The relativity of 'placebos': defending a modified version of Grünbaum's definition

Article *in* Synthese · January 2016

Impact Factor: 0.74 · DOI: 10.1007/s11229-015-1001-0

1 author:

Jeremy Howick
University of Oxford
62 PUBLICATIONS **778** CITATIONS

SEE PROFILE

Synthese
DOI 10.1007/s11229-015-1001-0

The relativity of 'placebos': defending a modified version of Grünbaum's definition

Jeremy Howick[1]

Received: 2 June 2015 / Accepted: 18 December 2015

Abstract Debates about the ethics and effects of placebos and whether 'placebos' in clinical trials of complex treatments such as acupuncture are adequate (and hence whether acupuncture is 'truly' effective or a 'mere placebo') rage. Yet there is currently no widely accepted definition of the 'placebo'. A definition of the placebo is likely to inform these controversies. Grünbaum's (1981, 1986) characterization of placebos and placebo effects has been touted by some authors as the best attempt thus far, but has not won widespread acceptance largely because Grünbaum failed to specify what he means by a therapeutic theory and because he does not stipulate a special role for expectation effects. Grünbaum claims that placebos are treatments whose 'characteristic features' do not have therapeutic effects on the target disorder. I show that with four modifications, Grünbaum's definition provides a defensible account of placebos for the purpose of constructing placebo controls within clinical trials. The modifications I introduce are: adding a special role for expectations, insisting that placebo controls control for all and only the effects of the incidental treatment features, relativizing the definition of placebos to patients, and introducing harmful interventions and nocebos to the definitional scheme. I also provide guidance for classifying treatment features as characteristic or incidental.

Keywords Philosophy of science · Methodology · Placebo · Randomized trial · Grünbaum · Philosophy of medicine · Ethics

✉ Jeremy Howick
jeremy.howick@phc.ox.ac.uk

[1] Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, UK

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

1 Introduction

> *Much of the literature about the placebo effect is, in effect, an effort to debunk, confuse, or minimize it … Efforts to try to actually move forward our understanding of this fundamental human phenomenon are very rare* (Moerman and Jonas 2002)

There is near universal consensus within medicine that 'gold standard' evidence for the existence of therapeutic effects is provided by the randomized controlled trial and many hold that the very highest carat evidential gold is carried by those randomized trials that are also double blind *and placebo controlled*. In sharp contrast, many believe that attempts to characterise what a 'placebo' *is* have foundered, there is no agreement on what effect—if any—placebos (whatever they exactly are) have, and there is on-going controversy regarding what counts as an adequate placebo control for complex treatments such as acupuncture, exercise, and electroconvulsive therapy (ECT). The failure to characterize the placebo has added to the confusion concerning questions of whether placebos are ethical in clinical practice (Foddy 2009) and clinical trials (Howick 2009a, b). While a single conceptualization of the placebo could help resolve all these problems, I will not assume this, and I will begin with the problem of designing and appraising placebo controls in clinical trials. In this paper I argue that a modified version of Grünbaum's conceptual scheme (Grünbaum 1981, 1986) is useful for providing standards for placebo controls.

I will proceed as follows: in Sect. 2 I will outline the problems with common characterizations of placebos. In Sect. 3 I explain the importance (and some difficulties) with control treatments, focusing on the importance of controlling for expectations. In Sect. 4 I outline explain Grünbaum's scheme in detail. In Sect. 5 I argue that with four modifications, Grünbaum's scheme resists my criticisms, as well as those from Greenwood (1997), Waring (2003), Hróbjartsson (2002), and Gøtzsche (1994). The modifications I introduce are: insisting on a special role for expectancy, adding 'harmful interventions', relativizing the definition of placebos to patients, and improving the definition of placebo controls to ensure that placebo controls control for all and only the effects of the incidental treatment features. A careful reading of Grünbaum suggests that the modifications may reflect his original intentions. I illustrate the usefulness of the modified version of Grünbaum's scheme with cases studies of 'placebo' acupuncture and 'placebo' vertebroplasty. In Sect. 6 I conclude that future research is warranted to explore the consequences of the definitional scheme I defend here to investigate the concept of the placebo in clinical practice, and the ethics of placebos.

2 Failed attempts to define the placebo

The Latin term 'placebo' means 'I shall please'; beyond this etymological fact, inadequate characterisations of the 'placebo' concept abound. An often-heard idea is that a 'placebo' is simply a 'dummy pill' or 'inert substance'. In 'The Powerful Placebo'—the most cited paper in the literature—Henry Knowles Beecher referred to placebos as 'pharmacologically inert substances', the administration of which, however, have 'real

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

therapeutic effects' (Beecher 1955). Without some fancy footwork regarding the term 'pharmacological', the (near) logical falsehood that 'a placebo is an inert substance with real effects' clearly threatens. In any case, the effect of applying a glycerine stick, for example, is 'pharmacologically inert' in the normal sense (in that nothing is absorbed into the blood stream), but it would surely not be counted as a placebo for chapped lips. Moreover some substances that are by no stretch of the imagination 'inert' are often intentionally prescribed simply for the 'placebo effect'. These include (regrettably) antibiotics for viral infections, sham surgery, and saline injections. Indeed as Grünbaum (1986) pointed out, even the proverbial sugar or bread pill will prove far from inert in patients with insulin dependent diabetes or with gluten intolerance, respectively.

The Oxford English Dictionary defines the placebos as a 'drug, medicine, therapy, *etc.*, prescribed more for the psychological benefit to the patient of being given treatment than for any direct physiological effect'. But this is only coherent if we presume a Cartesian distinction between mind and body, a view whose untenability every serious investigator accepts, yet which nonetheless continues to cloud much thought in this area. Even if we go along with the idea of a psychological/physiological distinction, the OED definition has the unacceptable consequence that any psychotherapeutic intervention—for example the administration of an antidepressant—*automatically* counts as a placebo intervention since it 'is prescribed … for the psychological benefit to the patient…'. Of course it is *possible* that some particular anti-depressant is a 'mere placebo' (Kirsch et al. 2008)—assuming I can in the end make sense of this notion—but this surely is to be decided by fact not definition. Finally, importing into the definition the *reasons why* a treatment is given is a mistake: the intentions of a clinician are one thing, the objective facts about physical processes another (though one hopes that the two are at least sometimes linked). So, for example, and assuming for the time being that there is a clear-cut notion of placebo, a homeopathic treatment surely cannot be ruled out as a placebo simply on the grounds that the homeopath prescribes it in the belief that it will have a 'direct physiological effect' and therefore with the 'intention' that it will have such an effect.

Arthur Shapiro made a number of often cited, but unsuccessful attempts to characterise the placebo in the 1970s. According to his 1978 characterisation (with Morris), claims that a placebo is any therapy or component of therapy that is deliberately used for:

> … its non-specific, psychological, or psychophysiological effect, or that is used for its presumed specific effect, but is without specific activity for the condition being treated. (Shapiro and Morris 1978)

There is, again, an unfortunate (though here readily eliminable) running together of epistemic and objective issues, and an unfortunate identification of 'non-specific' and 'psychological, or psychophysiological'—the latter conflation again implying that any (successful) psychotherapeutic intervention should count as a placebo. Indeed Irving Kirsch suggests just this, namely that all forms of psychotherapy are 'placebos' by definition (Kirsch 2005, p. 801). Whether Kirsch's proposal is defensible depends on whether a acceptable definition *necessarily* includes all forms of psychotherapy, which I argue below it does not. But even if we remove the reference to psychological

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

or psychophysiological effects, we are still not out of the woods: what exactly does 'specific' mean? There is good evidence that various kinds of 'placebo analgesia' (a) exist and (b) operate through the release of endorphins ('natural opiates') into the bloodstream (Benedetti 2009); and this seems just as 'specific' an activity as, say, that of, assuredly non-placebic, penicillin in killing the pneumococcus. The term 'specific' is also sometimes used to denote 'well-defined', or 'quantitatively precise'. But estimates of 'placebo' effects (if we accept them) illustrate that their effects can be quantified much in the same way nonplacebo effects are quantified (Howick et al. 2013a, b; Hróbjartsson and Gøtzsche 2010).

Some researchers sidestep the definitional problem by replacing 'placebo' with other terms. In his wonderful book *Meaning, Medicine, and the Placebo Effect*, Moerman argues that 'placebo effects' should be replaced by 'meaning responses'. He supports his thesis by citing a variety of cases where 'placebos' have different effects in different settings and cultures, and where different placebo modalities (colour, shape, size) have different effects. In one such study, the causes of death in 28,169 Chinese-Americans were matched with the causes of death in 412,632 randomly selected 'white' controls. They found that Chinese-Americans died 1.3 to 4.9 years earlier than whites if they had a combination of disease and birth year considered ill-fated by Chinese astrology (Phillips et al. 1993). In another study Moerman cites, different price tags were placed on the very same placebo pills ($0.10 and $2.50). The 'expensive' pills were shown to have greater analgesic benefits than the 'cheaper' pills (Waber et al. 2008). The effect in the Chinese astrology study is difficult to explain with conventional theories, and the effect of the 'expensive' pill cannot be due to the pill ingredients since these were the same. Moerman therefore attributes the effects to the 'meaning' of the treatment. He defines the *meaning response* as 'the psychological and physiological effects of *meaning* in the treatment of illness' (Moerman 2002).

But meaning will not do as a replacement for placebo for several reasons. For one, Moerman's understanding of the term 'placebo' appears at times to be mistaken. To wit, he uses the term 'inert' and 'specific' to describe 'placebos' and 'specific' to describe nonplacebos (Moerman 2002, p. 16). I exposed both of these to be erroneous above. Perhaps the most serious problem with Moerman's account is that conditioning and expectancy theories can account for all the phenomena Moerman describes in his book. It is beyond the scope of this paper to examine all the examples in Moerman's book, yet certainly expectancy can explain the examples of expensive pills and Chinese astrology described above. People expect more expensive pills to be more effective, and this can activate the neuronal reward mechanisms, reducing pain, anxiety, and a variety of other symptoms (Benedetti 2009). (Or, feeling that they *should* get better with more expensive pills, patients may report feeling better after taking the more expensive pills even if they do not feel any better.) Similarly, Chinese–Americans who have strong beliefs about the seriousness of the disease, given their astrological birth sign, could expect to have a negative outcome and adopt more fatalistic attitudes. Negative effects of placebos are often referred to as 'nocebo' effects. The fatalistic attitude could lead to refusal to take or adhere to treatment regimens as well as to effects on endogenous physiological processes, particularly through the immune system. Failure to adhere to treatment regimens has been shown to be an independent predictor of clinical outcomes (Simpson et al. 2006).

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

Unlike the meaning hypothesis, which Moerman himself acknowledges has not been tested directly in any experiments, conditioning and expectancy have been tested and confirmed in hundreds of studies starting with Pavlov's famous experiments. For example people feel stimulated when given what seems to be their favourite coffee, even if it had in fact secretly been replaced with decaffeinated coffee (Kirsch and Weixel 1988). This effect, it seems, can only be explained by those people's expectations. Numerous studies have examined expectation mechanisms (Benedetti 2009) and their clinical effects (Di Blasi et al. 2001). By contrast the term 'meaning' suffers from the problems listed above and is by Moerman's own admission unsupported by direct empirical tests.

Other researchers have replaced the term 'placebo' with 'context' to solve the definitional problem. In a widely cited paper adopting this approach, di Blasi et al. state:

> Such debates [about placebo effects] are understandable given the conceptual and operational difficulties associated with the term 'placebo effect' In this study, we use the neutral and broader term 'context effects' to refer to placebo effects deriving from patient—practitioner relationships. (Di Blasi et al. 2001).

