not also create for at least some active treatments.

The prospect of effective non-deceptive placebo treatments is an exciting possibility, and one that has the potential to render the arguments I make in majority of this thesis obsolete. However the question of whether or not non-deceptive placebos can be therapeutically effective is not one that is going to be definitively answered any time soon. Until there is enough research carried out on this topic to answer the question one way or the other, the received wisdom that placebo treatments need to be administered deceptively will still apply, as will the arguments I make here.

References

Amanzio M, Corazzini L, Vase L, Benedetti F (2009) “A systematic review of adverse events in placebo groups of anti-migraine clinical trials” *Pain* 146: 261-269

Barsky AJ, Saintfort R, Rogers MP, Borus JF (2002) “Nonspecific medication side effects and the nocebo phenomenon” *Journal of American Medical Association* 287(5): 622–627

Bingel U, Wanigasekera V, Wiech K, Ni Mhuircheartaigh R, Lee MC, Ploner M, Tracey I (2011) “The Effect of Treatment Expectation on Drug Efficacy: Imaging the Analgesic Benefit of the Opioid Remifentanil” *Science Translational Medicine* 3: 70ra14

Cannon W (1942) "Voodoo Death" *American Anthropologist* 44: 169–181.

Chung SK, Price DD, Verne GN, Robinson ME (2007) “Revelation of a personal placebo response: Its effects on mood, attitudes and future placebo responding” *Pain* 132: 281–288

Colloca L, Miller FG (2013) “Placebo and nocebo: Ethical challenges and solutions” In Colloca L, Flaten M, Meissner K, (eds) *Pain and placebo: From bench to bedside* Boston: Elsevier/Academic Press p278-286

Dworkin SF, Chen AC, LeResche L, Clark DW (1983) “Cognitive reversal of expected nitrous oxide analgesia for acute pain” *Anesthesia & Analgesia* 62(12): 1073-7.

Eaker E, Pinsky J, Castelli WP. (1992) "Myocardial infarction and coronary death among women: psychosocial predictors from a 20-year follow-up of women in the Framingham Study" *American Journal of Epidemiology* 135: 854–64

Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, et al. (2010) "Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome" *PLoS ONE* 5(12): e1559

Kennedy WP (1961) “The nocebo reaction” *Medical World* 95: 203–205

Leucht S, Heres S, Davis JM (2013) “Increasing Placebo Response in Antipsychotic Drug Trials: Let’s Stop the Vicious Circle” *American Journal of Psychiatry* 170(11): 1232-1234

Luparello T, Lyons HA, Bleecker ER, McFadden ER (1968) “Influences of suggestion on airway reactivity in asthmatic subjects” *Psychosomatic Medicine* 30: 819–25

Mondaini N, Gontero P, Giubilei G, et al (2007) “Finasteride 5 mg and sexual side effects: how many of these are related to a nocebo phenomenon?” *Journal of Sexual Medicine* 4(6): 1708-1712.

Myers, MG, Cairns JA, Singer J (1987) “The consent form as a possible cause of side effects” *Clinical Pharmacology and Therapeutics* 42(3): 250–253

Park LC, Covi L (1965) “Nonblind Placebo Trial: An Exploration of Neurotic Patients' Responses to Placebo When Its Inert Content Is Disclosed” *Archives of General*

Psychiatry 12: 336-345

Rief W, Nestoriuc Y, von Lilienfeld-Toal A (2009) “Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: A systematic review and meta-analysis” *Drug Safety* 32(11): 1041–1056.

Rodriguez-Raecke R, Doganci B, Breimhorst M, Stankewitz A, Büchel C, Birklein F, May A (2010) “Insular Cortex Activity Is Associated with Effects of Negative Expectation on Nociceptive Long-Term Habituation” *The Journal of Neuroscience* 30(34): 11363-11368

Rosenzweig P, Brohier S, Zipfel A (1993) “The placebo effect in healthy volunteers: influence of experimental conditions on the adverse events profile during phase I studies” *Clinical Pharmacology & Therapeutics* 54: 579-583.