But if 'context' is intended to replace 'placebo', and 'context' is defined as a 'placebo' it is unclear whether Di Blasi et al.'s strategy disambiguates the 'placebo' concept. Another problem is that their definition of 'context factor' is internally inconsistent because they include as 'context factors' some features that do not derive in any straightforward manner from patient-practitioner relationships. Factors influencing context effects include treatment characteristics (e.g. colour, size, shape, and price of pill), patient characteristics (e.g. beliefs, anxiety levels), patient-practitioner relationship (involving, e.g., empathy, compassion, suggestion), healthcare setting (room layout, home, hospital), and practitioner characteristics (status, sex, beliefs). Categorizing these factors is undoubtedly important, and I shall illustrate below how Grünbaum's scheme requires it. However the size, shape, colour, and price of a pill have little to do with the patient-practitioner relationship (the criteria for counting as a 'context factor'). Also if we accept the suggestion that context effects are placebo effects derived from patient-practitioner interaction, we are faced once again with the threat that all forms of talking therapies be categorized as placebos a priori.

In view of all this confusion about what would count as a placebo, it is again perhaps not surprising that the suggestion has recently arisen that there is no real concept of 'placebo' to be analysed. So for example, Gøtzsche concluded a study of 'The logic of the placebo' as follows: 'the placebo concept as presently used cannot be defined in a logically consistent way and leads to paradoxes' (Gøtzsche 1994). Gøtzsche allows that the term should nonetheless 'probably' be retained for pragmatic reasons to do with entrenchment of usage. Thus in his much-cited study with Hróbjartsson, he decided—in view of all the conceptual confusion—simply to adopt a 'practical' approach and characterize placebos 'practically as an [any!] intervention labelled as such in the report of a clinical trial.' But it hardly needs remarking that this approach is untenable. Suppose for example that someone reported using penicillin as a 'placebo' in a trial of some new antibiotic as a treatment for pneumonia. The response will of course be 'no one would, and if they did we would not take the trial seriously'. But this

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

reaction seems exactly to show that we work with some concept that involves judgments about what can and cannot count as 'appropriate' or 'legitimate' placebos and placebo controls. Moreover critics have complained that Hróbjartsson and Gøtzsche's 'practical approach' led to a mistaken estimate of 'placebo' effects precisely because of their failure to put strictures on what counts as a 'placebo'. Kirsch (2002), for example, notes that Hróbjartsson and Gøtzsche jumble (along with placebo pills and injections) relaxation (described as a 'placebo' in some studies and a treatment in others), leisure reading, answering questions about hobbies, newspapers, magazines, favourite foods and sports teams, talking about daily events, family activities, football, vacation activities, pets, hobbies, books, movies, and television shows as placebos. It is clear that if the classification of these treatments as 'placebos' is mistaken, then their estimates of 'placebo' effects is also likely to be mistaken.

In addition, Hróbjartsson and Gøtzsche go back on their alleged policy of accepting any treatment labelled as a 'placebo' in the report of a clinical trial. They, for example, exclude studies where 'it was very likely that the alleged placebo had a clinical benefit not associated with the ritual alone (e.g. movement techniques for postoperative pain)' Hróbjartsson and Gøtzsche (Hróbjartsson and Gøtzsche 2001, p. 1595). Here they seem to sneak in a definition of placebos as the effects of 'rituals', which is no improvement on earlier definitions: ritual feasting or fasting are not placebos.

The philosopher of science, Robin Nunn, is braver than Gøtzsche. Writing in the *BMJ* Nunn suggests that the linguistic confusion I have partially mapped is irredeemable: 'every way of looking at the placebo concept invites criticism, because it doesn't make sense' (Nunn 2009). According to Nunn, the difficulties in characterising the placebo concept should make us question if there is any such thing 'out there' to be adequately conceptualized: if something cannot be defined and does not make sense no matter how it is viewed, it's time to ask if it is really there at all. Nunn's view is that 'it' *isn't* 'really there': the term 'placebo' does not cut Nature at any joint. Examining the diverse variety of treatments that carry the label 'placebo' one is tempted to concur with Nunn because it is difficult to see what feature, if any, they share. Lactose pills, saline injections, sham devices, sham surgery, attention controls (sham talking therapy that involves listening but not reacting), sham manipulations of the body, and many other treatments have been dubbed as 'placebos' (Howick et al. 2013a, b). With that in mind Nunn suggests that medical science would be much improved and clarified if placebo-talk were eliminated altogether.

Turner (2012a, b) supports Nunn and argues that the *purpose* of placebo controlled trials is to create trials with two groups that are treated the same way *apart* from the fact that one receives an experimental intervention, while the other does not. He claims that his idea can be summed up by the following quote from Bradford Hill:

> To some patients a specific drug is given, to others it is not. The progress and prognosis of these patients are then compared. But in making this comparison in relation to the treatment the fundamental assumption is made—and must be made—that the two groups are equivalent in all respects, except for the difference in treatment (Hill 1951)

Turner's insight that we must think of the *function* of placebo controls in order to help constrain what 'legitimate' placebo controls are, is very useful, and one that Grünbaum

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

himself advocated. Moreover both Turner and Nunn are correct that adequate descriptions of treatments are required (Hoffmann et al. 2013, 2014; Howick 2009a, b). Once we have described the features of the treatment, to drop the term 'placebo' altogether, Turner argues. Yet does not follow from the fact that adequate descriptions of terms are useful, and that they can, in principle, be replaced by the descriptions, that we should give up on trying to provide an adequate characterization of a term. In fact a philosopher's role is precisely to clarify terminology where this is possible.

In short, dropping the term 'placebo' is too quick. For one, substantive issues lurk amidst this linguistic and conceptual confusion, as we shall see. Besides the concern about whether all effects achieved by so-called 'complementary and alternative medicine' (CAM) are 'merely' placebo effects, but moreover, as my initial remarks about the connection with randomized trial methodology indicate, important epistemic and ethical issues are involved along with the conceptual ones. Simply dropping the term will not make these issues go away. In addition neither Moerman nor di Blasi nor Nunn nor Turner show any evidence that they have considered Grünbaum's scheme. This is not a criticism of their proposals *per se*, but certainly suggests that Grünbaum's proposal must be considered before we accept dropping the term 'placebo' or replacing it with a different term. Grünbaum's proposal has also generated an on-going debate (Greenwood 1996, 1997; Waring 2003). Hróbjartsson admits it is 'by far the best proposal' (Hróbjartsson 2002, p. 432), yet rejects it—claiming it fails to be 'satisfying' (Hróbjartsson 2002, p. 432), mainly because Grünbaum fails to explicate what he means by a therapeutic theory. (Yet, somewhat ironically, Hróbjartsson Gøtzsche sidestep the problem by making a similar error by not—at least explicitly—putting any restrictions on what counts as a placebo control!) It is especially odd that neither Nunn nor Turner considered Grünbaum's analysis seriously because Grünbaum shared the view that currently used definitions are unacceptable. Referring to the various definitions on offer, Grünbaum reported uncovering a 'veritable Tower of Babel' (Grünbaum 1986). If I can defend an account of the 'placebo', therefore, then the premise of Nunn and Turner's arguments can be rejected and there is no need to drop the term.

It seems that the correct strategy for the philosopher is therefore to try again: to try to produce an acceptable account of placebos that does not fall prey to linguistic confusions. This is the task I undertake in this paper—building on Grünbaum's analysis, which Nunn and Turner ignore and which Hróbjartsson cite as 'by far the best proposal so far' but then go on to reject as 'unsatisfying'. This task of providing an adequate account of the notion of a placebo, I believe, goes beyond an exercise in analytic rigour (as important as that might be in itself), but also could have practical implications for clinical trial design. Before examining Grünbaum's proposal in detail, however a few words about the difference between 'placebos', 'placebo effects', and 'placebo controls' are required.

3 Placebo controls

There are three related but different notions in need of analysis: 'placebo', 'placebo effect' (or 'placebo response') and 'placebo control' (as employed in some clinical

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

trials). It might seem that logic dictates that we first decide what a placebo is (as linguistically it is a component in the other two concepts) and then we would be on the home straight: a placebo effect is an effect produced by a placebo and a placebo-controlled clinical trial is one in which the patients in the control arm are given a placebo. In fact however linguistic appearances are misleading here. It makes perfect sense to talk of a placebo effect when no placebo is involved, as we shall see; and moreover placebo controlling a trial has a methodological justification that is independent of whether or not the patients in the control arm of that trial in fact experienced any placebo effect. The place to start, I believe, is therefore with the notion of *placebo control*.

To see this clearly, let's first ask: why should clinical trials be controlled at all? Controlling a clinical trial involves looking for real evidence for the effectiveness of the treatment on trial by eliminating other plausible explanations of a possible positive result. So suppose, to take the hackneyed example, we are interested in whether taking regular vitamin C is an effective treatment for the common cold. The first suggestion might be to give vitamin C to a bunch of people suffering from colds and see what happens. Suppose that they all recover within five days. Although this result is certainly compatible with the 'vitamin C is effective' hypothesis, background knowledge tells us that colds often clear up within five days without any treatment. So the result fails to count (or at any rate, fails to count at all significantly) in favour of the vitamin C hypothesis because it fails to count *against* at least one (very) plausible rival: the natural history hypothesis. To test the 'vitamin C is effective' hypothesis, we need to control for 'natural history'. That is, we need a control group of patients with colds who are *not* given vitamin C.[1] Mackie (1974) expresses this intuition very clearly in reference to Mill: 'all these [Mill's] methods work by eliminating rival candidates for the role of cause'. The ideal (and in reality impossible) control group would involve comparing the effects of an intervention (say, vitamin C) in one person with the (hypothetical) counterfactual case where the very same person at the very same time did not take vitamin C, then compare the outcomes.

As a surrogate for the practically impossible, *control groups* are used. But of course there are an infinite number of differences between any two groups (or indeed people, or even the same person at different times). So the best we can do is ensure that the groups are 'equivalent' in terms of various factors that background knowledge suggests might make a difference. So for example the relative severities of the colds, the age distribution, the general health of the people in the two groups, and so on, should be at least closely the same in the two groups. Otherwise, if those in the experimental group were considerably younger on average than those in the experimental group, then a 'positive' result would produce very questionable evidence for the effectiveness of vitamin C, since background knowledge supplies a plausible alternative explanation of such a positive outcome-older people may tend to find it more difficult to 'shake off' colds than younger people.

[1] Note however that if the effect of the vitamin C were very large—suppose for example that everyone's cold cleared up *immediately*, then we could be quite certain that vitamin C caused the recovery. This judgment would be justified by a hypothetical control group: background knowledge informs us that colds take several days to go away on their own and without treatment.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

Of course there may be other factors—unknown (possible) confounders: factors which may affect recovery but which background knowledge gives us no reason to suppose do so. Clearly we cannot intentionally control for 'unknown' factors since they are unknown. Let's assume for the sake of the argument in this paper that, as is widely believed, using a randomizing device to decide which of the two matched blocks becomes the experimental group and which the control group, helps create similar groups (see Worrall 2002).

Surely at last a positive result in this ultra-controlled trial tells unambiguously in favour of the efficacy of the treatment? Going along with the idea that the randomization has controlled for unknown confounders, the positive result must be due to the treatment—all other possible rival explanations have been eliminated through making the two groups 'otherwise equivalent'. Not quite. The way I have envisaged it so far, both those involved in the trial and the administering clinicians know which is the experimental and which the control group (because only those in the treatment or experimental groups are given any 'medication'). But this knowledge can lead to confounding in the 'treatment' phase *even if the groups were equivalent at the outset*. Suppose, for example, that the clinicians are all members of the Linus Pauling Fan Club and really hope for a positive result for vitamin C. They—perhaps subconsciously—lavish a great deal of attention on those in the treatment group, but fail to engage with those in the control group. Obviously this potentially invalidates the trial—again because it makes plausible an alternative explanation of any superior outcome (or at any rate any superior outcome that is moderate in size): those in the experimental group might have had a better average outcome, not because of anything attributable to the vitamin C they ingested, but instead because the attention they received made them feel better about themselves in general. Clearly, the intervention (including any additional care provided) beyond the substance being tested must be (at least to a good approximation) the same in both groups.