Sandler AD, Bodfish JW (2007) “Open-label use of placebos in the treatment of ADHD: a pilot study” *Child Care and Health Development* 34: 104-110

Schafer SM, Colloca L, Wager TD (2014) “Conditioned placebo analgesia persists when subjects know they are receiving a placebo” *Journal of Pain* 16(5): 412-20

Schweiger A, Parducci A (1981) “Nocebo: The psychologic induction of pain” *The Pavlovian journal of biological science* 16(3): 140-143

Tuttle AH, Tohyama S, Ramsay T, Kimmelman J, Schweinhardt P, Bennett GJ, Mogil JS (2015) “Increasing placebo responses over time in U.S. clinical trials of neuropathic pain”

published as an ahead-of-print article on *Pain*'s website 22 August 2015. doi: 10.1097/j.pain.0000000000000333

Vase L, Robinson ME, Verne GN, Price DD (2003) "The contributions of suggestion, desire, and expectation to placebo effects in irritable bowel syndrome patients: An empirical investigation" *Pain* 105: 17–25

Weihrauch TR, Gauler TR (1999) "Placebo – Efficacy and Adverse Effects in Controlled Clinical Trials" *Arzneimittelforschung* 49(5): 385-393

Wells RE, Kaptchuk TJ (2012) "To Tell the Truth, the Whole Truth, May Do Patients Harm: The Problem of the Nocebo Effect for Informed Consent" *The American Journal of Bioethics* 12(3): 22-29

Conclusion

Love, like medicine, is only the art of encouraging nature

Laclos, *Les Liaisons Dangereuses*

In this thesis I have attempted to address what I believe are the main ethical problems raised by the prospect of deceptive placebo use in clinical practice. Where a doctor or nurse believes that a deceptive placebo presents a patient's best, or only, chance at some therapeutic benefit, I believe that they are justified in withholding that information necessary to promote a therapeutic placebo effect. I have argued that this deception can be carried out in a manner that is consistent with respect for patient autonomy, that meets the standards of informed consent, and that does not threaten the trust relationship between patient and practitioner. It is true that there will be patients whose informational needs will exceed the level at which a deceptive placebo is possible, or whose personality means that a deceptive placebo will risk the loss of their trust. But for such patients, it is simply not the case that a deceptive placebo treatment presents their best treatment option. Whether or not a deceptive placebo is a good option for any individual patient is a matter of practical judgement for the doctor or nurse involved.

There are other practical problems I have not touched on in this thesis. I have not delved into the details of how deceptive placebos might be administered, whether patients should be sent to the pharmacy for placebo pills, how much they should cost, and so on. There is the question of whether vitamins or herbal supplements, or even homeopathic treatments, might be used instead of inert sugar pills. There is also the problem of how to prevent patients from finding out that their course of medication is a placebo, in an age where

every brand name is easily googleable. And these are only a few of the practical problems raised by the prospect of deceptive placebo use in clinical practice. On the other hand, however, the surveys I reviewed in chapter three indicate that clinical placebo use is currently fairly prevalent, so doctors and nurses have probably already figured out ways around at least some of the practical problems involved in administering deceptive placebos.

In many of the cases where placebos are currently employed in clinical practice, using a deceptive placebo treatment is, and should be, impermissible. Placebos are used to mollify or punish difficult patients, patients are lied to about the nature of placebo treatments, and antibiotics are used as placebo pills. I do not wish to defend such cases here. However, the current regulatory environment, where ethical guidelines concerning clinical placebo use are inconsistent, nonexistent, or unworkable, does nothing to dissuade such behaviour. A consistent ethical approach is needed, one that clearly highlights the difference between justifiable placebo use, which is carried out in a manner that is mindful of patients' therapeutic needs, their autonomy, and their trust, and impermissible cases like the above. I hope that the arguments I have made in this thesis can provide such an approach.

Bibliography

Amanzio M, Corazzini L, Vase L, Benedetti F (2009) “A systematic review of adverse events in placebo groups of anti-migraine clinical trials” *Pain* 146: 261-269

American Medical Association (2014) *Code of Medical Ethics*, American Medical Association Press

Anderson LA, Dedrick RF (1990) “Development of the Trust in Physician Scale: a Measure to Assess Interpersonal Trust in Patient-physician Relationships” *Psychological Reports* 67(3): 1091-1100

Aronson J (1999) “Please, please me” *British Medical Journal* 318:716

Baghcheghi N, Koohestani H (2011) “Placebo use in clinical practice by nurses in an Iranian teaching hospital” *Nursing Ethics* 18(3): 364–373

Baier, A (1986) “Trust and Antitrust” *Ethics* 96(2): 231-260

Baier, A (1990) “Why Honesty Is a Hard Virtue” in Flanagan OJ, Rorty A (eds) *Identity, Character, and Morality*. Cambridge: MIT Press 259-282

Baier, A (1992) “Trusting People” *Philosophical Perspectives* 6: 137-153

Barnhill A (2011), “What it takes to defend deceptive placebo use” *Kennedy Institute of*

Ethics Journal 21(3): 219-250

Barnhill, A (2012) “Clinical Use of Placebos: Still the Physician”s Prerogative?” *Hastings Center Report* 42(3): 29-37

Barsky AJ, Saintfort R, Rogers MP, Borus JF (2002) “Nonspecific medication side effects and the nocebo phenomenon” *Journal of American Medical Association* 287(5): 622–627

Beauchamp TL, Childress J (2009) *Principles of Biomedical Ethics* (6th ed), New York: Oxford University Press

Beecher, HK (1955) “The Powerful Placebo” *Journal of the American Medical Association* 159(17):1602-1606

Benedetti F (2009) *Placebo Effects: Understanding the Mechanisms in Health and Disease*. Oxford: Oxford University Press

Berkeley, G (1747) *Siris: A Chain of Philosophical Reflexions and Inquiries concerning the Virtues of Tar-Water and divers other Subjects connected together and arising one from another.* Dublin: W Innys

Bingel U, Wanigasekera V, Wiech K, Ni Mhuircheartaigh R, Lee MC, Ploner M, Tracey I (2011) “The Effect of Treatment Expectation on Drug Efficacy: Imaging the Analgesic Benefit of the Opioid Remifentanil” *Science Translational Medicine* 3: 70ra14

BMJ (1996) “The Nuremberg Code (1947)” *British Medical Journal* 313: 1448

Bok S (1978) *Lying: Moral Choice in Public and Private Life.* New York: Pantheon Books

Bostick N, Sade R, Levine M, Stewart DM (2008) “Placebo Use in Clinical Practice: Report of the American Medical Association Council on Ethical and Judicial Affairs,” *Journal of Clinical Ethics* 19(1): 58-61

British Medical Association (2012) *Medical Ethics Today: the BMA's Handbook of Ethics and Law (3rd ed.)* Chichester: Wiley-Blackwell/BMJ Books

Brody H (1982) “The Lie that Heals: The Ethics of Giving Placebos” *Annals of Internal Medicine* 97: 112-118

Buckalew LW, Ross S (1981) "Relationship of perceptual characteristics to efficacy of placebos" *Psychological Reports* 49(3): 955–61.

Burt R, Cowart D (1998) “Confronting Death: Who Chooses, Who Controls?” *The Hastings Center Report* 28(1): 14-24

Cabot RC (1903) “The use of truth and falsehood in medicine: an experimental study” *American Medicine* 5: 344-349

Campbell B, Thom DH (1997) “Patient-physician Trust: An Exploratory Study” *Journal of Family Practice* 44(2): 169-176

Cannon W (1942) "Voodoo Death" *American Anthropologist* 44: 169–181.

Carson T (2010) *Lying and Deception: Theory and Practice*. Oxford: Oxford University Press

Chen GF, Johnson MH (2009) “Patients’ attitudes to the use of placebos: results from a New Zealand survey” *New Zealand Medical Journal* 122(1296): 35-46

Cherkin DC et al (2009) “A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain” *Archives of Internal Medicine* 169: 858-66

Cheyne C (2005) “Exploiting placebo effects for therapeutic benefit” *Health Care Analysis* 13(3): 177-188

Chung SK, Price DD, Verne GN, Robinson ME (2007) “Revelation of a personal placebo response: Its effects on mood, attitudes and future placebo responding” *Pain* 132: 281–288

Colloca L, Benedetti F (2005) “Placebos and painkillers: is mind as real as matter?” *National Review of Neuroscience* 6(7): 545-52