Just as doctors' behaviour can introduce differential treatment to experimental and control groups, and thus introduce alternative hypotheses for any perceived differences, so can patients' beliefs and behaviour. If a patient knew they were being left untreated (or indeed were being treated by a 'mere' placebo), they might covertly seek concomitant medication. Similarly the patients—especially those whose symptoms are more severe—might simply drop out of the trial. Differential rates of taking concomitant medication and differential dropout rates (especially if dropping out is related to the severity of symptoms) are potential confounders. Moreover the subjects taking the 'real' treatment know they are being given a 'real' treatment so expect to feel better; whereas those in the control group, who know they are missing out on the latest treatment (and taking a 'mere placebo'), are less likely to have any special expectation of an unusual improvement. This is not a mere philosopher's possibility: a growing body of evidence suggests that increased attention has a positive benefit, at least for some disorders (Kaptchuk et al. 2008). Indeed the recognition that some treatments may be efficacious simply though patient expectancy of improvement goes back to Hippocrates who stated: 'Some patients though conscious that their condition is perilous, recover their health simply through their contentment with the goodness of the physician.' And it was of course this challenge to Freud—that the efficacy of psychoanalytic treatment has nothing whatsoever to do with Freudian psychoanalytic

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

theory but rather had to do with the patients' beliefs that psychoanalysis might make them improve—that led to Grünbaum's resurrection of Freud's 'tally' argument and his consequent work on placebo controls.

One way to ensure similar care for the two groups might be to provide and enforce explicit protocols. But there is the view (how solidly based in previous real experience is another question) that, these things being very subtle, it is possible that—perhaps even unconsciously—such clinicians, while trying their best to be even-handed, in fact allow their own expectations of a better outcome in the experimental group to influence how they treat patients, and how they assess outcomes. This is especially worrying if outcomes are subjective.

The way to eliminate these further confounding differences in the interventions in the experimental and control groups that has been adopted in medicine is to 'blind' or 'mask' caregivers and participants with respect to which is the experimental and which is the control treatment (Howick 2011). If the caregiver doesn't know whether or not she is providing vitamin C or a control treatment, then she cannot provide different care to the experimental group. Likewise, if a participant doesn't know whether he is receiving the experimental treatment, he will have no reason to behave differently in ways that might confound the study, and his expectations regarding the likelihood of recovery will be the same.

But how do we blind caregivers and participants? Assuming the requirement of informed consent, the only way seems to be to give some 'treatment' to those in the control group as well—one that, so far as those receiving it are concerned, is indistinguishable from the treatment given to those in the experimental group. Such a control treatment in the pretend case would have to be the same as the treatment given to the patients in the experimental group apart from the fact that it contained no vitamin C. If, to preserve outward appearances of similarity, a bulking agent, for example, had to be added to the control treatment, then it should not contain any substances that affect recovery apart from vitamin C.[2] If it did then clearly it would 'over control' the study and raise the possibility of falsely inferring that vitamin C is inactive.

By keeping the intervention in both groups similar, blinded studies involving 'dummy' treatments control for potential *expectation effects*. We all can remember occasions when we have been feeling pretty good about things in general and to have shaken off colds more readily than normal and other times when we have felt comparatively miserable and the cold has seemed to linger on and on. Obviously expectations of a positive outcome are likely to be higher amongst those who know they have received the experimental treatment (unless they are unusually well-informed about the history of medicine) and it may be these expectations rather than anything distinctive about the vitamin C (as we will see below the distinctive features of a treatment are referred to as 'characteristic features' by Grünbaum) that were responsible for the positive average result. Empirical evidence supporting the claim that expectations can have effects is growing (Schulz et al. 1995; Savovic et al. 2012; Wood et al. 2008). The likely explanation of the improved results in the non-blind trials seems clearly to

[2] There are cases where bulking agents have had unexpected effects (Golomb 1995, 2002, 2009).

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

be that expectations played a role. It follows that the general philosophy of science principle requires that these expectations be controlled for.[3]

Nunn or Turner might, of course, object that we should call expectation effects expectation effects, and expectation controlled trials expectation controlled trials rather than using the potentially ambiguous term 'placebo'. Certainly they are correct that we should be clear about what we mean by placebos and that we should describe the placebos adequately. And my discussion above also suggests that any account of placebo controls needs to take expectations into account. However Nunn and Turner's general conclusion that we should drop the term 'placebo' only follows if we can't make sense of the term, which I claim to do in the remainder of this paper. Moreover the fact that a concept is ambiguous is not, in itself, a sufficient reason for removing them from our vocabulary. The term 'medical treatment' is ambiguous in much the same way the term 'placebo' is ambiguous. It does not follow that the term 'medical treatment' should be dropped. Moreover Nunn and Turner's potential suggestion that we should replace 'placebo control' with 'expectation control' also cannot be defended. I argue in Sect. 5.2 that expectations are not always necessary to control for and rarely sufficient.

Note that whether you regard a particular placebo control as adequate in some particular trial may depend on what theory you hold. Let's go back to the example of acupuncture for the treatment of pain. Suppose a practitioner holds the theory that inserting acupuncture needles to a certain depth is indeed efficacious for, say, back pain—but only if the needles are inserted at the corresponding 'chi' ('Qi' or 'acupuncture') points as specified by the theory of acupuncture supplied by traditional Chinese medicine. Call this 'real TCM acupuncture'. This person would be committed to the view that any effect on back pain of inserting acupuncture needles at points of the body other than the chi points are placebogenic. Hence for her a trial in which the experimental group receive real TCM acupuncture, while the control group receive treatment that involves the insertion of acupuncture needles to the same depth but at points other than the chi points is a placebo controlled trial. On the other hand, another practitioner who holds the different theory that inserting acupuncture needles always has some (overall) positive effect on back pain distinct from any expectations aroused, would *not* regard this trial as placebo controlled. For this second practitioner, a genuine placebo controlled trial would have to be one in which no needle was actually inserted.

It was with this point in mind that the Streitberger needle (a sham acupuncture needle that gives the appearance of penetrating the skin but that in fact does not—I will describe it in more detail below) was developed (Streitberger and Kleinhenz 1998). However whether even this is a placebo-controlled trial again depends on the exact theory held. If my second practitioner indeed holds the theory that actual insertion is necessary for any non-placebo-generated effect then this is indeed a placebo-controlled trial for her. Suppose however a third practitioner holds the different theory that simply applying needles to a person's skin has *some* non-'placebo'-generated effect (through 'acupressure') This third practitioner would not then view the trial in which control patients are treated with the Streitberger needle as fully placebo-controlled (though

[3] Medical scientists often talk of the non-blind studies 'exaggerating' the benefits of treatment, but since if the treatment is approved it will be carried out by practitioners in a non-blind way the non-blind results may well in fact give a more accurate measure of the 'real result' in 'the wild'.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

if she held the theory that acupuncture in any form has *greater* effect on pain than acupressure then she would expect a positive result in what would for her be an 'active treatment trial' (Moncrieff et al. 2004)).

This discussion shows, then, that one of Grünbaum's key insights in characterising the notion of a placebo (namely that the notion is implicitly relativized to therapeutic theory) certainly holds for the notion of 'placebo control'. Let's then turn to an explicit examination of Grünbaum's analysis to see if the account of placebo controls I have just developed is reflected in his definitional scheme.

4 Grünbaum's definitional scheme

Grünbaum offers two main insights that help clarify the placebo concept for the purpose of defining placebo controls. First, he suggests that the notion of a placebo needs to be doubly relativized—*first* to the condition treated (the effects, if any, of penicillin on flu are placebo effects, but the effects on bacterial pneumonia are not) and *secondly* to therapeutic theory. Grünbaum highlights the importance of relativizing to a disorder *D* using the well-worn example of a sugar pill:

> … none other than the much-maligned proverbial sugar pill furnishes a *reductio ad absurdum* of the notion that a medication can be generically a placebo *simpliciter*, without relativization to a target disorder. For even a lay person knows that the glucose in the sugar pill is anything but a generic placebo if given to a victim of diabetes who is in a state of insulin shock, or to someone suffering from hypoglycaemia. (1986, p. 35).

The need for the latter relativization should be clear from the above acupuncture discussion and is also strongly underlined by consideration of tests of psychotherapeutic claims. Grünbaum's claim is that an intervention operated as a placebo just in case the intervention made a difference but this difference was achieved via the treatment's 'incidental' features rather than its 'characteristic' features. Which of the treatment's features are seen as 'characteristic' and which 'incidental' will, in general, depend on what therapeutic theory is brought to bear. Hence what counts as a placebo control must be relativized to theory as well as disorder.

Often in 'somatic medicine' as Grünbaum calls it (though this again tends to encourage unfortunate dualist tendencies), there is so little controversy over the therapeutic theory presupposed that it might seem artificial to talk about a theory at all. To take an example that Grünbaum cites, the 'theory' that underwrites accepted treatment for gallstones will clearly make the surgical removal of the gallstones as characteristic. Other features such as the surgical consultation, the analgesia, etc. would be classified as incidental. But in the psychotherapeutic field the dependence on theory is often crucial. There, which aspects of a particular interaction with a patient are characteristic will clearly be theory-dependent so that one and the same feature of a given interaction may be judged characteristic by one theory and incidental by another. For instance, according to Freud the characteristic features of nonpharmacological treatment included lifting a patient's presumed repressions, while the incidental features included the patient's faith in the analyst, emotional support from an authority figure,

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

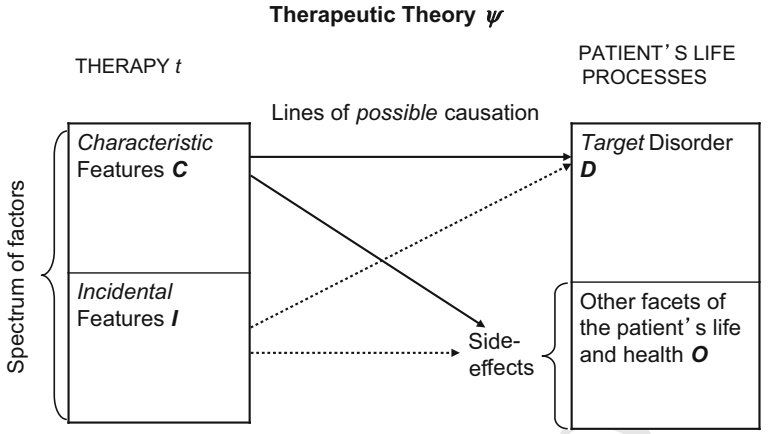


Fig. 1 Illustration of therapeutic theory ψ, used in clarifying the definition of 'placebo'. (Based on Grünbaum 1986, p. 22)

and the payment of a hefty fee (Grünbaum 1986, p. 24). Yet more pragmatic forms of talking therapies, such as cognitive behaviour therapy (CBT), do not regard these as characteristic. A problem with Grünbaum's scheme that I discuss below is that he fails to constrain therapeutic theories (and hence what counts as a characteristic feature).

Notice that Grünbaum's analysis has the (surely welcome) consequence that a treatment may be a nonplacebo overall and yet involve placebo features. This will occur whenever an overall treatment effect is achieved in part by the treatment's characteristic and in part by its incidental features. Grünbaum records that, for example, there is evidence that chemotherapy for certain kinds of cancer may produce enhanced positive effects if administered by an enthusiastic physician. The theory of the direct physiological effects of chemotherapy on tumours (not mediated through increased expectations) would then dictate which features of the overall treatment are 'characteristic'.

Another example will help illustrate this point. A therapeutic theory may state that the therapy t is the administration of Prozac according to some given regime, the target disorder D being major depressive disorder (MDD). The therapeutic theory might also specify that the chemical fluoxetine hydrochloride is the 'characteristic feature', C, of this therapy. The incidental features, I, of the therapy might include pill bulking agents, the potential disruption to the patient's life (they must take time every day to consume the pills), ingredients in the pill casing, the liquid with which the pills are swallowed, and perhaps most importantly expectations about the potential effects of fluoxetine hydrochloride and the patient/doctor interaction. The fact that all treatments, including apparently simple ones, have several treatment features is obscured by ordinary language. For example, it is common to refer to 'Prozac' as a treatment when what is actually meant is 'therapy involving fluoxetine hydrochloride, and that also includes other ingredients in the pill, the liquid with which the pill is swallowed, the beliefs and expectations of the patient, the label on the pill, etc'.