Colloca L, Lopiano L, Lanotte M, Benedetti F (2004) “Overt versus covert treatment for pain, anxiety and Parkinson's disease” *Lancet Neurology* 3: 679-684

Colloca L, Miller FG (2013) “Placebo and nocebo: Ethical challenges and solutions” In Colloca L, Flaten M, Meissner K, (eds) *Pain and Placebo: From Bench to Bedside.* Boston: Elsevier/Academic Press p278-286

Cullen KA, Hall MJ, Golosinskiy A (2009) “Ambulatory surgery in the United States, 2006” *National Health Statistics Reports* 11: 1-25

Cullen W (1772) *Clinical lectures 1772-3,* in RCPE Manuscript Cullen retrieved from http://www.jameslindlibrary.org/cullen-w-1772/

De Craen AJM et al (1999) “Placebo effect in the treatment of duodenal ulcer” *British Journal of Clinical Pharmacology* 48: 853-860

Diehl HS et al (1938) “Cold vaccines. An evaluation based on a controlled study” *Journal of the American Medical Association* 111: 1168-1173

Diener HC, Dowson AJ, Ferrari M, Nappi G, Tfelt-Hansen P (1999) “Unbalanced randomization influences placebo response: scientific versus ethical issues around the use of placebo in migraine trials” *Cephalalgia* 19(8): 699-700.

Dworkin SF, Chen AC, LeResche L, Clark DW (1983) “Cognitive reversal of expected nitrous oxide analgesia for acute pain” *Anesthesia & Analgesia* 62(12): 1073-7.

Eaker E, Pinsky J, Castelli WP. (1992) "Myocardial infarction and coronary death among

women: psychosocial predictors from a 20-year follow-up of women in the Framingham Study" *American Journal of Epidemiology* 135: 854–64

Evans W, Hoyle C (1933) “The comparative value of drugs used in the continuous treatment of angina pectoris” *Quarterly Journal of Medicine* 26:311-38

Evans, D (2003) *Placebo: Mind Over Matter in Modern Medicine.* London: HarperCollins

Faden RR, Beauchamp TL, King NMP (1986) *A history and theory of informed consent.* New York: Oxford University Press.

Fässler M, Gnädinger M, Rosemann T, Biller-Andorno, N (2009) “Use of placebo interventions among Swiss primary care providers” *BMC Health Services Research* 9: 144

Fässler M, Gnädinger M, Rosemann T, Biller-Andorno, N (2011) “Placebo interventions in practice: a questionnaire survey on the attitudes of patients and physicians” *British Journal of General Practice* 61: 101-107

Fässler M, Meissner K, Schneider A, Linde K (2010) “Frequency and circumstances of placebo use in clinical practice - a systematic review of empirical studies” *BMC Medicine* 8: 15

Finniss DG et al (2010) “Biological, clinical, and ethical advances of placebo effects” *Lancet* 375: 686-95

Flint A (1863) “A contribution toward the natural history of articular rheumatism; consisting of a report of thirteen cases treated solely with palliative measures” *American Journal of the Medical Sciences* 46:17-36.

Ford (ed.) (1898) *The Writings of Thomas Jefferson (vol. 10)* New York: Federal Edition

Franklin, B (1785) *Report of Dr. Benjamin Franklin, and other commissioners, charged by the King of France, with the examination of the animal magnetism, as now practised at Paris.* London: J Johnston

Geers AL, Helfer SG, Weiland PE, Kosbab K (2006) “Expectations and placebo response: a laboratory investigation into the role of somatic focus” *Journal of Behavioral Medicine* 29(2): 171-8.

Goebel MU et al. (2002) “Behavioral conditioning of immunosuppression is possible in humans” *Federation of American Societies for Experimental Biology Journal* 16:1869–73

Gold A, Lichtenberg P (2014) “The moral case for the clinical placebo” *Journal of Medical Ethics* 40: 219-224

Gold H (1946) “Cornell conferences on therapy: the use of placebos in therapy” *New York State Journal of Medicine* 46: 1718-27

Gold H, Kwit NT, Otto H (1937) “The xanthines (theobromine and aminophylline) in the treatment of cardiac pain” *Journal of the American Medical Association* 108: 2173-79

Golomb BA (2009) “Doctoring the Evidence: The case against lying to patients about placebos” *American Journal of Bioethics* 9: 34-6.

Goold SD (2001) “Trust and the Ethics of Health Care Institutions” *Hastings Center Report* 31(6): 26-33

Gordon A (1788) *The practice of physick*. Unpublished manuscript retrieved from http://www.jameslindlibrary.org/gordon-a-1788

Graves, TC (1920) “Commentary on a case of hystero-epilepsy with delayed puberty: treated with testicular extract” *Lancet* 196(5075): 1134-1135

Hall MA, Dugan E, Zheng B, Mishra AK (2001) “Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter?” *Milbank Quarterly* 79(4): 613-639

Hardin R (2006) *Trust*. Cambridge: Polity

Haygarth J (1801) *Of the imagination, as a cause and as a cure of disorders of the body: Exemplified by fictitious tractors and epidemical convulsions.* Bath: R. Cruttwell

Herman B (1993) *The Practice of Moral Judgement*. Cambridge: Harvard University Press

High Court of Australia (1992) "Rogers v Whitaker" *The Australian Law Journal* 67(1): 47-

55

Ho et al (1988) “Reduction of post-operative swelling by a placebo effect” J*ournal of Psychosomatic Research* 32:197–205

Hobbes T, in Tuck R (ed.) (1996) *Leviathan*. Cambridge: Cambridge University Press

Holt S, Gilbey A (2009) “Use of placebos by New Zealand doctors” *New Zealand Medical Journal* 122(1298): 84-86

Holton R (1994) “Deciding to Trust, Coming to Believe” *Australasian Journal of Philosophy* 72(1): 63-76

Honigfeld G (1964) “Non-specific factors in treatment: I. Review of placebo reactions and placebo reactors” *Diseases of the Nervous System* 25: 145-156

House of Commons Science and Technology Committee (2010) Committee press release (22 February 2010)

House of Commons Science and Technology Committee (2010) *Evidence Check 2: Homeopathy* London: The Stationary Office

Houston, WR (1938) “The Doctor Himself as a Therapeutic Agent” *Annals of Internal Medicine* 11(8): 1416-1425

Howick J, Bishop FL, Heneghan C, Wolstenholme J, Stevens S, Hobbs FDR, et al. (2013) “Placebo Use in the United Kingdom: Results from a National Survey of Primary Care Practitioners” *PLoS ONE* 8(3): e58247

Hróbjartsson A, Gøtzsche P (2001) “Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment” *New England Journal of Medicine* 344(21): 1594-1603

Hróbjartsson A, Gøtzsche P (2007) “Powerful Spin in the Conclusion of Wampold et al.’s Re-Analysis of Placebo Versus No-Treatment Trials Despite Similar Results as in Original Review”, *Journal of Clinical Psychology* 63(4): 373-377

Hróbjartsson A, Norup M (2003) “The use of placebo interventions in medical practice--a national questionnaire survey of Danish clinicians” *Evaluation and the Health Professions* 2003(26):153-65.

Hull SC et al. (2013) “Patients’ attitudes about the use of placebo treatments: telephone survey” *British Medical Journal* 346:f3757

Humphery N (2002) *The Mind Made Flesh.* Oxford: Oxford University Press

Hunter J (1788) *A treatise on the venereal disease*. London: G Nicol and J Johnson

Jacobs AK (2005) “Rebuilding an Enduring Trust in Medicine: A Global Mandate: Presidential Address American Heart Association Scientific Sessions 2004” *Circulation*

111(25): 3494-3498

Jefferson et al (2009), “Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis” *British Medical Journal* 339: b5106

Jones K (1996) “Trust as an Affective Attitude” *Ethics* 107(1): 4-25

Jones K (2012) “Trustworthiness” *Ethics* 123 (1): 61-85

Jones WH (ed) (1923) *Hippocrates with an English translation (Vol. II).* London: William Heinemann

Jütte R, Hoppe JD, Scriba P (2010) “'Stellungnahme des Wissenschaftlichen Beirats der Bundesärztekammer 'Placebo in der Medizin'” *Deutsches Ärzteblatt* 107(Heft 28–29): 1417-1421