The details of Grünbaum's scheme are best explained with the aid of a diagram (see Fig. 1). Beginning with the left-hand box in the diagram, we see that the therapeutic

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

theory, ψ, differentiates between characteristic (C) and incidental (I) features.[4] Even pill treatments that are often considered simple have several components, as we saw with the example of *Prozac* therapy above. For example, a therapeutic theory may state that the therapy t is the administration of Prozac according to some given regimen, the target disorder D being major depressive disorder (MDD). Other features would then be characterized as incidental.

The four arrows in the diagram represent *possible* effects. The top horizontal arrow represents the possible effect of the characteristic factors C on the target disorder D. The arrow that runs from the upper left to the lower right represents the possible side effects of the characteristic factors. The lower horizontal arrow represents potential effects of the incidental factors I on O, while the arrow from the bottom left to the upper right represents possible effects of the incidental factors I on the target disorder D. The four arrows of possible causal influences can be positive, negative, or, in some cases 'empty' i.e. represent no effects at all. Henceforth when speaking about effects of features (both incidental and characteristic), I am referring to *possible* effects unless otherwise specified. With the conceptual scheme in mind, Grünbaum defines placebos and related terms.

Nonplacebo a treatment process t is a nonplacebo for target disease D '*if (and only if) one or more of the characteristic factors do have a positive therapeutic effect on the target disease D*' (Grünbaum 1986, p. 23, italics original).

Hence the key feature of a nonplacebo is that its characteristic features must have a *positive* therapeutic effect on the target disorder D. The administration of Prozac therapy, would thus be characterized as a nonplacebo for depression if and only if fluoxetine hydrochloride had some *positive* therapeutic effect for depression. Grünbaum then proceeds on this basis to characterise notion of placebos and related terms:

Generic Placebo a treatment process t is a generic placebo if none of the characteristic treatment factors C are remedial for D (Grünbaum 1986, p. 33). Generic placebos come in two types: intentional and inadvertent.[5]

Intentional placebo a treatment process t is an *intentional placebo* if and only if it satisfies the following four conditions—the fourth normally holding but, strictly speaking, being optional:

(a) t is a generic placebo,
(b) the practitioner believes that the characteristic factors C all fail to be remedial for D (the practitioner believes that t is a generic placebo),
(c) the practitioner believes that some patients will benefit from the treatment due to one or more of its incidental features,
(d) [optional] the practitioner 'abets, or at least acquiesces in, [the patient's] belief that t has remedial efficacy for D by virtue of some constituents that belong to the set of characteristic factors [C]' (1986, p. 24).

[4] Grünbaum uses F to designate characteristic factors and C to designate incidental factors. I use the more natural 'C' and 'I'.

[5] Henceforth unless otherwise specified when referring to the term 'placebo' I mean 'generic placebo'.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

Inadvertent placebo a treatment process *t* is an *inadvertent* placebo if and only if it satisfies the first two of the following three conditions—the third normally holding but, strictly speaking, being optional:

(a) t is a generic placebo,
(b) the practitioner believes that some of the characteristic features *C* are remedial for *D*,
(c) [optional] the patient believes that the remedial effects on *D* are due to some characteristic feature of the treatment *t*.

Placebo effect a placebo effect is either (a) one produced by the incidental features of some treatment (even when the treatment as a whole is a nonplacebo), or (b) any effect of a generic placebo. In Grünbaum's words:

> On the basis of the explications I have given, it is appropriate to speak of an effect as a 'placebo effect' under two sorts of conditions: (a) even when the treatment [process] t is a nonplacebo, effects on *D*—be they good, bad, or neutral—that are produced by *C*'s incidental factors count as placebo effects, precisely because these factors wrought them; and (b) when t is a generic placebo whose characteristic factors have harmful or neutral effects on *D*, these effects as well count as placebo effects. Hence, if *t* is a placebo, then all of its effects qualify as placebo effects. (Grünbaum 1986, p. 32)

Placebo control a placebo control is an intentional generic placebo that is generally harmless. In Grünbaum's words:

> A treatment type *t* functions as a 'placebo control' in a given context of experimental inquiry, which is designed to evaluate the characteristic therapeutic efficacy of another modality t* for a target disorder D, just when the following requirements are jointly satisfied: (1) *t* is a generic placebo for D, as defined under the first condition (a) in the definition above of' 'intentional placebo'; (2) the experimental investigator conducting the stated controlled trial of t* believes that t is not only a generic placebo for D, but also is generally quite harmless to those victims of D who have been chosen for the control group. (Grünbaum 1986, p. 26)

It is immediately clear how Grünbaum's scheme solves many of the problems with previous attempts at defining the placebo. The scheme allows for placebos to be active and have specific effects, provided that the characteristic features do not cause these effects. It also allows for psychological factors to be both placebic and nonplacebic (it depends on the therapeutic theory).

At the same time there are several problems with Grünbaum's scheme, some of which have been noted by critics such as Greenwood (1997) and Waring (2003). These include:

(1) Grünbaum fails to define characteristic features,
(2) Grünbaum's definition do not allow for any intrinsically privileged role for expectations,
(3) Grünbaum's explicit definition of placebo controls does not require inclusion of *all* incidental features,

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

(4) Grünbaum allows harmful interventions to be classified as placebos,
(5) the definitions should be, but are not, explicitly relativized to individuals.

Each of these objections warrants a clarification to Grünbaum's original scheme.

5 Problems with Grünbaum's scheme, and suggested solutions

5.1 Answering Greenwood's objection that Grünbaum allows pharmacologically active treatment features to be characterized as placebos

Greenwood argues that Grünbaum's concept of the placebo has the absurd consequence of allowing pharmacologically active substances to be regarded as placebic. If a factor in t is declared 'incidental' by ψ but is pharmacological rather than psychological while none of the factors of t declared characteristic by ψ has any effect, than t counts as a placebo on Grünbaum's scheme. This, says Greenwood, violates our intuitions:

> Consider the hypothetical case of a drug treatment [process] t for disorder D. According to therapeutic theory T of drug treatment [process] t for disorder D, the pharmacological components a, b, and c are "characteristic" or "active" components [C]; the pharmacological components d and e are "incidental" or "inert" components [I]. Say it turned out to be the case that components a, b and c are not remedial for D, but that component e alone is responsible for the total remedial effect. In this case, where the effect is produced by pharmacological component e alone, we would have an instance of a placebo effect, according to Grünbaum's definition *even though no part of the effect is produced by psychological factors such as therapist/doctor commitment or client/patient expectancy*. I think that to call such a*pharmacologically produced* effect a "placebo effect" is a misuse of language. Any account that has such as consequence is off to a very bad start (Greenwood 1997, p. 500, emphasis original).

Although Greenwood does not provide a real example to illustrate the apparently unhappy consequences of Grünbam's scheme, he surely has in mind a case such as the following. Imagine some treatment for bacterial pneumonia had the following treatment features:

a: the pill casing,
b: a bulking agent,
c: water with which the pills are swallowed,
d: patient/doctor expectancy,
e: antibiotics.

Imagine further that the therapeutic theory classified *d* and *e* as incidental while *a*, *b*, and *c* were classified as characteristic. Grünbaum's scheme would refer to this treatment as a 'placebo' for treating pneumonia, and this would be a misuse of language. To be sure the example of antibiotics is loaded—'antibiotic' is a heavily theory-laden term—substances don't just come with 'antibiotic' written on them. A better example might be to replace 'antibiotic' with 'pharmacological substance X'. In an actual example of a mistakenly labelled *incidental* feature, olive oil was once used in placebo

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

capsules for trials of cholesterol-lowering agents before there was evidence that olive oil reduced cholesterol (Golomb 1995). Although olive oil was not considered characteristic by the therapeutic theory at the time, it may have had effects nonetheless. The therapeutic theory, in the case of substance X and (in the past), olive oil, failed to correctly identify the characteristic features.

Greenwood's argument reveals the serious problem that Grünbaum fails to place any strictures on what counts as a therapeutic theory (and hence what can legitimately be classified as a characteristic feature). Hróbjartsson (2002), and Walach (2011) also note this problem. At least in principle, antibiotics *could* mistakenly be classified as incidental for treating bacterial pneumonia on Grünbaum's scheme, which seems absurd. The failure to constrain therapeutic theories can also lead to mistaken classifications of treatments as nonplacebos. Imagine we design a treatment that involves a saline injection and a positive and deceptive suggestion (telling a patient that the injection 'involves a powerful drug that is very effective') for treating pain. Imagine further that we classify the saline as incidental and the positive suggestion as characteristic. Background knowledge tells us that the 'characteristic' feature of such a treatment is likely to be effective, leading one who adheres strictly to Grünbaum's scheme to classify the treatment as a non-placebo. This seems absurd.

The apparently absurd consequences of Grünbaum's failure to put strictures on what counts as a characteristic feature is serious, and can be solved by appealing to the importance of controlling for expectancy. To solve this problem, I will therefore define a characteristic feature a feature which:

(1) is not expectancy *that* a treatment is effective,
(2) has an incremental benefit on the target disorder over a legitimate placebo control in a well controlled trial.

Since antibiotics are not expectancy, and since they have a benefit over and above antibiotic placebo, they need to be classified as characteristic. On the other hand, positive suggestions (inducing positive expectations) are not characteristic (with a possible exception, see below). It is clear from this definition that we will not always know whether a particular feature has been correctly classified until after a placebo controlled trial in which expectations that the experimental intervention are effective have been controlled for. In fact a main purpose of conducting placebo controlled trials in the first place is to determine whether interventions' characteristic features have benefits over an above 'placebo' effects. Even after having conducted a trial, however, we might have to revise the classification of a feature as incidental or characteristic.

The fact that we have to revise our classification of features as incidental or characteristic is not a problem with Grünbaum's scheme, but a consequence of the fact that *all* scientific theories being tentative and revisable in light of (hopefully reliable) new insights and evidence. Grünbaum explicitly acknowledged this: 'if some of the incidental constituents of t are remedial but presently elude the grasp of ψ, the current inability of ψ to pick them out from the treatment process hardly lessens the objective specificity of their identity, mode of action, or efficacy' (1986, p. 33). Grünbaum need merely add that in practice, some of the factors named as incidental according to a therapeutic theory would be better described, by a 'truer' theory, as characteristic. The potential necessity to revise the classification of a feature in light of evidence is not a

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

problem with Grünbaum's scheme *per se* but a problem with the fallibility of science in general. Yet Greenwood is correct that Grünbaum failed to restrict what could count as a characteristic feature, and that this is problematic. My definition of characteristic features remedies the problem.

5.2 Waring and Greenwood's objection that Grünbaum fails to make a special place for expectations

Both Waring (2003) and Greenwood (1997) complain that Grünbaum fails to capture the intuition that placebos are allegedly associated with psychological rather than incidental factors. They both suggest replacing Grünbaum's definition of placebos with one that is more closely tied to factors such as patient expectation and practitioner enthusiasm. Waring states: 'psychological factors such as a patient's expectations of benefit seem closer to what we intend by the placebo concept rather than remedial failure' (Waring 2003, p.14). Greenwood states: 'we [might] have an instance of a placebo effect, according to Grünbaum's definition, *even though no part of the effect is produced by psychological factors such as therapist/doctor commitment or client/patient expectancy*' (Greenwood 1997, p. 499, emphasis original).