Kanaan R (2009) “When Doctors Deceive” *American Journal of Bioethics* 9: 29-30

Kant I, in Beck LW (ed.) (1949) *Critique of Practical Reason and Other Writings in Moral Philosophy*. Chicago: University of Chicago Press

Kant I, in Gregor M, (ed.) (1996) *The Metaphysics of Morals.* Cambridge: Cambridge University Press

Kant I, in Gregor M, Timmerman J (eds.) (2012) *Groundwork of the Metaphysics of Morals*

(revised ed.) Cambridge: Cambridge University Press

Kaptchuk T (1998) “Powerful placebo: the dark side of the randomised controlled trial” *Lancet* 351: 1722–25

Kaptchuk TJ, Friedlander E, Kelley JM, Sanchez MN, Kokkotou E, et al. (2010) "Placebos without Deception: A Randomized Controlled Trial in Irritable Bowel Syndrome" *PLoS ONE* 5(12): e1559

Kennedy WP (1961) “The nocebo reaction” *Medical World* 95: 203–205

Kerr et al. (2008) “William Cullen and a missing mind-body link in the early history of placebos” *Journal of the Royal Society of Medicine* 101(2): 89-92

Kienle, G and Kiene, H (1997) “The Powerful Placebo Effect: Fact or Fiction?” *Journal of Clinical Epidemiology* 50(12): 1311-1318

Kihlbom U (2008) “Autonomy and Negatively Informed Consent” *Journal of Medical Ethics* 34: 146-149

King CJ (1983) “F v R” 33 SASR 189 SASC

Kirsch et al (2008), “Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration” *PLoS Medicine* 5(2): e45.

Kirsch I (ed) (1999) *How Expectancies Shape Experience*. Washington DC: American Psychological Association.

Kirsch I, Sapirstein G (1998) “Listening to Prozac but hearing placebo: a meta-analysis of antidepressant medication” Prevention and Treatment 1(2): 2a

Kirsch I, Weixel LJ (1988) “Double-blind versus deceptive administration of a placebo” Behavioral Neuroscience 102(2): 319-23.

Kirsch, I. (1991). “The placebo effect as a conditioned response: Failures of the 'litmus test.' *Behavioral and Brain Sciences* 14: 200-204.

Korsgaard C (1996) *Creating the Kingdom of Ends.* Cambridge: Cambridge University Press

Kupferschmidt, K (2011) “More placebo use promoted in Germay” *Canadian Medical Association Journal* 183(10): E633–E634

Leucht S, Heres S, Davis JM (2013) “Increasing Placebo Response in Antipsychotic Drug Trials: Let’s Stop the Vicious Circle” *American Journal of Psychiatry* 170(11): 1232-1234

Lichtenberg P, Heresco-Levy U, Nitzan U (2004) “The ethics of the placebo in clinical practice” *Journal of Medical Ethics* 30: 551–554

Lidstone SC et al. (2010) “Effects of expectation on placebo-induced dopamine release in

Parkinson disease" *Archives of General Psychiatry* 67(8): 857-865

Linde K, et al. (2014) "The use of placebo and non-specific therapies and their relation to basic professional attitudes and the use of complementary therapies among German Physicians – a cross-sectional survey" *PLoS ONE* 9(4): e92938

Luparello T, Lyons HA, Bleecker ER, McFadden ER (1968) "Influences of suggestion on airway reactivity in asthmatic subjects" *Psychosomatic Medicine* 30: 819–25

Mallinckrodt CH, Zhang L, Prucka WR, Millen BA (2010) "Signal detection and placebo response in schizophrenia: parallels with depression" *Psychopharmacology Bulletin* 43(1): 53-72.

Manson NC, O'Neill O (2007) *Rethinking informed consent in bioethics.* Cambridge: Cambridge University Press

Masson D (ed.) (1968) *The Collected Writings of Thomas de Quincey (vol. 4).* New York: AMS Press

Miller J (2007) "The Other Side of Trust in Health Care: Prescribing Drugs with the Potential for Abuse" *Bioethics* 21(1): 51-60

Moerman D (2002) *Medicine, Meaning, and the Placebo Effect.* Cambridge: Cambridge University Press

Mondaini N, Gontero P, Giubilei G, et al (2007) “Finasteride 5 mg and sexual side effects: how many of these are related to a nocebo phenomenon?” *Journal of Sexual Medicine* 4(6): 1708-1712.