To respond to this objection we must first distinguish between psychological factors in general, and expectations. If Waring and Greenwood's objection is interpreted as an argument that all psychological factors are placebos, this implies classifying all psychological therapy as placebic *a priori* which is a mistake, as we saw above. The second interpretation is that placebos must involve features such as doctor commitment or patient expectancy. In this regard I believe Greenwood and Waring are correct. Expectations deserve a special place in any account of placebo controls, and indeed elsewhere in his paper Grünbaum himself acknowledges this:

> Turning now to placebo controls, we must bear in mind that to assess the remedial merits of a given therapy t* for some [disorder] D, it is imperative to disentangle from each other two sorts of possible positive effects as follows: (1) those desired effects on D, if any, actually wrought by the characteristic factors of t*; and (2) improvements produced by the *expectations* aroused in both the doctor and the patient by their belief in the therapeutic efficacy of t*. To achieve just such a disentanglement, the baseline measure (2) of expectancy effect can be furnished by using a generic placebo t in a control group of persons suffering from D. (Grünbaum 1986, p. 26, italics added)

Unfortunately, Grünbaum's formal definition of placebo controls (as generally harmless generic placebos) fails to reflect what he writes about the importance of expectations elsewhere. There are three good reasons to support the view that (the caveat below notwithstanding) expectation effects are placebo effects. First, it captures a common intuition about what a placebo effect is. The association between placebo effects and expectation effects has been documented in historical accounts of the placebo (Kaptchuk 1998), and is reflected in Waring and Greenwood's objection. It is also arguably justified etymologically: telling someone they will get better (inducing a positive expectation) is likely to please all but the most negative people. Second, basic science evidence converges on the view that the main mechanism of

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

action of placebo treatments (however they are defined) is conscious or subconscious expectancy and subsequent reward mechanisms (Benedetti 2009). Third, the usage of placebos in clinical trials' key purpose is to keep expectations (and hence expectation effects) the same in both groups. The philosopher's job is to clarify and elucidate natural language rather than reinvent it wherever possible, and expectation effects are used in natural language as placebo effects. I therefore maintain that not specifying the special role of expectancy in an account of placebos is therefore a mistake, and my definition of characteristic features takes this into account.

It is important to note, however, that there are exceptional cases where controlling for expectations is neither necessary nor sufficient. Controlling for expectation is not necessary in at least the two following examples. The first involves unconscious patients who are given injections. Such patients would not have any expectations about the efficacy of the injection and therefore expectations would not have any effects on these patients. An unconscious patient has no (conscious) expectations by definition so these expectations do not need to be controlled for.[6] Yet placebos might affect their treatment for two reasons. First, even unconscious patients' bodies have been conditioned to respond to stimuli that have been shown to have some healing benefit (Benedetti et al. 2003). This conditioned response is an explanation for how 'open label' placebos (placebos given to patients who know the treatments are placebos) can be effective (Kaptchuk et al. 2010). Second, using a placebo control in an unconscious patient will help to rule out the potentially confounding influence of needle insertion and bulking agents, and to control for experimenter biases. Experimenter enthusiasm, for example, could have some effect on unconscious patients, and are part of what we mean when we talk about placebo effects.

There are also certain *types* of expectation that may not be placebic. To illustrate, consider the example of 'Positive Psychology' (PP). The theory behind PP is that patients should focus on the positive aspects of their lives. This encourages them to have more positive expectations. Positive Psychology therapists provide patients with cognitive tools that help them change negative thoughts and expectations into positive ones. For purposes of this argument, assume that PP's therapeutic theory classifies positive expectations about recovery arising as a result of a PP consultation as the only characteristic feature and all other treatment features as incidental. Now imagine that PP became very popular, with beautiful Hollywood stars using and endorsing it. PP's popularity could (again, at least in principle) lead to patients having positive expectations about the effectiveness of PP before even having a PP session. These positive expectations could lead to some benefit independent of the PP session itself. On the other hand, a qualified PP therapist might induce a further benefit for the patient by providing a strategy that helps them modify their thought pattern. That is, there are two potential sources of expectations that could be responsible for effectiveness of a

[6] It is important to note the difference between expectancy and conditioning. Expectancy and conditioning are activated by overlapping but different stimuli, are known to operate via different mechanisms and have different effects (Stewart-Williams and Podd 2004). But by definition in the case of the unconscious patients we are controlling for *unconscious* expectations. Unconscious expectancy is generally regarded as distinct from (conscious) expectancy and is referred to by a different name: conditioning. Hence while the conditioned response of the unconscious patient and the additional practitioner enthusiasm may be incidental (placebo) factors, they are not the same thing as (conscious) expectancy.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

PP session: (i) expectations *that* PP is effective (arising from, for example, Hollywood hype), and (ii) expectations generated in the patient by the PP therapy (arising from the things a qualified PP therapist says). Only the first, I argue, should be classified as placebic. The second has a separate cause, could have a distinct mechanism of action, and is, I submit, more accurately classified as a non-placebic (characteristic) feature.

A real example from my experience will help clarify the difference between the two types of expectation. When I was an athlete I lost a hard fought race and I wanted to win the next one. To do this I had to improve my ability to focus. I called my first coach whose name is Scott. Now Scott is a great coach and I had positive expectations that my focus would improve after talking to him. These positive expectations arose before I spoke to him and were therefore independent of anything he actually said. They were analogous to the 'Hollywood hype' in the PP example. These expectations could, at least in principle, have led to an improved focus regardless of what he said, and would legitimately be classified as placebo expectations. However in addition to the positive expectations that arose at the mere though of speaking to Scott, he gave me some cognitive tools that helped me develop positive expectations about potentially negative situations. The one I remember most is that whenever something negative happened he would remind me to tell myself that, 'adversity is an opportunity'. He then gave me examples of great athletes who used setbacks to regroup themelves and come back stronger than ever. I used these cognitive strategies (telling myself that adversity is an opportunity and recalling real cases of great athletes who had been through hard times and succeeded) to turn my negative expectations about the future into positive ones. These latter expectations that were induced by a specific cognitive tool—perhaps similar to those used by PP therapists—were independent, at least in principle, from expectations I had about the benefits of interacting with him.

Since the expectations generated by 'Hollywood hype' surrounding PP are placebic, they need to be controlled for. If the expectations induced by the PP therapist during a PP therapy session do not have any incremental benefits over and above the expectations *that* PP is an effective method ('Hollywood hype'), then any benefits of PP are not due to any 'characteristic features' of PP, but due to the expectations patients have about the effectiveness of PP, and PP can safely be classified as a placebo.

If PP effects were only due to 'Hollywood hype', one could replace an actual PP session with a 'sham' PP session and have the same results. In fact this is not the case. In one study, five PP interventions designed to induce positive expectations (showing gratitude, listing three good things in life, identifying a time when the patient did their best, identifying strengths, and using strengths in a new way) were compared with a 'placebo' control that involved writing about early memories (Mitchall et al. 2009). The patients were blinded to the treatment condition, so expectations *that* PP therapy was effective ('Hollywood hype') were the same in both groups. A systematic review including this and 38 additional randomized trials of PP found that PP outperformed the sham PP (Bolier et al. 2013). The systematic review reported that PP had a significant overall effect for subjective well being (standardized mean difference = 0.34) as well as depression (standardized mean difference = 0.23). This suggests that PP therapy has an incremental benefit over and above 'Hollywood hype' expectations and that it contains a feature that counts as characteristic according to the criteria laid out in Sect. 5.1 above.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

Besides not always being necessary, controlling for expectations is also—again albeit in exceptional cases—not sufficient. This type of case can be clearly illustrated with case studies of acupuncture and vertebroplasty.

5.2.1 Case study of acupuncture illustrating why controlling for expectations is not a sufficient condition for a treatment to be a placebo

Derived from traditional Chinese medicine, acupuncture is a form of treatment for various disorders that involves insertion of fine needles into particular 'Qi' points. The needles are very thin and usually penetrate to a depth of a quarter to three quarters of an inch (5–40 millimetres) depending on the location. The needle penetration into the skin is barely perceptible, and acupuncture is widely used. Some researchers advocate a theory involving lines of energy flowing through the body, or 'meridians' (Kaptchuk 2002). However these theories lack a widely accepted or established empirical base, at least according to conventional science. We saw above that it is possible to hold different theories about how acupuncture might work, and these different theories will lead to different specifications of what the characteristic features of acupuncture are. Still, it is possible to list common features (characteristic or incidental) of acupuncture therapy, which *might* play a role in outcome. These include:

1. Patients and practitioner beliefs about, attitude towards and expectations of relief from needling and acupuncture.
2. The acupuncture consultation.
3. Needle insertion (anywhere in the body, not at the 'acupuncture' points indicated by the relevant theory of acupuncture).
4. Needle stimulation (of acupuncture points) at what the relevant theory sees as the correct location.
5. Pressure at any point on the body.
6. Pressure *at what the relevant theory sees as the correct location*.

One device touted as a 'placebo' or 'sham' acupuncture procedure involves the Streitberger Needle (Streitberger and Kleinhenz 1998). This is a blunt needle embedded in a moveable shaft (see Fig. 2). When the device is pressed on the skin, the shaft moves and gives the appearance of penetrating the skin. In order to hold the device in place, plastic rings are taped to the patient's skin at the acupuncture points. To maintain the deception, the rings are also used for the real acupuncture. Some researchers claim that the sham needle is 'validated', by which they mean a trial involving treatment with the sham device is capable of remaining successfully double masked thus keeping expectation levels the same in treatment and control groups.[7] Hence by 'validation' they seem to mean that the Streitberger needle successfully controls for expectations that the therapy is 'real' acupuncture.

Trials comparing real acupuncture with acupuncture involving the Streitberger needle typically only show small benefits of real acupuncture (Furlan et al. 2005). At the

[7] I ignore here the issue of whether the device is indistinguishable from 'real' acupuncture and hence whether it has been 'validated' in the sense proponents claim; evidence suggests it is not (see Howick 2011).

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

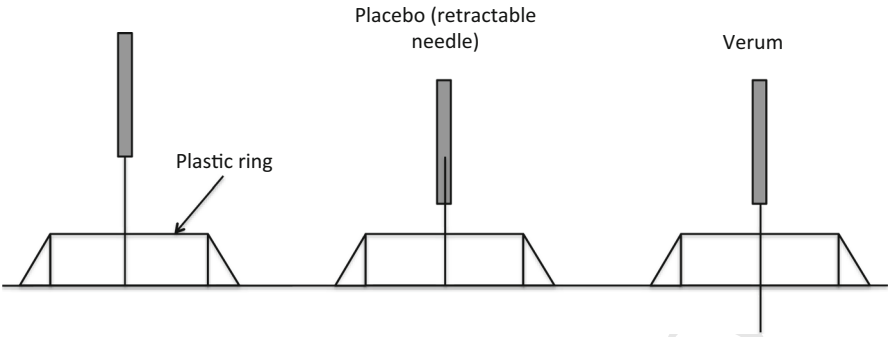


Fig. 2 The Streitberger Needle (simplified model)

same time, evidence suggests that treatment involving the Streitberger needle is more effective than placebo pills (Linde et al. 2010), while both real and sham acupuncture is more effective than conventional treatment for back pain (Furlan et al. 2005). The larger effects of the Streitberger needle compared with conventional pill placebos can be interpreted in two ways: either treatment with the Streitberger needle produces an especially large placebo effect (Ernst 2006), or it is not a 'real' placebo (Paterson and Dieppe 2005).