Montgomery GH, Kirsch I (1997) “Classical conditioning and the placebo effect” *Pain* 72: 107-113

Moseley JB et al (2002) “A Controlled Trial of Arthroscopic Surgery for Osteoarthritis of the Knee” *New England Journal of Medicine* 347(2): 81-88

Myers, MG, Cairns JA, Singer J (1987) “The consent form as a possible cause of side effects” *Clinical Pharmacology and Therapeutics* 42(3): 250–253

Nickel PJ (2007) “Trust and Obligation-ascription” *Ethical Theory and Moral Practice* 10(3): 309-319

Nitzan U, Litchenberg P (2004) “Questionnaire survey on the use of placebo” *British Medical Journal* 329: 944-46.

O'Neill O (1984) “Paternalism and partial autonomy” *Journal of Medical Ethics* 10: 173-178

O'Neill O (1989) *Constructions of reason: explorations of Kant's practical philosophy.* Cambridge: Cambridge University Press

O'Neill O (2002) *A Question of Trust: The BBC Reith Lectures 2002*. Cambridge:

Cambridge University Press

O'Neill O (2003) “Some limits of informed consent” *Journal of Medical Ethics* 29: 4-7

Park LC, Covi L (1965) “Nonblind Placebo Trial: An Exploration of Neurotic Patients' Responses to Placebo When Its Inert Content Is Disclosed” *Archives of General Psychiatry* 12: 336-345

Pearson SD, Raeke LH (2000) “Patients' Trust in Physicians: Many Theories, Few Measures, and Little Data” *Journal of General Internal Medicine* 15(7): 509-513

Pellegrino ED, Thomasma DC (1993) *The Virtues in Medical Practice.* New York: Oxford University Press

Pollo A, Amanzio M, Arslanian A, Casadio C, Maggi G, Benedetti F (2001) “Response expectancies in placebo analgesia and their clinical relevance” *Pain* 93(1): 77-84.

Potter NN (2002) *How Can I Be Trusted?: A Virtue Theory of Trustworthiness.* Lanham: Rowman & Littlefield

Prendergast BJ et al (2002) “Photorefractoriness of immune function in male Siberian hamsters (Phodopus sungorus)” *Journal of Neuroendocrinology* 14(4): 318-29

Preston RA et al (2000) “Placebo-Associated Blood Pressure Response and Adverse Effects in the Treatment of Hypertension: Observations From a Department of Veterans

Affairs Cooperative Study" *Archives of Internal Medicine* 160(10): 1449-1454

Quincy J, Hooper R (1811) *Quincy's Lexicon-Medicum: A New Medical Dictionary.* London

Rief W, Nestoriuc Y, von Lilienfeld-Toal A (2009) "Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: A systematic review and meta-analysis" *Drug Safety* 32(11): 1041–1056.

Rodriguez-Raecke R, Doganci B, Breimhorst M, Stankewitz A, Büchel C, Birklein F, May A (2010) "Insular Cortex Activity Is Associated with Effects of Negative Expectation on Nociceptive Long-Term Habituation" *The Journal of Neuroscience* 30(34): 11363-11368

Rosenzweig P, Brohier S, Zipfel A (1993) "The placebo effect in healthy volunteers: influence of experimental conditions on the adverse events profile during phase I studies" *Clinical Pharmacology & Therapeutics* 54: 579-583.

Sandler AD, Bodfish JW (2007) "Open-label use of placebos in the treatment of ADHD: a pilot study" *Child Care and Health Development* 34: 104-110

Scanlon T (1990) "Promises and Practices" *Philosophy & Public Affairs* 19(3): 199-226

Schafer SM, Colloca L, Wager TD (2014) "Conditioned placebo analgesia persists when subjects know they are receiving a placebo" *Journal of Pain* 16(5): 412-20

Schneider C (1998), *The practice of autonomy: patients, doctors, and medical decisions.*

New York: Oxford University Press

Schweiger A, Parducci A (1981) “Nocebo: The psychologic induction of pain” *The Pavlovian journal of biological science* 16(3): 140-143

Shapiro AK (1968) "Semantics of the placebo". *Psychiatric Quarterly* 42(4): 653-695