If we accept a therapeutic theory stipulating that needle penetration is the *only* characteristic feature of acupuncture, then the Stretiberger needle is little more than a placebo. The sham acupuncture trials certainly demonstrate that needle penetration does not add very much additional benefit. However it is also possible that a therapeutic theory classifying needle insertion as the exclusive characteristic feature is mistaken, according to my definition above. That is, according to my definition of characteristic features, the Streitberger needle might include some features that are best described as characteristic. To see how, recall the case of the polypill cited above. A control treatment for the polypill that is the same as the polypill other than it does not contain aspirin is not what we would like to call a placebo control. In another example, co-amilofruse is the generic name for a drug that contains two agents that are known to have positive effects on hypertension and oedema, namely amiloride and frusemide. If the 'placebo' control were identical to 'real' co-amilofruse apart from the fact that it was missing amiloride (but contained frusemide), then a trial involving a placebo control that contained frusemide might be successful at controlling for expectations, and measuring the effects of amiloride. Yet it would not be an adequate placebo control for co-amilofruse, because it contains a feature (frusemide) that is positively effective not via some expectational route. To test whether co-amilofruse was more effective than a placebo, a control treatment could contain *neither* amiloride nor frusemide. [8]

[8] These are not the only treatments for which it is difficult to construct adequate placebo controls. How, for example, would we design a 'placebo' control for exercise? That is, how could we make people expect to be doing exercise (and experience the effects of all other incidental features—whatever those turn out to be) without actually doing exercise (Howick 2011)? One might suggest we could hypnotize people to believe they are doing exercise. However hypnosis has its own effects (Lee et al. 2010). Thornier still is the example of electroconvulsive therapy (ECT), which is the electrical induction of seizures in patients. ECT

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

With this in mind, I can now argue that treatment with the Streitberger needle may not be a 'true' placebo control. This is because there is independent evidence that acupressure is effective for treating pain independently of the expectational effects of acupressure (Lee and Done 2004). Given that the Streitberger Needle (as well of course as real acupuncture) exerts pressure, this suggests that a sensible therapeutic theory—one that applies the criteria for classifying characteristic features as features that are effective and not due to expectation effects (as specified above in Sect. 5.1) would classify the exerted pressure as characteristic rather than incidental. To be sure the pressure exerted by real or Streitberger acupuncture needles could be less intense than the pressure exerted as part of real acupressure therapy. However we cannot rule out that even the less intense pressure is effective for treating pain in advance of further empirical studies. Moreover, it is argued that the acupuncture consultation (which is often much longer than a conventional consultation) should be classified as characteristic (Paterson and Dieppe 2005, p. 1203). There is certainly a robust body of evidence supporting the view that longer more empathetic consultations can have relevant positive effects when compared with other ('placebo') consultations (Hojat et al. 2011).

The debate about how to classify features of acupuncture could be decided more easily if there were an accepted therapeutic theory for acupuncture. But in fact there is no accepted (from a conventional point of view) therapeutic theory. Without an accepted therapeutic theory, such arguments (and therefore defending claims that a particular treatment is a 'placebo' control for acupuncture) are difficult to sustain. The point of the Streitberger needle example is simply to show that controlling for expectations, in some cases, is not sufficient.

The problem, therefore, with accepting the 'validity' of the Streitberger needle is the belief that controlling for expectations is sufficient for a treatment to count as a placebo control. While expectations about the effectiveness of a therapy need to be controlled for as an incidental feature, controlling for these expectations is arguably not sufficient. Treatment with the Streitberger needle controls for expectations but may do so at the cost of including some features such as acupressure and extensive consultations that could, in at least one reasonable interpretation, be best classified as characteristic.

5.2.2 Case study of vertebroplasty

Vertebroplasty involves making a small incision in someone's back then injecting bone glue (cement) into a vertebra that has been damaged. In a clinical trial researchers from Australia took 78 patients with spinal fractures of the kind that are often treated by vertebroplasty (Buchbinder et al. 2009). Half of them the real thing while the other half got placebo vertebroplasty, where surgeons cut the skin and touched the bone to

is used to treat patients suffering from major depressive disorder who have not responded to other forms of therapy. It is difficult to imagine an adequate 'placebo' control for ECT. Worse, even if we could design an adequate placebo control, as Blease points out, ECT has so many deleterious side-effects (confusion, memory loss, fatigue, headaches, and general cognitive impairment (Blease 2013a, b) that it could arguably be unethical.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

simulate the glue injection, but did not inject any cement. The sham procedure performs as well as the 'real' surgery. Other studies have confirmed these results (Miller et al. 2011). Worse, the cement glue used can leak (Martin et al. 2012), possibly causing more fractures (Sisodia 2013).

The failure of vertebroplasty to outperform sham vertebroplasty proves that *one of* the characteristic features of vertebroplasty, namely injecting cement into a vertebra—has no benefit. This suggests that the (expensive and common) procedure should be replaced by less expensive and less dangerous procedures. However it is also possible to conceive of 'sham' vertebroplasty as a nonplacebo. When the body senses a wound—as it does when surgeons make an incision, the body instigates what is called a 'wound healing cascade' (Sinno and Prakash 2013), which includes various processes including the activation of fibrin (a kind of endogenous glue), inflammation, and new tissue growth. These processes could hypothetically benefit the damaged vertebrae adjacent to the vertebrae. If we could classify these self-healing processes induced by the sham incision as characteristic, then the sham vertebroplasty may not be a placebo. Another possibility is that the stronger analgesic drugs used as part of the (real or sham) procedures do a better than usual job of reducing pain symptoms. This, in turn, allows the patient to freely move and engage in physical activity. Physical activity, in turn, has been shown to have (non-placebo) benefits for reducing symptoms of low back pain (van Middelkoop et al. 2011).

One might, of course, object that any possible effects of the incision (the wound healing cascade) or the surgical analgesic (leading to increased physical activity) are placebo effects because they result from endogenous healing processes. I accept this as a potentially reasonable objection, and my point here is merely to point out the possibility that treatment involving a sham incision may not merely have expectation effects, and to make the more general point that controlling for expectations is not a sufficient criteria for classifying a treatment as a placebo.

5.2.3 Word of caution about (possibly) mistaken placebo controls

It is important to establish two things about possible (but in my view mistaken) imputed implications of pointing out possible problems with evaluating the effects of treatments whose characteristic features are difficult to identify such as vertebroplasty and acupuncture (and exercise). First, the fact that the Streitberger needle or sham surgery *might* not count as legitimate placebo controls according to my proposed definition does not imply that the specific features under test are effective. The acupuncture and vertebroplasty placebo controlled trials clearly show that needle penetration adds very little to the benefits of acupuncture for pain, and injecting cement into a vertebra is not effective for fractures. Acupuncturists, vertebroplasty surgeons, and indeed practitioners of any discipline whose treatments fail to demonstrate superiority to a control treatment could be tempted to call the methodology of the trial into account. Like the failed carpenterwho blames his tools, these practitioners could maintain that their therapies are effective (non-placebos) but blame the randomized trial methodology. Yet the fact that the Streitberger needle and sham vertebroplasty might not be legitimate placebos does not mean that 'anything goes' or that interventions can be exempt from evaluation in rigorous trials. Instead, equally rigorous and tightly controlled ran-

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

domized trials that use non-placebo treatments as a controls can be employed. Such trials are common. For example a recent systematic review of acupuncture for pain found 5 trials (346 patients) that compared acupuncture with other (drug and non-drug) interventions. The trials found that acupuncture was as effective (in one of the trials) or more effective (in four trials) (Furlan et al. 2005). Similarly, a systematic review identified 5 randomized trials comparing vertebroplasty with usual care (conservative management); the review reported no statistically significant benefit of vertebroplasty for pain.[9]

5.3 Grünbaum's definition of placebo control is inadequate for not requiring the inclusion of all incidental feature effects

It does not follow from the fact that patients in the control arm in some trial were given a Grünbaumian generic placebo (as treatments that the investigator correctly believes to be generic placebos and which moreover, if they have effects at all, are harmless) that the trial isolates and measures the incremental benefits of the characteristic features. This is because, according to a strict interpretation of Grünbaum's scheme, a generic placebo need not be a treatment that replicates *all* the incidental features of a treatment process: it need only be a treatment without characteristic effects. But if some effective incidental features are not replicated in the control but in fact have an effect on outcome, then the trial would not determine 'the incremental remedial potency of the characteristic in t* but would determine the *combined* effects of the characteristic features of the trial treatment plus those of the missing incidental effects.

The example of 'active' placebos highlights Grünbaum's error. Tricyclic antidepressants have been shown to be more effective than placebo antidepressants in trials described as double-blind (Furukawa et al. 2003). However patients who enrol in such trials are ethically required to be informed about the likely effects and side effects of the experimental treatment. Patients who subsequently experience such side effects (in the case of tricyclic antidepressants a common one is dry mouth) subsequently could be 'unblinded' because they correctly believe and expect they are taking the 'real' drug as opposed to the placebo. The expectations could activate the neuronal reward mechanisms and cause some recovery from depression. Such (partial) recovery could be independent of any characteristic effects of the drugs. To further confound such a trial, patients who do not experience the side effects could then believe they are merely receiving the placebo and have neutral or negative expectations. This could, at least in principle, exacerbate their depression, at least relative to those with positive expectations.

To test whether these different expectations that arise due to 'unblinding' could influence the results, Moncrieff et al. compared results from standard 'placebo' con-

[9] Placebo controlled trials may have some advantages compared with head to head trials that compare one intervention with another, and vice-versa. Discussion of the debate of the relative methodological advantages of placebo compared with other standard treatment controls is beyond the scope of the current paper. Suffice it to say that each design has relative advantages and disadvantages, and that there is no widespread consensus about the absolute superiority of one method is superior to another. See Howick (2009a, b) for further discussion.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

trolled trials with results from what they called (rather unfortunately given that all placebos can be active) 'active placebo' controlled trials. 'Active' placebos are not only sensibly indistinguishable from the test treatment and lack its characteristic features, but also contain some ingredients that imitate some (in the ideal case, *all*) of the experimental treatment's side-effects (Moncrieff 2003; Moncrieff and Wessely 1998; Moncrieff et al. 2004). They found that the apparent characteristic benefit of antidepressant drugs is smaller in trials with 'active placebo' controls. The most plausible explanation for this phenomenon is that both participants and caregivers correctly identify 'inactive' placebos as placebos. This knowledge then leads lower expectations in the 'placebo' group about the likelihood of recovering.[10]

This all means then, that if a treatment is to be a placebo control in the sense of being optimally designed to detect the 'incremental effect' of the features deemed characteristic by the accepted therapeutic theory, it cannot simply be a Grünbaumian generic placebo. It must also have all the effects of the experimental treatment other than the effects of the characteristic features of the treatment on the target disorder so that it produces the incidental expectations effects this may require the use of 'active' placebos.[11] My revised definition of placebo controls takes this into account. This involves a shift in the description of placebo controls from incidental *features* to *effects* of incidental features. In practice, of course, the best way to ensure that all and only the effects of the incidental features are produced by the placebo control is to arrange for the placebo control to have the features.

5.4 Grünbaum allows harmful interventions to be classified as placebos

Since the only distinguishing feature of placebos, according to Grünbaum, is that it not contain any characteristic features that have *positive* effects on the target disorder, treatments whose characteristic features have *negative* effects on the target disorder count as generic placebos. This is directly at odds with ordinary usage. Imagine a therapeutic theory that classified deep scratching of the skin as the only characteristic feature in a treatment for haemophilia. This treatment would be classified as a placebo for treating haemophilia on Grünbaum's scheme. Similarly, treatments whose characteristic features have no effects on D but that have negative effects on other life processes are classified as placebos. This implies that, for example, therapy aimed at treating pain that did not do so but that caused blindness would be classified as a placebo.

Of course sometimes it can be a positive aspect of the analysis of some term that it challenge and correct ordinary usage. But there seems absolutely no advantage to doing that in this case. I will therefore introduce the term 'harmful intervention' to refer to treatments whose characteristic features have harmful effects on a target disorder

[10] Introducing 'active' placebos presents two new problems. First, it is ethically questionable to introduce harm to the control group in a controlled trial. Second (and this has usually gone unnoticed) when measuring incidence of side effects in clinical trials, comparisons between outcomes in treatment and control groups are made. But if the side effects were introduced to the control group, we would not expect any differences. This can lead to mistaken claims about the side effect profile of a new intervention.

[11] This insight about placebo controls has been suggested by (Howick 2011) and Turner (2012a, b).