Shapiro AK, Shapiro E (1997), *The Powerful Placebo.* Baltimore: Johns Hopkins University Press

Shaw DM (2009) “Prescribing placebos ethically: the appeal of negatively informed consent” *Journal of Medical Ethics* 35: 97-99

Sinyor M, Levitt AJ, Cheung AH, Schaffer A, Kiss A, Dowlati Y, Lanctôt KL (2010) "Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses" *Journal of Clinical Psychiatry* 71: 270–279

Skirbekk H, Middelthon A, Hjortdahl P, Finset A (2011) “Mandates of Trust in the Doctor-patient Relationship” *Qualitative Health Research* 21(9): 1182-1190

Stewart-Williams S and Podd J (2004) “The Placebo Effect: Dissolving the Expectancy Versus Conditioning Debate” *Psychological Bulletin* 130(2):324-40.

Talbot M (2000), “The Placebo Prescription” *New York Times* January 9, 2000

Thom DH, Hall MA, Pawlson LG (2004) “Measuring Patients' Trust in Physicians When Assessing Quality of Care” *Health Affairs* 23(4): 124-132

Thomas L (1990) “Trust, Affirmation, And Moral Character: A Critique Of Kantian Morality” *Nous*, 24:294-295

Tilburt JC et al. (2008) “Prescribing 'placebo treatments': results of national survey of US internists and rheumatologists” *British Medical Journal* 2008(337): a1938

Trimmer et al (2013) “Understanding the placebo effect from an evolutionary perspective” *Evolution and Human Behavior* 34(1): 8-15

Trotter T (1792) *Observations on the scurvy.* London: Longman

Tuttle AH, Tohyama S, Ramsay T, Kimmelman J, Schweinhardt P, Bennett GJ, Mogil JS (2015) “Increasing placebo responses over time in U.S. clinical trials of neuropathic pain” published as an ahead-of-print article on the *Pain* website 22 August 2015. doi: 10.1097/j.pain.0000000000000333

Vase L, Riley JL, Price DD (2002) “A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia” *Pain* 99(3): 443-52

Vase L, Robinson ME, Verne GN, Price DD (2003) “The contributions of suggestion, desire, and expectation to placebo effects in irritable bowel syndrome patients: An

empirical investigation" *Pain* 105: 17–25

Voudouris NJ et al (1990) "The role of conditioning and verbal expectancy in the placebo response" *Pain* 43: 121-128

Walton M (1998) *The Trouble with Medicine: Preserving the Trust Between Patients and Doctors.* St Leonards: Allen & Unwin

Wampold BE et al (2005) "The Placebo Is Powerful: Estimating Placebo Effects in Medicine and Psychotherapy From Randomized Clinical Trials" *Journal of Clinical Psychology* 61(7): 835-854

Wampold BE et al. (2007) "The Story of Placebo Effects in Medicine: Evidence in Context" *Journal of Clinical Psychology* 63(4): 379-390

Webster Marketon JI and Glaser R "Stress Hormones and Immune Function" *Cell Immunology* 252: 16-26

Weihrauch TR, Gauler TR (1999) "Placebo – Efficacy and Adverse Effects in Controlled Clinical Trials" *Arzneimittelforschung* 49(5): 385-393

Wells RE, Kaptchuk TJ (2012) "To Tell the Truth, the Whole Truth, May Do Patients Harm: The Problem of the Nocebo Effect for Informed Consent" *The American Journal of Bioethics* 12(3): 22-29

Wood A (1999) *Kant's Ethical Thought.* Cambridge: Cambridge University Press

World Medical Association (1964) “The Declaration of Helsinki” *British Medical Journal* 313: 1448-1449

World Medical Association (1989) “The Declaration of Helsinki” https://history.nih.gov/research/downloads/helsinki.pdf

World Medical Association (2000) "Helsinki Declaration VI" *Bulletin of Medical Ethics* 162: 8–11

World Medical Association (2008) “The Declaration of Helsinki” http://www.wma.net/e/policy/b3.htm

Zubieta JK et al. (2005) "Placebo Effects Mediated by Endogenous Opioid Activity on μ-Opioid Receptors" *Journal of Neuroscience* 25(34): 7754-7762