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

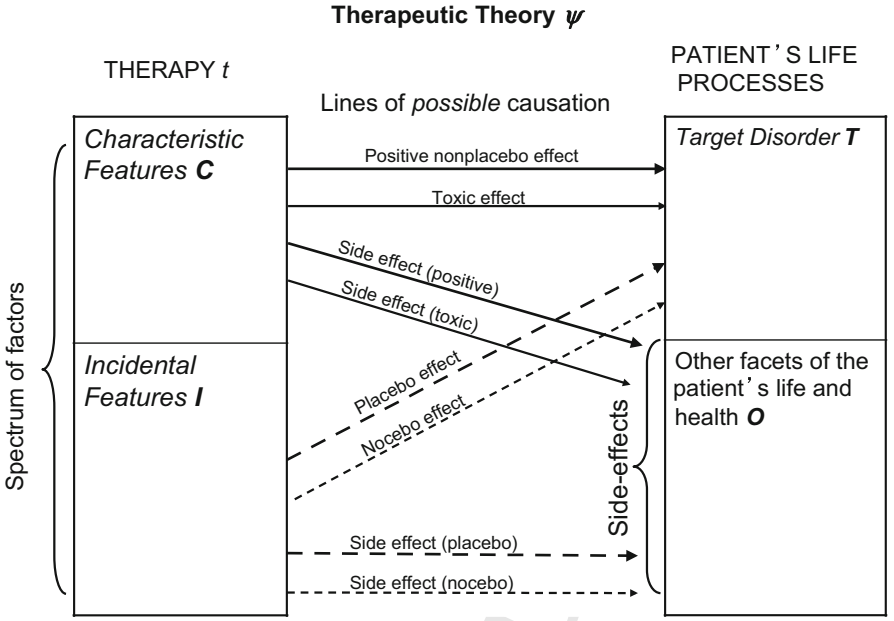


Fig. 3 Revised illustration of the therapeutic theory, used in clarifying definitions of 'placebo', nonplacebo, harmful intervention, placebo effects, and nocebo effects

or other life processes. And similarly use the term 'nocebo' (which is Latin for 'I shall harm') and 'nocebo effects' to refer to the negative effects of incidental features (See Fig. 3, below).

5.5 Waring's 'paradoxical effects' objection and the necessity of relativizing the definition of placebos to patients

Waring uses the example of drugs that elicit 'paradoxical responses' to argue that Grünbaum's scheme has the unreasonable consequence that the very same treatment can be classified both as a placebo and as a nonplacebo. This, he argues, illustrates a contradiction in Grünbaum's scheme. A paradoxical response is an exacerbating response on the target disorder produced by a drug that is normally remedial. He states:

> [C]onsider the newer generation of Selective Serotonin Reuptake Inhibitors (SSRIs). There is evidence that they might induce acutely anxious and even suicidal behaviour in certain patients suffering from anxiety and depression (Waring 2003, p. 12).

So, for example, although SSRIs may be effective for *most* patients suffering from depression, they allegedly cause a worsening of depressive symptoms in others, or so Waring argues. Waring's point is well known in pharmacology; Hauben and Aronson have identified no fewer than 60 drugs with paradoxical effects (Hauben and Aronson

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

2006). Waring contends calling paradoxical effects 'placebic' is a 'misuse of language' (Waring 2003, p. 12).

Importantly, a paradoxical effect is more than a negative side effect. Like the phenomena of hormesis it is a negative effect on the same disorder that the treatment sometimes cures. To use a 'toy' but dramatic and illustrative, example, swimming might be a wonderful treatment for obesity or rehabilitation, or general well being *but only for those patients who know how to swim*. Swimming could lead to death by drowning, a clear exacerbation of well being, for non-swimmers. Whether Prozac is an antidepressant, whether swimming improves health, and (more generally) whether a treatment (feature) is a placebo is relative to the patient. By necessity, then, the therapeutic theory must specify, in addition to which factors are incidental, which patients for which the treatment is a nonplacebo.[12]

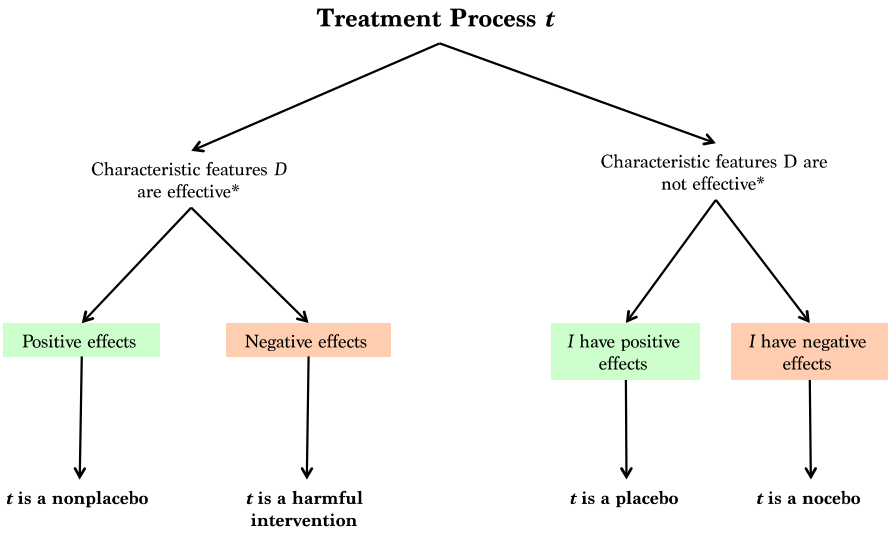
It is especially important to note the relativization to patients given that judgments about treatment effects are usually made based on average statistical differences between groups that receive experimental and control treatments. Average treatment benefits are compatible with great variation in treatment responses, including paradoxical responses (Howick 2011).

The same principle applies to whether a feature is considered harmful. Prozac supposedly has the side effect of causing sexual dysfunction in some men (which includes weakened sensation and difficulty maintaining an erection). This will generally be viewed as a negative feature. However by desensitizing relevant body parts, the very same side effect is beneficial for patients suffering from premature ejaculation (Arafa and Shamloul 2007). Likewise to some patients the possible side effect of gastro-intestinal bleeding after taking a non-steroidal anti-inflammatory drug (NSAID) might outweigh its analgesic benefits, but to an Olympic athlete in contention for a gold medal the side effect may be worth the risk. In short, whether a treatment feature, counts as beneficial or harmful (or the degree to which such a feature is viewed as beneficial or harmful) must also be relativized to an individual patient's physiology, values, and circumstances. A fortiori, whether a treatment process as a whole offers a net benefit will also be relative to an individual patient.

A careful reading of Grünbaum indicates that he presumed what counts as a placebo should be relativized to patients. When describing intentional placebos he makes explicit reference to particular 'victims': 'A treatment process t … will be said to be an 'intentional' placebo with respect to a target disorder D, suffered by a patient V and treated by a dispensing practitioner P' (1986, p. 24). Or later, when referring to both types of placebo (intentional and inadvertent), he states: 'Both explications are *relativized to disease victims of a specifiable sort*, as well as to therapists (practitioners) of certain kinds' (1986, p. 35, emphasis added). Yet it is fair to say that here too, Grünbaum's scheme did not adequately reflect his intentions. My revised definitional scheme therefore explicitly relativizes the definition of placebos to particular patients.

[12] It is, of course, problematic to determine in advance which patients will benefit, and which might be harmed by a treatment (although there are examples where genetic testing helps in this respect). This interesting epistemological problem, however, is orthogonal to my current ontological investigation into the nature of placebo controls.

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X



* In some cases a treatment will not have any effects on the target disorder. In these cases t can be classified as ineffective; there is no need to further classify it as a placebo or nonplacebo.

Fig. 4 Distinction between nonplacebo, harmful intervention, placebo, and nocebo, in relation to whether they are effective

5.6 The modified version of Grünbaum's scheme

The revised definitions take into account the problems with Grünbaum's scheme discussed above. It adds four lines of possible causation to the original (see Figs. 3 and 4), and introduces a definition of placebo controls that reflects Grünbaum's intentions.

Nonplacebo a treatment process t is a nonplacebo for target disease D, therapeutic theory ψ, and patients X *if (and only if) one or more of the characteristic factors do have a positive therapeutic effect on the target disease D*

Harmful intervention A treatment process t is a harmful intervention relative to a target disorder D, therapeutic theory ψ, and patients X if and only if (a) the characteristic features C do not have remedial effects on D and the characteristic features C have negative effects on the target disorder D or other life processes O.

Generic Placebo (revised) a treatment process t is a placebo when none of the characteristic treatment factors C are effective (remedial or harmful) in patients X for D.

Generic nocebo a treatment process t is a generic nocebo if it is a generic placebo whose incidental effects exacerbate the target disorder D in patients X or other life processes O.

Intentional placebo a treatment process t is an *intentional placebo* if and only if it satisfies the following four conditions—the fourth normally holding but, strictly speaking, being optional:

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

(a) *t* is a (revised) generic placebo
(b) to (d): (unchanged)

Inadvertent placebo (unchanged)

Placebo effect a placebo effect is either (a) a remedial effect produced by the incidental features of some treatment (even when the treatment as a whole is a nonplacebo), or (b) any effect of a (revised) generic placebo.

Nocebo effect a nocebo effect is either (a) a negative effect produced by the incidental features of some treatment (even when the treatment as a whole is a nonplacebo), or (b) any negative effect of a generic nocebo.

Placebo control (revised) A treatment functions as an adequate placebo control when it controls for *all* the effects of the experimental treatment *other than* the remedial effects of the characteristic features of the experimental treatment on the target disorder. Under conditions of informed consent, the placebo control must also mimic the sensory appearance of the experimental treatment in order to control for the effects of expectation *that* the treatment being given is (or in the case of a double blind trial) could be the experimental treatment.* This implies that the placebo control cannot contain any characteristic features that produce effects on the target disorder.
*Controlling for expectations is not sufficient, and in some exceptional cases—those in which the expectations in question arise from, for example, cognitive strategies taught by a therapist or coach—they are not necessary

Characteristic feature A characteristic feature is a feature which:

(1) is not expectancy *that* a treatment is effective, and
(2) that has an incremental benefit on the target disorder over a legitimate placebo control in a well controlled trial.

6 Conclusion and implications

Mistaken definitions of placebos have led to questionable estimates of placebo effects, unjustified 'placebo' control treatments, and confused debates about the ethics of placebos. Hróbjartsson and Gøtzsche's suggestion to accept any treatment labelled as a 'placebo' has unwanted consequences, and Nunn and Turner's suggestion to drop the term 'placebo' is only warranted if we can't define the placebo which I argued here is not the case. My modified version of Grünbaum's scheme captures what we mean by placebo controls and sheds light on complex cases such as that of acupuncture 'placebos' whereas other proposals leave us in the dark. Grünbaum's main insights are: (1) all treatments are complex and the features of interventions can be classified into 'characteristic' and 'incidental', and (2) what counts as a placebo is *relative* to a therapeutic theory, target disorder, and patient. The main problems with Grünbaum's scheme are that he fails to specify what he means by a therapeutic theory and because he does not specify that expectation effects are placebo effects. I showed that with four modifications, Grünbaum's definition provides a defensible account of placebos for the purpose of constructing placebo controls within clinical trials. The modifications

Journal: 11229-SYNT Article No.: 1001 TYPESET DISK LE CP Disp.:2016/1/13 Pages: 34 Layout: Small-X

I introduce are: adding a special role for expectations, insisting that placebo controls control for all and only the effects of the incidental treatment features, relativizing the definition of placebos to patients, and introducing harmful interventions and nocebos to the definitional scheme. I also provide guidance for classifying treatment features as characteristic or incidental. Future work is now warranted to investigate the implications of this definition for investigating the ethics of placebos in clinical practice and clinical trials, and to measure placebo effects more accurately.

Acknowledgments Special thanks to my Ph.D. supervisor John Worrall, with whom this paper was developed. Most of the ideas and much of the writing of the first two sections of the paper is John's. Thanks also to Nancy Cartwright and Michael Hardimon who both commented on earlier drafts of this paper, to Sir Iain Chalmers for insisting that I use real examples wherever possible, to the University of California at San Diego Graduate Conference attendees (organized by Michael Hardimon), to attendees of the "Evidence in Science and Epistemology" conference in Helsinki (especially Maria Lasonen-Aarnio, Petri Ylikoski, and Markus Lammenranta), to Have Carel and Alexander Bird for hosting me to give a talk about this paper at the University of Bristol and for providing useful comments. Andrew Turner, Frank Miller, and Wiebe Van Der Hoek gave useful feedback on earlier versions of the manuscript. Students taking part in the Oxford Philosophy of Medicine Seminar Series at *All Souls*, Oxford, including Matthew Clark, Meghana Mishra provided useful insights.

Funding information This work was partly supported by the National Institute for Health Research School for Primary Care Research (NIHR SPCR). The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health.

References

Arafa, M., & Shamloul, R. (2007). A randomized study examining the effect of 3 SSRI on premature ejaculation using a validated questionnaire. *Therapeutics and Clinical Risk Management*, *3*(4), 527–531.

Beecher, H. K. (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.

Benedetti, F., Pollo, A., et al. (2003). Conscious expectation and unconscious conditioning in analgesic, motor, and hormonal placebo/nocebo responses. *The Journal of Neuroscience*, *23*(10), 4315–4323.

Blease, C. (2013). Electroconvulsive therapy: The importance of informed consent and 'placebo literacy'. *Journal of Medical Ethics*, *39*(3), 173–174.

Blease, C. R. (2013). Electroconvulsive therapy, the placebo effect and informed consent. *Journal of Medical Ethics*, *39*(3), 166–170.

Bolier, L., Haverman, M., et al. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, *13*, 119.

Branthwaite, A., & Cooper, P. (1981). Analgesic effects of branding in treatment of headaches. *British Medicine Journal (Clinical Research Edition)*, *282*(6276), 1576–1578.

Buchbinder, R., Osborne, R. H., et al. (2009). A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *New England Journal of Medicine*, *361*(6), 557–568.

Di Blasi, Z., Harkness, E., et al. (2001). Influence of context effects on health outcomes: A systematic review. *The Lancet*, *357*(9258), 757–762.

Ernst, E. (2006). Acupuncture-a critical analysis. *Journal of Internal Medicine*, *259*(2), 125–37.

Foddy, B. (2009). A duty to deceive: Placebos in clinical practice. *The American Journal of Bioethics*, *9*(12), 4–12.

Furlan, A. D., van Tulder, M., et al. (2005). Acupuncture and dry-needling for low back pain: An updated systematic review within the framework of the cochrane collaboration. *Spine (Phila Pa 1976)*, *30*(8), 944–963.

Furukawa, T., McGuire, H., et al. (2003). Low dosage tricyclic antidepressants for depression. *Cochrane Database System Review*, *3*, CD003197.

Journal: **11229-SYNT** Article No.: **1001** ☐ TYPESET ☑ DISK ☐ LE ☐ CP Disp.:**2016/1/13** Pages: **34** Layout: **Small-X**

Golomb, B. (1995). Paradox of placebo effect. *Nature*, *375*(6532), 530.

Golomb, B. (2002). When Are Medication Side Effects Due to the Nocebo Phenomenon? Journal of the American Medical Association.

Golomb, B. (2009). Control Theory: Placebo-controlled drug trials have problems. Active-controlled drug trials are not always the solution. *American Journal of Bioethics*, *67*(69), 9–9.

Gøtzsche, P. (1994). Is there logic in the placebo? *Lancet*, *344*(8927), 925–6.

Greenwood, J. D. (1996). Freud's "Tally" argument, placebo control treatments, and the evaluation of psychotherapy. *Philosophy of Science*, *62*, 605–621.

Greenwood, J. D. (1997). Placebo control treatments and the evaluation of psychotherapy: A reply to Grunbaum and Erwin. *Philosophy of Science*, *64*(September), 497–510.

Grünbaum, A. (1981). The placebo concept. *Behaviour Research and Therapy*, *19*(2), 157–167.

Grünbaum, A. (1986). The placebo concept in medicine and psychiatry. *Psychological Medicine*, *16*(1), 19–38.

Hauben, M., & Aronson, J. K. (2006). Paradoxical reactions: Under-recognized adverse effects of drugs. *Drug Safety*, *29*(10), 970.

Hill, A. B. (1951). The clinical trial. *British Medical Bulletin*, *7*(4), 278–282.

Hoffmann, T. C., Erueti, C., et al. (2013). Poor description of non-pharmacological interventions: Analysis of consecutive sample of randomised trials. *British Medical Journal*, *347*, f3755.

Hoffmann, T. C., Glasziou, P. P., et al. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *British Medical Journal*, *348*, g1687.

Hojat, M., Louis, D. Z., et al. (2011). Physicians' empathy and clinical outcomes for diabetic patients. *Academic Medicine*, *86*(3), 359–364.

Howick, J. (2009). Questioning the methodologic superiority of 'placebo' over 'active' controlled trials. *The American Journal of Bioethics*, *9*(9), 34–48.

Howick, J. (2009). Reviewing the unsubstantiated claims for the methodological superiority of 'placebo' over 'active' controlled trials: reply to open peer commentaries. *The American Journal of Bioethics*, *9*(9), W5–7.

Howick, J. (2011). *The philosophy of evidence-based medicine*. Oxford: Wiley-Blackwell.

Howick, J., & Bishop, F. L. (2013). Placebo Use in the United Kingdom: Results from a national survey of primary care practitioners. *PLoS One*, *8*(3), e58247.

Howick, J., Friedemann, C., et al. (2013). Are treatments more effective than placebos? A systematic review and meta-analysis. *PLoS One*, *8*(5), e62599.

Hróbjartsson, A. (2002). What are the main methodological problems in the estimation of placebo effects? *Journal of Clinical Epidemiology*, *55*(5), 430–435.

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), 1594–1602.

Hróbjartsson, A., & Gøtzsche, P. (2010). Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews*, *1*, CD003974.

Kaptchuk, T. J. (1998). Intentional ignorance: A history of blind assessment and placebo controls in medicine. *Bulletin of the History of Medicine*, *72*(3), 389–433.

Kaptchuk, T. J. (2002). The placebo effect in alternative medicine: Can the performance of a healing ritual have clinical significance? *Annals of Internal Medicine*, *136*(11), 817–825.

Kaptchuk, T. J., Friedlander, E., et al. (2010). Placebos without deception: A randomized controlled trial in irritable bowel syndrome. *PLoS One*, *5*(12), e15591.

Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., et al. (2008). Components of placebo effect: Randomised controlled trial in patients with irritable bowel syndrome. *BMJ, 336*(7651), 999–1003.

Kirsch, I. (2002). Antidepressants and placebos: Secrets, revelations, and unanswered questions. *Prevention & Treatment*, *5*, 33.

Kirsch, I. (2005). Placebo psychotherapy: Synonym or oxymoron? Journal of Clinical Psychology.

Kirsch, I., Deacon, B. J., 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., & Weixel, L. J. (1988). Double-blind versus deceptive administration of a placebo. *Behavioral neuroscience*, *102*(2), 319–23.

Lee, A., & Done, M. L. (2004). Stimulation of the wrist acupuncture point P6 for preventing postoperative nausea and vomiting. *Cochrane Database System Review*, *3*, CD003281.

Journal: **11229-SYNT** Article No.: **1001** ☐ TYPESET ☑ DISK ☐ LE ☐ CP Disp.:**2016/1/13** Pages: **34** Layout: **Small-X**

Lee, R., Middleton, D., et al. (2010). A review of events that expose children to elemental mercury in the United States. *Ciencia & Saude Coletiva*, *15*(2), 585–598.

Linde, K., Niemann, K., et al. (2010). Are sham acupuncture interventions more effective than (other) placebos? A re-analysis of data from the Cochrane review on placebo effects. *Forschende Komplementarmedizin*, *17*(5), 259–264.

Mackie, J. L. (1974). *The cement of the universe*. Oxford: Clarendon Press.

Martin, D. J., Rad, A. E., et al. (2012). Prevalence of extravertebral cement leakage after vertebroplasty: Procedural documentation versus CT detection. *Acta Radiologica*, *53*(5), 569–572.

Miller, F. G., Kallmes, D. F., et al. (2011). Vertebroplasty and the placebo response. *Radiology*, *259*(3), 621–625.

Mitchall, J., Stanimirovic, R., et al. (2009). A randomised controlled trial of a self-guided internet intervention promoting well-being. *Computers in Human Behavior*, *25*(3), 749–760.

Moerman, D. E. (2002). *Meaning, medicine, and the "placebo effect"*. Cambridge: Cambridge University Press.

Moerman, D. E., & Jonas, W. B. (2002). Deconstructing the placebo effect and finding the meaning response. *Annals of Internal Medicine*, *136*(6), 471–476.

Moncrieff, J. (2003). A comparison of antidepressant trials using active and inert placebos. *International Journal of Methods in Psychiatric Research*, *12*(3), 117–127.

Moncrieff, J., & Wessely, S. (1998). Active placebos in antidepressant trials. *The British Journal of Psychiatry*, *173*, 88.

Moncrieff, J., Wessely, S., et al. (2004). Active placebos versus antidepressants for depression. *Cochrane Database System Review*, *1*, CD003012.

Nunn, R. (2009). It's time to put the placebo out of its misery. *British Medical Journal*, *338*, 1015.

Paterson, C., & Dieppe, P. (2005). Characteristic and incidental (placebo) effects in complex interventions such as acupuncture. *British Medical Journal*, *330*(7501), 1202–1205.

Phillips, D. P., Ruth, T. E., et al. (1993). Psychology and survival. *Lancet*, *342*(8880), 1142–1145.

Savovic, J., Jones, H. E., et al. (2012). Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Annals of Internal Medicine*, *157*(6), 429–438.

Schulz, K. F., Chalmers, I., Hayes, R. J., & Altman, D. G. (1995). Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA*, *273*(5), 408–412.

Shapiro, A., & Morris, L. A. (1978). The placebo effect in medical and psychological therapies. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavioural change: An empirical analysis* (pp. 369–410). New York: Wiley.

Simpson, S. H., Eurich, D. T., et al. (2006). A meta-analysis of the association between adherence to drug therapy and mortality. *British Medical Journal*, *333*(7557), 15.

Sinno, H., & Prakash, S. (2013). Complements and the wound healing cascade: An updated review. *Plastic Surgery International*, *2013*, 146764.

Sisodia, G. B. (2013). Methods of predicting vertebral body fractures of the lumbar spine. *World Journal of Orthopedics*, *4*(4), 241–247.

Stewart-Williams, S., & Podd, J. (2004). The placebo effect: dissolving the expectancy versus conditioning debate. *Psychological Bulletin*, *130*(2), 324–340.

Streitberger, K., & Kleinhenz, J. (1998). Introducing a placebo needle into acupuncture research. *Lancet*, *352*(9125), 364–365.

Turner, A. (2012). Evidence based medicine, placebos and the homeopathy controversy, University of Nottingham. Ph.D.: 294.

Turner, A. (2012). 'Placebos' and the logic of placebo comparison. *Biology & Philosophy*, *27*(3), 419–432.

van Middelkoop, M., Rubinstein, S. M., et al. (2011). A systematic review on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. *European Spine Journal*, *20*(1), 19–39.

Waber, R. L., Shiv, B., et al. (2008). Commercial features of placebo and therapeutic efficacy. *Journal of the American Medical Association*, *299*(9), 1016–1017.

Walach, H. (2011). Placebo controls: Historical, methodological and general aspects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *366*(1572), 1870–1878.

Waring, D. (2003). Paradoxical drug response and the placebo effect: A discussion of Grunbaum's definitional scheme. *Theoretical Medicine and Bioethics*, *24*(1), 5–17.

Journal: **11229-SYNT** Article No.: **1001** ☐ TYPESET ☑ DISK ☐ LE ☐ CP Disp.:**2016/1/13** Pages: **34** Layout: **Small-X**

Wood, L., Egger, M., et al. (2008). Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal*, *336*(7644), 601–605.

Worrall, J. (2002). What evidence in evidence-based medicine? *Philosophy of Science, 69*(Supplement), S316–S330.

Journal: **11229-SYNT** Article No.: **1001** ☐ TYPESET ☑ DISK ☐ LE ☐ CP Disp.:**2016/1/13** Pages: **34** Layout: **Small-X**